

293



# Universidad Nacional Autónoma de México

FACULTAD DE CIENCIAS

PRUEBAS DE HIPOTESIS EN LA FORMA ESTRUCTURAL DEL  
MODELO LINEAL MULTIECUACIONAL BASADAS EN ESTIMA-  
DORES DE MINIMOS CUADRADOS INDIRECTOS.

## TESIS PROFESIONAL

P r e s e n t a

ROBERTO ORTIZ SANCHEZ

Para obtener el Título de

A C T U A R I O



## **UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso**

### **DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

# I N D I C E

Pág.

Introducción .....	1
Capítulo I .....	5
1. Presentación del Modelo .....	6
2. Identificación .....	12
Capítulo II .....	21
1. El Método de Mínimos Cuadrados Indirectos .....	22
2. Estimador Unico de M.C.I. para Ecuaciones Sobreidenti- ficadas .....	30
Capítulo III .....	39
1. Pruebas de Identificabilidad: Aspectos Generales ....	40
2. El Método de Máxima Verosimilitud con Información Li- mitada .....	50
3. Pruebas de Hipótesis de la Condición de Rango para -- Identificación .....	61
4. Una Interpretación Alternativa sobre las Pruebas de - Identificabilidad .....	73
Capítulo IV .....	79
Pruebas de Hipótesis Basadas en el Estimador M.C.I.K.	80
Consideraciones Finales .....	84
Bibliografía .....	90

## INTRODUCCION

Los estudios econométricos abarcan hoy en día una gran variedad de problemas relacionados con los extensos campos de la teoría económica. Es difícil situar con exactitud los orígenes de una ciencia; sin embargo, podemos decir que la econometría, como un método de análisis del desarrollo económico, nació después de la Primera Guerra Mundial, apoyada por países occidentales dado el creciente poder del capitalismo. Entre sus primordiales objetivos se encontraba el pronosticar los ciclos económicos para evitar pérdidas a las empresas y al Estado derivadas de las crisis.

Actualmente las principales etapas en una investigación econométrica son, primero, la construcción o especificación de un modelo que se considera responsable de generar una serie de datos observados; segundo, la determinación de la identificabilidad de cada relación en el modelo; tercero, la inferencia estadística, que comprende la estimación de los parámetros del modelo y la contrastación o verificación de hipótesis relacionadas con dichos parámetros y, finalmente, la utilización de los resultados econométricos, que se vinculan generalmente con la toma de decisiones y el control. En este último caso, juegan un papel muy importante tanto el pronóstico, como el análisis de la estructura de un fenómeno económico para un período dado.

Es dentro del problema de la inferencia estadística en econometría que se encuentra inmersa la naturaleza del presente trabajo.

Bajo la perspectiva del modelo lineal clásico, las técnicas de estimación han sido suficientemente desarrolladas por la estadística matemática. El problema es diferente para el caso del modelo de ecuaciones simultáneas en donde los métodos de estimación están basados en su mayoría en teorías asintóticas, y de tal suerte que la elección de tal o cual método dependerá específicamente de las características del problema tratado.

Enmarcado dentro del modelo lineal general de ecuaciones simultáneas, este trabajo pretende hacer un análisis un tanto exhaustivo del método de mínimos cuadrados indirectos como una alternativa a los que proporcionan estimadores pertenecientes a la clase  $k$ . Si bien este método es poco utilizado por sus propias limitaciones, la interpretación dada para éstos en 1977 por D. Khazzoom, permite hacer un estudio más razonable sobre sus propiedades estadísticas, y señalarlo como una buena opción, sobre todo por su considerable sencillez, en el trabajo econométrico aplicado.

La discusión se presenta como sigue: en el primer capítulo se establecen las propiedades y definiciones que caracterizan el modelo lineal de ecuaciones simultáneas así como el problema de la identificación. En el capítulo II se presenta el método de mínimos cuadrados indirectos tanto en su forma usual como bajo la interpretación de Khazzoom. En el capítulo III se dará una especial importancia a las pruebas de identificabilidad, destacándose éstas como una fase preliminar en la investigación y en la práctica, y finalmente en el capítulo IV, de una manera muy general, se

mencionará el procedimiento de elaborar pruebas de hipótesis de ecuaciones sobreidentificadas utilizando las propiedades del estimador de Khazzoomm. Evidentemente el presente será sólo un trabajo informativo, pero que puede dar la pauta a investigaciones posteriores.

CAPITULO I

## 1. PRESENTACION DEL MODELO

Los métodos de regresión tradicionales en la inferencia estadística sufren de ciertas limitaciones cuando abordan los problemas relacionados con la teoría económica. Es bien sabido que cuando se trabaja bajo las hipótesis usuales en el modelo lineal clásico, el método de mínimos cuadrados ordinarios (M.C.O.) proporciona estimadores insesgados, consistentes y eficientes de los parámetros; sin embargo, desde que se introdujo el concepto de simultaneidad de las relaciones económicas (Haavelmo, 1943), y con esto los modelos de ecuaciones simultáneas, la aplicación directa de MCO queda muy restringida pues, al no satisfacer algunas de las hipótesis tradicionales, proporciona generalmente estimaciones inconsistentes de los parámetros estructurales. Por este motivo, dicho método sólo puede aplicarse a relaciones individuales de un modelo multiecuacional, cuando esta relación posea sólo una variable endógena o cuando el modelo sea recursivo. Dadas estas limitaciones, se han propuesto en la literatura varios métodos que son apropiados para estimar las ecuaciones estructurales de modelos multiecuacionales; de éstos, los más recurridos son: el método de mínimos cuadrados indirectos (M.C.I.), mínimos cuadrados bivariados (M.C.2.), mínimos cuadrados trivariados (M.C.3.), variables instrumentales (V.I.), máxima verosimilitud con información limitada (M.V.I.L.) y máxima verosimilitud con información completa (M.V.I.C.).

En realidad lo único que se ha hecho, han sido adaptaciones de los principios básicos de los mínimos cuadrados y de la máxima verosimilitud, para con esto superar las dificultades originadas por los modelos multiecuacionales. Cabe resaltar que no existe un criterio unánime que pueda reconocer las excelencias de un modelo frente a otro, y más aún, las ventajas e inconvenientes dependerán del trabajo específico que se esté realizando. Es por ello que no profundizaremos por el momento en los comentarios sobre los métodos antes mencionados y los aquí descritos.

Introduciremos ahora la notación y supuestos que caracterizan a un modelo lineal de ecuaciones simultáneas en su forma estructural. El modelo consta de;

- i) Identidades
- ii) Relaciones de equilibrio
- iii) Funciones de comportamiento

Las identidades son las definiciones que nos proporciona la teoría económica; por su parte las relaciones de equilibrio corresponden a ciertos supuestos sobre el comportamiento de determinados fenómenos económicos; por lo que, tanto identidades como relaciones de equilibrio tienen todos sus parámetros conocidos de antemano.

Las funciones de comportamiento expresan, como su nombre lo indica, algún aspecto del comportamiento de los agentes económi--

cos, ya sean individuales, sectoriales, de mercado, etc. Intervienen en su definición un conjunto de variables "relevantes" (endógenas y predeterminadas), un conjunto de parámetros desconocidos y otro de componentes estocásticos.

En lo que resta del trabajo haremos implícitamente una sustitución de las identidades y las relaciones de equilibrio en las funciones de comportamiento y nos ocuparemos exclusivamente de un modelo estático; por tanto las variables del sistema estarán clasificadas solamente en variables endógenas y variables exógenas.

Consideremos entonces cierta teoría económica que nos plante la existencia de un sistema de ecuaciones simultáneas de la forma

$$YB + X\Gamma = U \quad (1.1.1)$$

donde:

$Y$  es una matriz  $(T, G)$  de observaciones en las variables endógenas ( $T$  observaciones sobre las  $G$  variables endógenas)

$X$  es una matriz  $(T, K)$  de observaciones en las variables exógenas ( $T$  observaciones sobre las  $K$  variables exógenas)

$U$  es una matriz  $(T, G)$  de perturbaciones estocásticas (no observables)

$B$  es una matriz  $(G, G)$  de parámetros desconocidos correspondientes a las variables endógenas.

$\Gamma$  es una matriz  $(K, G)$  de parámetros desconocidos correspondientes a las variables exógenas.

La idea que se sigue del modelo es que, si dispusiéramos de una muestra infinita que nos proporcionara el conocimiento de los valores fijos para los parámetros poblacionales  $\beta$  y  $\Gamma$ , el sistema simultáneo de ecuaciones podría ser solucionado en forma única para las variables dependientes, en términos de las variables exógenas y las perturbaciones. Este criterio lo mantendremos para --- cualquier muestra finita, esto es, supondremos para garantizar la solubilidad del sistema que:

$$\beta \text{ es no singular} \quad (1.1.2)$$

Bajo este supuesto podemos entonces resolver o pasar del sistema en su forma estructural a su forma reducida.

$$Y\beta + X\Gamma = U \iff Y\beta = -X\Gamma + U$$

postmultiplicando por  $\beta^{-1}$  tenemos

$$Y = X\Pi + V \quad (1.1.3)$$

donde

$$\Pi = -\Gamma\beta^{-1} \quad (1.1.4)$$

$\Pi$  es la matriz de dimensión  $(K, G)$  de coeficientes de la forma reducida y

$$V = UB^{-1} \quad (1.1.5)$$

es la matriz de dimensión  $(T, G)$  de perturbaciones en la forma reducida.

### SUPUESTOS Y ESPECIFICACIONES ESTOCASTICAS

Denotando por  $U'_t$  al  $t$ -ésimo renglón de la matriz  $U$ ; los supuestos bajo los cuales opera el modelo son, además de (1.1.2):

- i)  $|X_{t\delta}| < \infty$ , para  $\infty$  constante finita;  $n(X) = K$ ,  $K < T$   
 ii)  $E(U_t) = 0$ ;  $Cov(U_t, U_\delta) = E(U_t U'_\delta) = \Sigma \delta_{t,\delta}$  (1.1.6)  
 donde  $\Sigma = (\sigma_{ij})$  es una matriz simétrica positiva definida

$$y \quad \delta_{t,\delta} = \begin{cases} 0, & t \neq \delta \\ 1, & t = \delta \end{cases}$$

posteriormente para hacer inferencia exacta supondremos que

$$U_t \sim NID(0, \Sigma)$$

- iii) Por (1.1.5) tendremos que  $V_t \sim NID(0, \Omega)$

$$\text{donde } \Omega = B^{-1} \Sigma B'^{-1}$$

Serán importantes también los siguientes supuestos sobre propiedades asintóticas

$$\begin{aligned}
 p\lim \frac{1}{T} X'X &= Q, \quad Q \text{ simétrica y positiva definida} \\
 p\lim \frac{1}{T} U'U &= \Sigma \\
 p\lim \frac{1}{T} X'U &= 0, \text{ y por tanto} && (1.1.7) \\
 p\lim \frac{1}{T} V'V &= \Omega \\
 p\lim \frac{1}{T} X'V &= 0
 \end{aligned}$$

El sistema de ecuaciones (1.1.1) y el anterior conjunto de supuestos es lo que se designa como un modelo lineal general en su forma estructural.

Como apuntamos anteriormente, si quisiéramos estimar los parámetros estructurales por M.C.O., los estimadores serían inconsistentes. Sin embargo, nótese que la forma reducida describe un modelo lineal clásico para el cual se obtienen estimaciones consistentes de sus parámetros utilizando M.C.O. Supóngase entonces que estimamos la forma reducida y utilizamos la relación (1.1.4) para con esta obtener los parámetros estructurales estimados. Este razonamiento (utilizado por el método de M.C.I.) para su ejecución, estará condicionado al concepto de identificación que como veremos es una propiedad que puede o no satisfacer el modelo.

## 2. IDENTIFICACION

Cabe resaltar que la identificación o identificabilidad es una -- propiedad del modelo y no del tratamiento estadístico que se le -- de a éste. Es por tanto el problema de identificación anterior -- al de estimación; y consiste en saber si un modelo es lo bastan-- te restrictivo como para poder determinar, en forma única, los va-- lores de sus parámetros estructurales. En particular, si el cono-- cimiento de la distribución poblacional de las observaciones nos permitiera deducir el valor de un parámetro, este estará identifi-- cado, de igual manera si todos los parámetros de una relación es-- tán identificados, dicha relación estará identificada, y si todas las relaciones del modelo lo están, el modelo será identificable.

Hemos visto la estrecha relación que existe entre un modelo en su forma estructural y en su forma reducida, y hemos estableci-- do de hecho el "puente" de la forma estructural a la reducida a -- través del sistema de parámetros (1.1.4). Poder cruzar ese puen-- te en ambos sentidos es garantizar la identificabilidad del mode-- lo. Insistiremos en este punto a continuación.

Es conveniente distinguir para nuestro análisis entre modelo y estructura. Por un modelo nos referimos a la forma específi-- ca de las relaciones que conectan a las variables en cuestión. -- Por una estructura (asociada al modelo) nos referiremos a la asig-- nación de valores numéricos específicos para los parámetros estruc-- turales  $\beta$ ,  $\Gamma$ , y  $\Sigma$ , así como a la distribución de las perturba---

ciones.

El problema de la identificación surge cuando tenemos estructuras "equivalentes en observación", o sea, estructuras que pueden tener las mismas implicaciones de un fenómeno observable bajo determinadas circunstancias. Pueden existir en este caso, varias estructuras que generen la misma forma reducida o que den pie a la observación del mismo conjunto de datos, y que asimismo puedan generar la misma distribución de las observaciones. Recordemos que todo modelo incorpora, de acuerdo a la teoría económica, cierto conocimiento a priori acerca del mecanismo que genera las observaciones; de esta manera si una estructura tiene suficientes restricciones (de acuerdo con dicho conocimiento a priori), el modelo estará identificado y esto se logrará cuando esa estructura y sólo esa, sea compatible a la vez con los datos y con la teoría económica. Lo que significa poder deducir de manera única los parámetros estructurales a partir de los parámetros reducidos. -- Nuestro propósito es entonces suprimir tantas hipótesis como sea posible mediante una selección preliminar, basándonos en que algunas de éstas son incompatibles con la teoría económica. La forma más usual de representar nuestro conocimiento a priori son las exclusiones de variables o "restricciones cero", que implican que ciertas variables no se encuentran en determinadas ecuaciones, es decir, que algunos coeficientes estructurales son cero.

Llamemos al conjunto de parámetros estructurales  $(\beta, \Gamma, \Sigma)$  la estructura del modelo y reescribamos lo anterior en forma más es-

quemática.

Dos estructuras serán equivalentes en observación (E.O.) si ambas determinan la misma distribución de las variables endógenas en el modelo. Con esto es fácil demostrar la siguiente proposición: dos estructuras  $\{B_1, \Gamma_1, \Sigma_1\}$  y  $\{B_2, \Gamma_2, \Sigma_2\}$  son E.O. si y sólo si existe una transformación  $R$  no singular que satisfice: --

$$B_2 = RB_1, \quad \Gamma_2 = R\Gamma_1, \quad \Sigma_2 = R\Sigma_1R'$$

La estructura  $\{B, \Gamma, \Sigma\}$  está identificada si y sólo si la única transformación admisible es  $R = I_G$ . Entendiéndose por una transformación admisible aquella para la cual la estructura  $\{B_2, \Gamma_2, \Sigma_2\}$  satisface las restricciones impuestas sobre  $\{B_1, \Gamma_1, \Sigma_1\}$ .

Veremos ahora un criterio alternativo que permitirá comprobar la identificabilidad de una estructura con un método analítico simple.

Partimos del modelo lineal general en su forma estructural.

$$YB + X\Gamma = U$$

o equivalentemente

$$ZA = U \tag{1.2.1}$$

donde

$$Z = (Y, X) \quad A = \begin{pmatrix} B \\ \Gamma \end{pmatrix}$$

$Z$  es una matriz  $(T, G + K)$  y  $A$  es una matriz  $(G + K, G)$ .

El sistema (1.2.1) lo podemos reescribir como

$$(I_G \otimes Z) \alpha = u \quad (1.2.2)$$

donde  $\alpha = \text{vec } A$ ,  $u = \text{vec } U$  y  $\otimes$  denota producto kronecker. Hagamos las siguientes consideraciones:

$$Z = (V, X) = (X\Pi + V, X) + X(\Pi, I_K) + (V, 0) = XW + V^* \quad (1.2.3)$$

donde  $W = (\Pi, I_K)$ ,  $V^* = (V, 0)$ ;

$W$  es  $(K, G + K)$  y  $\text{r}(W) = K$ ,  $V^*$  es  $(T, G + K)$

$E(Z | X) = E(XW + V^* | X) = XW$ , entonces

$$E(I_G \otimes Z) \alpha = (I_G \otimes XW) \alpha = (I_G \otimes X) (I_G \otimes W) \alpha = 0 \quad (1.2.4)$$

bajo el supuesto de que  $\text{r}(X) = K$ , se sigue que  $(I_G \otimes X)$  tiene todas sus columnas linealmente independientes; por lo que el sistema de ecuaciones (1.2.4.) tiene solución si y solo si

$$(I_G \otimes W) \alpha = 0 \quad (1.2.5)$$

(1.2.5) define un sistema lineal homogéneo en el cual  $\alpha \in \text{Ker}(I_G \otimes W)$  y  $\dim \text{Ker}(I_G \otimes W) = G(G+K) - GK = G^2$ , i.e.  $\alpha$  pertenece a un subespacio de dimensión  $G^2$ . Lo que deseamos es que  $\alpha$  esté determinada en forma única para satisfacer así la identificabilidad de nuestro modelo. Para ver si esto es posible, incorporaremos nuestro conocimiento a priori representado, por ejemplo, por una matriz  $\Phi$ , en

la forma de un sistema de restricciones lineales del tipo:

$$\phi \alpha = C \quad (1.2.6)$$

de tal forma que cada restricción impuesta a la forma estructural se traduce en una relación lineal entre los coeficientes de una misma ecuación. Supóngase que  $\phi$  incorpora  $m$  restricciones; se sigue entonces que  $\phi$  es  $(m, G(G+K))$ ,  $\alpha$  es  $(G(G+K), 1)$  y  $C$  es  $(m, 1)$ ; con esto reescribimos el sistema (1.2.5) incorporando (1.2.6) de la siguiente manera:

$$\Psi \alpha = C^* \quad (1.2.7)$$

donde  $\Psi = \begin{pmatrix} I_G \otimes W \\ \phi \end{pmatrix}$  y  $C^* = \begin{pmatrix} 0 \\ C \end{pmatrix}$

$\Psi$  es  $(GK + m, G(G+K))$  y  $C^*$  es  $(GK + m, 1)$ .

En el sistema de ecuaciones no homogéneo (1.2.7)  $\alpha$  estará determinada en forma única si:

$$\rho(\Psi) = G(G+K) \quad (1.2.8)$$

esta relación se denomina condición de rango para la indentificación y es una condición necesaria y suficiente para poder garantizarla. Nótese que de (1.2.7) y (1.2.8):

$$m + GK \geq G(G+K) = G^2 + GK \implies m \geq G^2 \quad (1.2.9)$$

esta relación se denomina condición de orden para la indentificación y es condición necesaria pero no suficiente para poder establecerla. Si  $m = G^2$  y se satisface (1.2.8) el modelo está exacta

mente identificado. Si  $m > G^2$  y se satisface (1.2.8) el modelo está sobreidentificado; y si  $m < G^2$  entonces (1.2.8) no se satisface y el modelo está subidentificado.

Veremos que cuando las restricciones existen sólo hacia el interior de las ecuaciones y no entre ecuaciones, es decir, cuando solamente consideramos restricciones de exclusión, podemos verificar la identificabilidad del modelo ecuación por ecuación; y este estará identificado si y sólo si, todas las ecuaciones están identificadas.

Hemos establecido hasta ahora dos criterios para probar si el modelo está identificado:

- i) La única transformación admisible es  $I_G$  (1.2.10)
- ii)  $\lambda(\psi) = G(G+K)$

Se sigue entonces que para el análisis ecuación por ecuación, la estructura para estas estará definida por el conjunto  $(B_{\bullet i}, \Gamma_{\bullet i}, \sigma_{ii})$ . Por i) en (1.2.10) se deduce fácilmente que la  $i$ -ésima ecuación estará identificada, si la  $i$ -ésima columna de todas las transformaciones admisibles es  $I_{\bullet i}$ .

Haciendo el desarrollo análogo que nos llevó a establecer ii) en (1.2.10) pero ahora ecuación por ecuación tendremos:

$$ZA_{\bullet i} = U_{\bullet i} \quad \text{donde} \quad A_{\bullet i} = \begin{pmatrix} B_{\bullet i} \\ \Gamma_{\bullet i} \end{pmatrix}$$

$Z$  es  $(T, G+K)$ ,  $A_{\circ i}$  es  $(G+K, 1)$  y  $U_{\circ i}$  es  $(T, 1)$

$$E(ZA_{\circ i}|X) = XWA_{\circ i} = 0$$

con  $W$  definida como en (1.2.3). Dado que  $\kappa(X) = K$ , este sistema tiene solución si y sólo si  $WA_{\circ i} = 0$  con  $A_{\circ i} \in \text{Ker}(W)$  y  $\dim \text{Ker}(W) = G+K-K = G$ . Para intentar determinar  $A_{\circ i}$  en forma única, incorporamos el conocimiento a priori en la forma:

$$\Phi_i A_{\circ i} = C_i$$

supongamos ahora que  $\Phi_i$  incorpora  $m_i$  restricciones, entonces

$$\Psi_i A_{\circ i} = C_i^* \quad \text{donde} \quad \Psi_i = \begin{pmatrix} W \\ \Phi_i \end{pmatrix} \text{ y } C_i^* = \begin{pmatrix} 0 \\ C_i \end{pmatrix} \quad (1.2.11)$$

$W$  es  $(K, G+K)$  y  $\Phi_i$  es  $(m_i, G+K)$ , por tanto para que  $A_{\circ i}$  esté determinado en forma única en (1.2.11) es necesario que

$$\kappa(\Psi_i) = G+K \quad (1.2.12)$$

que es la condición de rango para la  $i$ -ésima ecuación (necesaria y suficiente). Se sigue de (1.2.12) que

$$m_i + K \geq G+K \implies m_i \geq G \quad (1.2.13)$$

que es la condición de orden para la identificación (necesaria pero no suficiente). De tal forma que, si  $m_i = G$  y se satisface (1.2.12) la ecuación está exactamente identificada; si  $m_i > G$  y se satisface (1.2.12) la ecuación está sobreidentificada, y si  $m_i < G$  la ecuación está subidentificada.

Para cualquier ecuación en el modelo si multiplicamos esta - por algún escalar diferente de cero, su estructura permanece básicamente inalterada. Por lo que podemos asumir una regla de normalización: en cada ecuación estructural el coeficiente de una de - las variables dependientes, digamos  $\beta_{ii}$ , puede ser tomado a priori como  $\beta_{ii} = -1$ .

Supongamos entonces que  $m_i$  consta de una normalización y ---  $m_i - 1$  exclusiones, de esta forma

$$B_{\bullet i} = \begin{pmatrix} -1 \\ \beta_{\bullet i} \\ 0 \end{pmatrix} \quad \text{y} \quad \Gamma_{\bullet i} = \begin{pmatrix} \gamma_{\bullet i} \\ 0 \end{pmatrix}$$

donde  $\beta_{\bullet i}$  es  $(G_{1i}, 1)$ ; 0 correspondiente a  $B_{\bullet i}$  es  $(G_{2i}, 1)$ ,  $\gamma_{\bullet i}$  es  $(K_{1i}, 1)$  y 0 correspondiente a  $\Gamma_{\bullet i}$  es  $(K_{2i}, 1)$  con  $G = 1 + G_{1i} + G_{2i}$  - y  $K = K_{1i} + K_{2i}$ , esto es

$$m_i = 1 + G_{2i} + K_{2i} \quad (1.2.14)$$

sabemos por la condición de orden para identificación que

$$m_i > G \Rightarrow 1 + G_{2i} + K_{2i} > G \Rightarrow K_{2i} > G - G_{2i} - 1 = G_{1i}$$

$$\text{ie } K_{2i} > G_{1i} \quad (1.2.15)$$

lo que significa que para que una ecuación esté identificada, mínimamente se debe satisfacer que el número de variables exógenas excluidas en la ecuación sea mayor o igual al número de variables endógenas incluidas como explicativas, o bien, por (1.2.15)

$$K_{2,i} = K - K_{1,i} \geq G_{1,i} \Rightarrow K \geq G_{1,i} + K_{1,i} \quad (1.2.16)$$

esto es, el número de variables exógenas en el modelo tiene que ser mayor o igual al número de variables que aparecen en la ecuación.

Este sencillo procedimiento que nos permite comprobar si la condición de orden para identificación se satisface, y que se reduce solamente al conteo de variables en las ecuaciones, es el mecanismo más usado en la práctica. Sin embargo, el criterio fundamental de la identificación es la condición de rango, y este se podría verificar sólo si conociéramos los valores reales de los parámetros estructurales; lo que únicamente es posible aproximar -- (con ciertas medidas de probabilidad) después de la estimación. Será en el capítulo III cuando se mencionen bajo estas condiciones, algunos métodos para hacer pruebas de hipótesis sobre la condición de rango para identificabilidad, para con esto tener una medida estadística más confiable sobre este importante requisito que debe satisfacer el modelo.

En el próximo capítulo se presentará en forma detallada el método de mínimos cuadrados indirectos (M.C.I.) y se verá la estrecha relación que hay entre la existencia de los estimadores -- que pertenecen a esta categoría y el concepto de identificación.

## CAPITULO II

## 1. EL METODO DE MINIMOS CUADRADOS INDIRECTOS

El método de mínimos cuadrados indirectos (M.C.I.), como se señaló anteriormente, es una de las opciones con las que se aborda el problema de la estimación de parámetros estructurales en un modelo lineal de ecuaciones simultáneas. El método debe su nombre a que deduce las estimaciones de los parámetros estructurales a partir de la estimación de los parámetros de la forma reducida; es decir, no estima directamente, sino que recurre a una transformación con la cual reparametriza el problema dejando intacta la información muestral.

Citaremos primeramente de una manera informal, las condiciones que debe satisfacer el modelo para garantizar la existencia del estimador M.C.I., (ver Dhrymes 1970, p.p. 279-289). Con este fin, de aquí en adelante, para efectos prácticos y sin pérdida de generalidad, trabajaremos solamente con la  $i$ -ésima ecuación estructural dada por:

$$\forall B_{\bullet i} + \lambda \Gamma_{\bullet i} = u_{\bullet i}$$

donde  $B_{\bullet i}$ ,  $\Gamma_{\bullet i}$  y  $u_{\bullet i}$  son vectores de dimensión  $(G, 1)$ ,  $(K, 1)$ ,  $(T, 1)$  y corresponden a la  $i$ -ésima columna de las matrices  $B$ ,  $\Gamma$  y  $u$  respectivamente. En esta ecuación supondremos que  $B_{\bullet i}$  y  $\Gamma_{\bullet i}$  incorporan toda la información correspondiente a la exclusión o inclusión de determinadas variables para cada una de las ecuaciones, asumiendo también la regla de normalización definida como  $\beta_{ii} = -1$ .

De esta forma, para la primera ecuación por ejemplo, tendremos

$$B_{o1} = \begin{pmatrix} -1 \\ \beta_{o1} \\ 0 \end{pmatrix} \quad \Gamma_{o1} = \begin{pmatrix} \gamma_{o1} \\ 0 \end{pmatrix} \quad (2.1.1)$$

donde: -1 corresponde a la normalización,  $\beta_{o1}$  y  $\gamma_{o1}$  son los vectores de coeficientes desconocidos correspondientes a las  $G_1$  variables endógenas y  $K_1$  variables exógenas que aparecen en la ecuación estructural y son por ende de dimensión  $(G_1, 1)$  y  $(K_1, 1)$  respectivamente. El subvector cero de la izquierda es  $(G_2, 1)$  y el de la derecha es  $(K_2, 1)$  con  $G = G_1 + G_2 + 1$  y  $K = K_1 + K_2$ . Es decir estamos suponiendo por convención que la primera ecuación de estructura contiene a  $y_1, \dots, y_{G_1+1}$  como variables endógenas y a  $x_1, \dots, x_{K_1}$  como variables exógenas. Denotando por  $y_{o1}$  la variable cuyo coeficiente se ha igualado a -1 y poniendo dicha variable al otro lado de la igualdad, la ecuación podría también escribirse como:

$$y_{o1} = Y_1 \beta_{o1} + X_1 \gamma_{o1} + U_{o1} \quad (2.1.2)$$

donde  $y_{o1}$  es un vector  $(T, 1)$ ,  $Y_1$  es la matriz  $(T, G_1)$  de observaciones de las variables endógenas que aparecen en la ecuación estructural,  $X_1$  es la matriz  $(T, K_1)$  de observaciones de las variables exógenas que aparecen en la ecuación estructural y  $\beta_{o1}$ ,  $\gamma_{o1}$  y  $U_{o1}$  son tal como las definimos anteriormente. Las observaciones de las  $G_2$  variables endógenas y las  $K_2$  variables exógenas que aparecen en el modelo pero no en esta ecuación, estarán representadas por las matrices  $Y_2$  y  $X_2$  de dimensiones  $(T, G_2)$  y  $(T, K_2)$  -

respectivamente, es decir, hemos efectuado las particiones - - -  
 $X = [X_1, X_2]$  y  $Y = [y_{o1}, Y_1, Y_2]$ .

Sabemos que la relación entre parámetros estructurales y reducidos está dada por

$$\Pi = -\Gamma B^{-1} \quad (2.1.3)$$

por tanto, para la primera ecuación, el subconjunto de relaciones que nos interesa será

$$\Pi \begin{pmatrix} 1 \\ -\beta_{o1} \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma_{o1} \\ 0 \end{pmatrix} \quad (2.1.4)$$

haciendo una partición adecuada de la matriz  $\Pi$  tenemos

$$\Pi = \begin{bmatrix} \pi_{o1} & \Pi_{11} & \Pi_{12} \\ \pi_{o2} & \Pi_{21} & \Pi_{22} \end{bmatrix} \quad (2.1.5)$$

en donde  $\pi_{o1}$  es  $(K_1, I)$ ,  $\Pi_{11}$  es  $(K_1, G_1)$ ,  $\Pi_{12}$  es  $(K_1, G_2)$ ,  $\pi_{o2}$  es  $(K_2, I)$ ,  $\Pi_{21}$  es  $(K_2, G_1)$  y  $\Pi_{22}$  es  $(K_2, G_2)$ . De esta forma podemos reescribir el sistema (2.1.4) como

$$\begin{bmatrix} \pi_{o1} & \Pi_{11} & \Pi_{12} \\ \pi_{o2} & \Pi_{21} & \Pi_{22} \end{bmatrix} \begin{pmatrix} 1 \\ -\beta_{o1} \\ 0 \end{pmatrix} = \begin{pmatrix} \gamma_{o1} \\ 0 \end{pmatrix} \Rightarrow \quad (2.1.6)$$

$$\pi_{o1} - \Pi_{11} \beta_{o1} = \gamma_{o1} \Rightarrow \quad \Pi_{11} \beta_{o1} + \gamma_{o1} = \pi_{o1}$$

$$\pi_{o2} - \Pi_{21} \beta_{o1} = 0 \quad \Pi_{21} \beta_{o1} = \pi_{o2}$$

Como puede apreciarse se trata de un sistema recursivo de ecuaciones, por lo que (2.1.6) tendrá solución si y sólo si podemos resolver:

$$\Pi_{21} \beta_{01} = \pi_{02} \quad (2.1.7)$$

esto es, si podemos resolver (2.1.7) en forma única, entonces existiría un único estimador para  $\beta_{01}$  y por tanto para  $\gamma_{01}$ .  $\Pi_{21}$  es  $(K_2, G_1)$ , por tal motivo se tiene que una condición necesaria y suficiente para la existencia de estimadores M.C.I. es que (ver Dhrymes, ibid)

$$r(\Pi_{21}) = G_1 \quad (2.1.8)$$

Para el caso en que  $K_2 = G_1$  y se satisfaga (2.1.8),  $\Pi_{21}$  será no singular, con lo que podemos obtener un único estimador M.C.I. para  $\beta_{01}$  y  $\gamma_{01}$  en la forma:

$$\hat{\beta}_{01} = (\hat{\Pi}_{21})^{-1} \hat{\pi}_{02} \quad (2.1.9)$$

$$\hat{\gamma}_{01} = \hat{\pi}_{01} - \hat{\pi}_{11} (\hat{\Pi}_{21})^{-1} \hat{\pi}_{02}$$

si  $K_2 > G_1$  y se cumple (2.1.8), existe al menos una submatriz  $\Pi_{21}^*$  no singular de dimensión  $(G_1, G_1)$  en  $\Pi_{21}$ , de tal manera que suprimiendo en (2.1.7) todas las ecuaciones no incluidas en  $\Pi_{21}^*$ , obtenemos una solución explícita para  $\beta_{01}$  y  $\gamma_{01}$  en la forma

$$\begin{aligned} \hat{\beta}_{01} &= (\hat{\Pi}_{21}^*)^{-1} \hat{\pi}_{02}^* \\ \hat{\gamma}_{01} &= \hat{\pi}_{01} - \hat{\pi}_{11} (\hat{\Pi}_{21}^*)^{-1} \hat{\pi}_{02}^* \end{aligned} \quad (2.1.10)$$

donde  $\hat{\pi}_{21}^*$  es  $(G_1, G_1)$ ,  $\hat{\pi}_{22}^*$  es  $(G_1, 1)$  y  $\kappa(\hat{\pi}_{21}^*) = G_1$ . Como este resultado depende de las ecuaciones que decidamos suprimir, tenemos a lo más  $K_2/G_1 \cdot (K_2 - G_1)!$  distintos estimadores M.C.I.

Si  $\kappa(\pi_{21}) < G_1$ , entonces no existen estimadores de M.C.I. para  $\beta_{01}$  y  $\gamma_{01}$ , dado que la solución para estos no estaría sólo en términos de la muestra, sino también de parámetros desconocidos, a saber, los restantes  $G_1 - \kappa(\pi_{21})$  parámetros de  $\beta_{01}$ ; lo que estrictamente hablando, no es una solución estadística.

El planteamiento antes descrito sobre la existencia o no existencia de estimadores M.C.I., lleva implícito el concepto de identificación presentado en el capítulo I. De hecho si la ecuación en cuestión satisface (2.1.8), entonces se dice que la ecuación está identificada; en particular si  $K_2 = G_1$  la ecuación estará justamente identificada y si  $K_2 > G_1$  estará sobreidentificada. La condición que requiere que  $K_2 \geq G_1$  es la condición de orden y la que exige  $\kappa(\pi_{21}) = G_1$  es la condición de rango. Si  $\kappa(\pi_{21}) < G_1$ , la ecuación estará subidentificada o no identificada. Nótese que si  $K_2 < G_1$ , se tendrá que  $\kappa(\pi_{21}) < G_1$ . Para comprobar las anteriores afirmaciones, hagamos los siguientes planteamientos.

Sabemos por (1.2.12) del capítulo anterior que la condición de rango para identificación de la primera ecuación es

$$\kappa(\psi_1) = G + K$$

Es fácil comprobar el siguiente resultado que nos simplifica

rá el análisis:

$$\psi_1 = \begin{pmatrix} \omega \\ \phi_1 \end{pmatrix} = \begin{pmatrix} \Pi & I_K \\ \phi_{11} & \phi_{12} \end{pmatrix} = \begin{pmatrix} 0 & I_K \\ \phi_{1AB^{-1}} & \phi_{12} \end{pmatrix} \begin{pmatrix} I_G & 0 \\ \Pi & I_K \end{pmatrix} \quad (2.1.11)$$

donde

$A = \begin{pmatrix} B \\ \Gamma \end{pmatrix}$  y  $\phi_1 = (\phi_{11}, \phi_{12})$ ;  $\phi_{11}$  es  $(m_1, G + K)$ , así que  $\phi_{11}$  es  $(m_1, G)$  y  $\phi_{12}$  es  $(m_1, K)$ . Dado que  $\begin{pmatrix} I_G & 0 \\ \Pi & I_K \end{pmatrix}$  es de rango

completo, se sigue entonces:

$$\lambda(\psi_1) = \lambda \begin{pmatrix} 0 & I_K \\ \phi_{1AB^{-1}} & \phi_{12} \end{pmatrix} = K + \lambda(\phi_{1AB^{-1}}) = K + G \quad (2.1.12)$$

la última igualdad se da si se satisface la condición de rango, por tanto:

$$\lambda(\psi_1) = G + K \iff \lambda(\phi_{1AB^{-1}}) = G \quad (2.1.13)$$

Supongamos tal y como hemos hecho anteriormente, que las  $m_1$  restricciones consisten en  $G_2$  variables endógenas excluidas,  $K_2$  variables exógenas excluidas y una normalización. De esta forma escribamos

$$\phi_1 = \begin{pmatrix} \phi_{10} \\ \phi_1 \end{pmatrix} \quad (2.1.14)$$

donde el vector  $\phi_{10}$  de  $(1, G + K)$  corresponde a la normalización y la matriz  $\phi_1$  de  $(m_1 - 1, G + K)$  a las exclusiones, así:

$$\phi_{10} = I_{10} \quad \text{y} \quad \phi_1 = \begin{pmatrix} 0 & I_{G_2} & 0 & 0 \\ 0 & 0 & 0 & I_{K_2} \end{pmatrix}$$

nótese que

$$AB^{-1} = \begin{pmatrix} B \\ \Gamma \end{pmatrix} \quad B^{-1} = \begin{pmatrix} I_G \\ -\Pi \end{pmatrix} \quad (2.1.15)$$

Particionando  $\Pi$  como en (1.2.15) e  $I_G$  conformablemente, se tiene:

$$\begin{aligned} \phi_1 AB^{-1} &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & I_{G_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{K_2} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & I_{G_1} & 0 \\ 0 & 0 & I_{G_2} \\ -\pi_{01} & -\Pi_{11} & -\Pi_{12} \\ -\pi_{02} & -\Pi_{21} & -\Pi_{22} \end{pmatrix} = \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & I_{G_2} \\ -\pi_{02} & -\Pi_{21} & -\Pi_{22} \end{pmatrix} \quad (2.1.16) \end{aligned}$$

Hemos establecido entonces, que la primera ecuación estará identificada si y sólo si  $\lambda(\phi_1 AB^{-1}) = G$  y esto sucederá si y sólo si:

$$\lambda(\Pi_{21}) = G_1 \quad (2.1.17)$$

Esta afirmación se aprecia claramente por la última igualdad en (2.1.16), en donde la dimensión de la matriz del extremo derecho y de las submatrices que la conforman en la partición, nos dan esta respuesta. La condición (2.1.17) es utilizada en algunos textos para definir identificación; particularmente nosotros trabajaremos con dicha expresión en secciones posteriores.

Dhrymes demuestra que si la ecuación en cuestión es identificable, es decir, si se satisface (2.1.17), entonces el (o todos)-los estimadores de M.C.I. de los parámetros estructurales para, en este caso la primera ecuación estructural, son consistentes.

El método de M.C.I. por lo general, sólo se aplica a sistemas estructurales exactamente identificados, ya que para modelos sobreidentificados proporciona como hemos visto, soluciones múltiples. Sin embargo gran parte de su interés radica en que estima los parámetros estructurales solamente a través de la estimación de la forma reducida, y en que es un método que se presenta como alternativa a los estimadores clase-*k*.

Khazzoom (1976) propone un procedimiento basado en el uso de la inversa Moore-Penrose para obtener un único estimador M.C.I. - para el caso sobreidentificado. Deriva su consistencia y menciona la relación entre éste y el estimador de mínimos cuadrados bi-típicos (M.C.2) y variables instrumentales; exhibe su distribución asintótica y finalmente resume algunos experimentos Monte Carlo con los que prueba su comportamiento.

En la próxima sección se exploran básicamente, algunos de los resultados obtenidos por Khazzoom.

## 2. ESTIMADOR UNICO DE M.C.I. PARA ECUACIONES SOBREIDENTIFICADAS

Es conveniente como prerrequisito a esta sección, presentar las características principales de las inversas generalizadas y en particular de la inversa Moore-Penrose.

Una inversa generalizada o  $g$ -inversa de una matriz  $D$  de ----  
( $m, n$ ), se define como una matriz  $D^-$  de ( $n, m$ ) que satisface .

$$D D^- D = D \quad (2.2.1)$$

La matriz  $D^-$  no es única, de hecho existen un número infinito de matrices que satisface (2.2.1). Su importancia es debida principalmente a su aplicación a matrices singulares, tanto cuadradas como rectangulares. (Para demostraciones y discusión sobre su existencia y no unicidad, véase por ejemplo: Searle 1971. p.p.1-7)

La matriz inversa generalizada Moore-Penrose, se define como sigue: para cualquier matriz  $D$  de ( $m, n$ ) y rango  $r$ , existe una matriz única que denotaremos por  $D^+$ , y que satisface las siguientes cuatro condiciones:

- i)  $D D^+ D = D$
  - ii)  $D^+ D D^+ = D^+$
  - iii)  $(D D^+)' = D D^+$
  - iv)  $(D^+ D)' = D^+ D$
- (2.2.2)

donde  $D^+$  es de dimensión  $(n, m)$  y rango  $r$ . (Para la demostración de existencia y unicidad véase por ejemplo Rao y Mitra 1971 p.p. - 50-55).

Un caso especial de la inversa Moore-Penrose se tiene cuando  $D$  posee rango completo por columnas, esto es,  $D'D$  es no singular, y en este caso

$$D^+ = (D'D)^{-1}D' \quad (2.2.3)$$

las propiedades (2.2.2) pueden ser comprobadas directamente como sigue

- i)  $D D^+ D = D(D'D)^{-1}D'D = D$
- ii)  $D^+ D D^+ = (D'D)^{-1}D'D(D'D)^{-1}D' = (D'D)^{-1}D' = D^+$
- iii)  $D^+ D = (D'D)^{-1}D'D = I$  simétrica
- iv)  $D D^+ = D(D'D)^{-1}D'$  simétrica

Con esta información veremos ahora el estimador propuesto por Khazzoom.

En la sección anterior se hizo notar que la existencia del estimador M.C.I. dependía de la solución del siguiente sistema recursivo de ecuaciones: (deben satisfacerse para los valores reales de los parámetros, pero sólo podemos conocer sus estimaciones).

$$\begin{aligned} \hat{\pi}_{11} \hat{\beta}_{01} + \hat{\gamma}_{01} &= \hat{\pi}_{01} \\ \hat{\pi}_{21} \hat{\beta}_{01} &= \hat{\pi}_{02} \end{aligned} \quad (2.2.4)$$

que lo podemos escribir como

$$\begin{bmatrix} \hat{\Pi}_{11} & I_{K_1} \\ \hat{\Pi}_{21} & 0 \end{bmatrix} \begin{bmatrix} \hat{\beta}_{01} \\ \hat{\gamma}_{01} \end{bmatrix} = \begin{bmatrix} \hat{\pi}_{01} \\ \hat{\pi}_{02} \end{bmatrix} \quad (2.2.5)$$

en donde, si  $\hat{\Pi}_{21}$  es cuadrada y no singular, entonces (2.2.5) tiene una única solución que corresponde al caso justamente identificado; y si  $K_2 > G_1$  y  $\kappa(\hat{\Pi}_{21}) = G_1$  (caso sobreidentificado), hay más de una forma de estimar consistentemente los parámetros estructurales. Supongamos que se cumple este último caso. Sea

$$\hat{D} = \begin{bmatrix} \hat{\Pi}_{11} & I_{K_1} \\ \hat{\Pi}_{21} & 0 \end{bmatrix} ; \quad \hat{\delta}_{01} = \begin{pmatrix} \hat{\beta}_{01} \\ \hat{\gamma}_{01} \end{pmatrix} , \quad \hat{\pi}_0 = \begin{pmatrix} \hat{\pi}_{01} \\ \hat{\pi}_{02} \end{pmatrix}$$

con lo que reescribimos (2.2.5) como :

$$\hat{D} \hat{\delta}_{01} = \hat{\pi}_0 \quad (2.2.6)$$

$\hat{D}$  es  $(K, G_1 + K_1)$  y  $\kappa(\hat{D}) = G_1 + K_1$ ; premultiplicando (2.2.6) -- por  $\hat{D}'$  se tiene

$$\hat{D}' \hat{D} \hat{\delta}_{01} = \hat{D}' \hat{\pi}_0$$

donde  $(\hat{D}' \hat{D})$  es una matriz no singular, por tanto

$$\hat{\delta}_{01} = (\hat{D}' \hat{D})^{-1} \hat{D}' \hat{\pi}_0$$

y por (2.2.3)

$$\hat{\delta}_{01} = \hat{D}^+ \hat{\pi}_0 \quad (2.2.7)$$

el vector  $\hat{\delta}_{o1}$  es único en virtud de la unicidad de  $\mathcal{D}^+$ .

Recordando que el estimador de M.C.O. para  $\Pi$  es

$$\hat{\Pi} = (X'X)^{-1}X'Y$$

es posible obtener una expresión alternativa para  $\delta_{o1}$ .

Sea  $E = [I_K, 0]$ ;  $X'E' = [X_1, X_2]$   $\begin{bmatrix} I_K \\ 0 \end{bmatrix} = X_1$ ;  $Z = (Y, X)$   
entonces

$$\begin{aligned} \hat{\mathcal{D}} &= \begin{bmatrix} \hat{\Pi}_{11} & I_K \\ \hat{\Pi}_{21} & 0 \end{bmatrix} = [\hat{\Pi}_1 \ E'] = [(X'X)^{-1}X'Y_1 \quad (X'X)^{-1}X'XE'] \\ &= (X'X)^{-1}X'(Y_1, X_1) = (X'X)^{-1}X'Z_1 \end{aligned}$$

por otra parte  $\hat{\pi}_o = (X'X)^{-1}X'y_{o1}$ , de donde

$$\hat{\delta}_{o1} = (\hat{\mathcal{D}}'\hat{\mathcal{D}})^{-1}\hat{\mathcal{D}}'\hat{\pi}_o = [Z_1'X(X'X)^{-1}(X'X)^{-1}X'Z_1]^{-1}Z_1'X(X'X)^{-1}(X'X)^{-1}X'y_{o1}$$

$$\hat{\delta}_{o1} = [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'y_{o1} \quad (2.2.8)$$

sabemos por (2.1.2) que  $y_{o1} = Z_1\delta_{o1} + u_{o1}$ ; sustituyendo esta expresión en (2.2.8) se sigue

$$\hat{\delta}_{o1} = \delta_{o1} + [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'u_{o1}$$

Las expresiones (2.2.7), (2.2.8) y (2.2.9) corresponden al estimador de mínimos cuadrados indirectos propuesto por Khazzoom (lo identificaremos como "M.C.I.K").

Para examinar su consistencia tenemos que evaluar los límites en probabilidad de  $\hat{\delta}_{o1}$ , i.e.

$$p\lim_{T \rightarrow \infty} \hat{\delta}_{o1} = \delta_{o1} + p\lim_{T \rightarrow \infty} [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'U_{o1} \quad (2.2.10)$$

por el teorema de Klintchine es claro que

$$\begin{aligned} p\lim_{T \rightarrow \infty} [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'U_{o1} &= \\ [p\lim_{T \rightarrow \infty} \frac{1}{T} Z_1'X p\lim_{T \rightarrow \infty} (\frac{1}{T} X'X)^{-1} p\lim_{T \rightarrow \infty} (\frac{1}{T} X'X)^{-1} p\lim_{T \rightarrow \infty} \frac{1}{T} X'Z_1]^{-1} & \cdot \\ p\lim_{T \rightarrow \infty} \frac{1}{T} Z_1'X p\lim_{T \rightarrow \infty} (\frac{1}{T} X'X)^{-1} p\lim_{T \rightarrow \infty} (\frac{1}{T} X'X)^{-1} p\lim_{T \rightarrow \infty} \frac{1}{T} X'U_{o1} & \end{aligned} \quad (2.2.11)$$

en cuya expresión hemos supuesto (ver 1.1.7) que:

$$p\lim_{T \rightarrow \infty} \frac{1}{T} X'X = Q; Q \text{ matriz simétrica y positiva definida.}$$

Escribamos a  $X = (X_1, X_2)$  y a  $Y_1 = X\Pi_1 + V_1$ , entonces

$$p\lim_{T \rightarrow \infty} \frac{1}{T} X'X = p\lim_{T \rightarrow \infty} \frac{1}{T} \begin{bmatrix} X_1' X_1 & X_1' X_2 \\ X_2' X_1 & X_2' X_2 \end{bmatrix} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} = Q$$

con este argumento se sigue que

$$\begin{aligned} p\lim_{T \rightarrow \infty} \frac{1}{T} Z_1'X &= p\lim_{T \rightarrow \infty} \frac{1}{T} (V_1'X_1)'(X_1X_2) = p\lim_{T \rightarrow \infty} \frac{1}{T} \begin{bmatrix} V_1' X_1 & V_1' X_2 \\ X_1' X_1 & X_1' X_2 \end{bmatrix} \quad (2.2.12) \\ p\lim_{T \rightarrow \infty} \frac{1}{T} V_1'X_1 &= p\lim_{T \rightarrow \infty} \frac{1}{T} (X\Pi_1 + V_1)'X_1 = \Pi_1' \begin{pmatrix} Q_{11} \\ Q_{21} \end{pmatrix}, \end{aligned}$$

porque  $p\lim_{T \rightarrow \infty} \frac{1}{T} X_1' V_1 = 0$  y

$$p\lim_{T \rightarrow \infty} \frac{1}{T} V_1' X_2 = p\lim_{T \rightarrow \infty} (X \Pi_1 + V_1)' X_2 = \Pi_1' \begin{pmatrix} Q_{12} \\ Q_{22} \end{pmatrix}$$

porque  $p\lim_{T \rightarrow \infty} \frac{1}{T} X_2' V_1 = 0$ . Finalmente, dado que

$$p\lim_{T \rightarrow \infty} \frac{1}{T} X' U_{01} = 0 \quad (2.2.13)$$

se puede concluir entonces que todos los límites en probabilidad de (2.2.11) existen y son finitos, por tanto

$$p\lim_{T \rightarrow \infty} \hat{\delta}_{01} = \delta_{01}$$

lo cual muestra la consistencia del estimador M.C.I. de Khazoom, (M.C.I.K.).

Para la elaboración de pruebas de hipótesis, no basta saber que el estimador M.C.I.K. es consistente; se necesita tener conocimiento sobre la distribución del estimador, o por lo menos de su distribución asintótica. Desafortunadamente la distribución exacta del estimador M.C.I.K. para pequeña muestra se desconoce; sin embargo si es posible contar con su distribución límite. Cuando nos referimos a ésta, no hablamos directamente de la distribución de  $\hat{\delta}_{01}$ , ya que como hemos citado se trata de un estimador consistente, que converge por tanto en el límite a una constante. Para hablar de convergencia estocástica lo que se hace es aproximar la distribución de  $\hat{\delta}_{01}$  por medio de una transformación que le "reste velocidad de convergencia", es decir, el problema es determinar la distribución de  $\sqrt{T}(\hat{\delta}_{01} - \delta_{01})$  cuando  $T \rightarrow \infty$ . Con este criterio estableceremos la distribución asintótica del estimador M.C.-

I.K. para, en este caso, la primera ecuación estructural.

$$\begin{aligned}
 \text{plim } \sqrt{T} (\hat{\delta}_{01} - \delta_{01}) &= [\text{plim } \frac{1}{T} Z_1' X \text{plim } \left( \frac{X'X}{T} \right)^{-2} \text{plim } \frac{1}{T} X' Z_1]^{-1} \\
 &\cdot \text{plim } \frac{1}{T} Z_1' X \text{plim } \left( \frac{X'X}{T} \right)^{-2} \text{plim } \frac{1}{\sqrt{T}} X' u_{01} \quad (2.2.14)
 \end{aligned}$$

en donde por analogía con el desarrollo anterior, todos los límites existen y son finitos. Sabemos además que, bajo nuestros supuestos y por el teorema central del límite

$$\frac{1}{\sqrt{T}} X' u_{01} \overset{a}{\sim} N(0, \sigma_{11} Q) \quad (2.2.15)$$

por tanto

$$\sqrt{T} (\hat{\delta}_{01} - \delta_{01}) \overset{a}{\sim} N(0, \sigma_{11} \text{plim } \Psi) \quad (2.2.16)$$

donde

$$\begin{aligned}
 \Psi &= \left[ \left( \frac{Z_1' X}{T} \left( \frac{X'X}{T} \right)^{-2} \frac{X' Z_1}{T} \right)^{-1} \frac{Z_1' X}{T} \left( \frac{X'X}{T} \right)^{-2} \right] \frac{X' X}{T} \cdot \\
 &\quad \left[ \left( \frac{X'X}{T} \right)^{-2} \frac{X' Z_1}{T} \left( \frac{Z_1' X}{T} \left( \frac{X'X}{T} \right)^{-2} \frac{X' Z_1}{T} \right)^{-1} \right] \\
 &= T (Z_1' X' + X' Z_1)^{-1} Z_1' X' (X'X)^{-1} X' Z_1 (Z_1' X' + X' Z_1)^{-1} \quad (2.2.17) \\
 &= T (X' Z_1)' + X' X' + (X' Z_1)' + \dots
 \end{aligned}$$

El estimador propuesto por Khazzoom pertenece a la clase de estimadores definida por:

$$\hat{\delta}^* = (Z_1' C Z_1)^{-1} Z_1' C y$$

donde

$$C = \begin{cases} X(X'X)^{-1}X & \text{para M.C.I.K.} \\ (1-k)I + kX(X'X)^{-1}X' & \text{para clase - k} \end{cases}$$

Para terminar de enmarcar el problema de la inferencia estadística para el estimador M.C.I.K., es necesario contar con un estimador consistente de la varianza. El camino más usual de definir este es en la forma:

$$\hat{\sigma}_{11} = \frac{1}{T} \hat{u}'_{o1} \hat{u}_{o1}$$

donde

$$\hat{u}_{o1} = y_{o1} - Z_1 \hat{\delta}_{o1} = u_{o1} - Z_1 (\hat{\delta}_{o1} - \delta_{o1})$$

por tanto

$$\hat{\sigma}_{11} = \frac{1}{T} u'_{o1} u_{o1} - 2 \frac{1}{T} u'_{o1} Z_1 (\hat{\delta}_{o1} - \delta_{o1}) + (\hat{\delta}_{o1} - \delta_{o1})' \frac{1}{T} Z'_1 Z_1 (\hat{\delta}_{o1} - \delta_{o1})$$

utilizando los supuestos y especificaciones estocásticas del modelo y dado que  $plim \hat{\delta}_{o1} = \delta_{o1}$  se sigue

$$plim_{T \rightarrow \infty} \hat{\sigma}_{11} = plim_{T \rightarrow \infty} \frac{1}{T} u'_{o1} u_{o1} = \sigma_{11}$$

con lo cual se establece la consistencia de  $\hat{\sigma}_{11}$ .

Hemos resumido hasta aquí la forma y propiedades del estimador de M.C.I. tanto para el caso justamente identificado como para el caso sobreidentificado. La forma específica de elaborar pruebas de hipótesis utilizando el estimador M.C.I.K., es decir,

cuando el modelo está sobreidentificado se dejará para el capítulo IV. Sin embargo recuérdese que si el modelo no estuviera identificado no existirían estimadores de M.C.I. Es por tal motivo - que en el próximo capítulo nos ocuparemos de ese aspecto que se - deja generalmente fuera de consideración en el trabajo econométrico aplicado y que se refiere a la verificación bajo la teoría de las pruebas de hipótesis estadísticas de si la condición de rango para identificación se satisface.

CAPITULO III

## 1. PRUEBAS DE IDENTIFICABILIDAD : ASPECTOS GENERALES

El desarrollo de la teoría econométrica, se apoya generalmente en una serie de supuestos que permiten la construcción y manejo de -- piezas de teoría formal. Así por ejemplo en la elaboración de métodos tanto de estimación, evaluación y pronóstico de modelos estructurales (todos ellos posteriores a la especificación), se trabaja bajo el supuesto de que el modelo está identificado. En la práctica, este supuesto se basa generalmente en pensar en la condición de orden para identificación como si tal fuera suficiente para poder establecerla (aunque estrictamente no lo sea) y si este se cumple, la condición de rango se da por satisfecha. Conviene recordar que la identificación depende de los valores reales de -- los parámetros estructurales, los cuales son desconocidos para el investigador; pero también conviene cuestionarse sobre cuantas veces el trabajo econométrico aplicado ha carecido de validez por -- trabajar quizá con estructuras equivalentes en observación. Por tal motivo, suponer que un modelo posee tales o cuales propiedades es algo que debe de tomarse con ciertas reservas.

Para modelos en los que el grado de sobreidentificación impuesto por la condición de orden es "grande", resulta razonable el pensar que la probabilidad de que estas ecuaciones estén realmente subidentificadas es muy baja. Para resolver esta incertidumbre con respecto a si la especificación del modelo ha sido adecuada, algunos au-

tores han sugerido elaborar pruebas de hipótesis estadísticas que nos permitan constatar, de una manera más adecuada, si la condición de rango es satisfecha o no. "... es natural aplicar pruebas de identificabilidad antes de proceder con el cálculo de las estimaciones de las varianzas muestrales, las cuales requieren un mayor esfuerzo computacional, y abandonar los cálculos posteriores si la indicación de no identificabilidad es fuerte". (Koopmans y Hood 1953, p.p. 184).

Las pruebas a las que nos referimos corresponden a los trabajos originalmente realizados por Anderson y Rubin (1949, 1950) y por Koopmans y Hood (1953). En estos trabajos la condición de rango para identificación es probada ecuación por ecuación al no existir un método accesible que nos permita probarlas simultáneamente.

Retomando la notación utilizada en los capítulos anteriores, y esperando no ser muy repetitivo, recordaremos que la primera ecuación del modelo

$$Y B + X\Gamma = U \quad (3.1.1)$$

está dada por

$$Y B_{01} + X\Gamma_{01} = U_{01} \quad (3.1.2)$$

o bien, si consideramos explícitamente la normalización, por

$$y_{01} = Y_1 \beta_{01} + X_1 \gamma_{01} + u_{01} \quad (3.1.3)$$

Será conveniente no imponer por ahora esta regla de normalización

en nuestra ecuación, esto es, ignorar que el coeficiente de  $y_{01}$  es  $-1$ . De esta forma la ecuación (3.1.3) la escribiremos como:

$$[y_{01} \quad v_1] \begin{pmatrix} \beta_{01} \\ \beta_{02} \end{pmatrix} + X_1 \gamma_{01} = V_1^* \beta_{01}^* + X_1 \gamma_{01} = u_{01} \quad (1.3.4)$$

es decir, estamos suponiendo sin pérdida de generalidad que

$$B_{01} = \begin{pmatrix} \beta_{01}^* \\ \beta_{02} \end{pmatrix} \quad \Gamma_{01} = \begin{pmatrix} \gamma_{01} \\ \gamma_{02} \end{pmatrix} \quad (3.1.5)$$

donde  $\beta_{01}^* = \begin{pmatrix} \beta_{01} \\ \beta_{01} \end{pmatrix}$  es el subvector de  $B_{01}$  que consiste en sus primeros  $G_1 + 1$  elementos y  $\gamma_{01}$  es el subvector de  $\Gamma_{01}$  que consiste en sus primeros  $K_1$  elementos. Estamos suponiendo además que las variables excluidas (restricciones cero) están asociadas a los parámetros

$$\begin{pmatrix} \beta_{02} \\ \gamma_{02} \end{pmatrix} = 0 \quad (3.1.6)$$

donde  $\beta_{02}$  es  $(G_2, 1)$  y  $\gamma_{02}$  es  $(K_2, 1)$  con  $G = 1 + G_1 + G_2$  y  $K = K_1 + K_2$ .

Dado que  $B$  es no singular podemos resolver como sabemos la forma estructural (3.1.1) en la forma reducida:

$$Y = X\Pi + V \quad (3.1.7)$$

para la que particionando conformablemente se tiene

$$(V_1^* \ V_2) = (X_1 \ X_2) \begin{pmatrix} \Pi_{11}^* & \Pi_{12} \\ \Pi_{21}^* & \Pi_{22} \end{pmatrix} + (V_1 \ V_2) \quad (3.1.8)$$

donde  $V_1^*$  es  $(T, G_1 + 1)$ ,  $V_2$  es  $(T, G_2)$ ,  $X_1$  es  $(T, K_1)$ ,  $X_2$  es  $(T, K_2)$ ,  $\Pi_{11}^*$  es  $(K_1, G_1 + 1)$ ,  $\Pi_{12}$  es  $(K_1, G_2)$ ,  $\Pi_{21}^*$  es  $(K_2, G_1 + 1)$ ,  $\Pi_{22}$  es  $(T, G_1 + 1)$  y  $V_2$  es  $(T, G_2)$ . Postmultiplicando (3.1.8) por  $(\beta_{\theta 1}^* \ \beta_{\theta 2}^*)$  obtenemos el sistema de parámetros

$$\begin{pmatrix} \Pi_{11}^* & \Pi_{12} \\ \Pi_{21}^* & \Pi_{22} \end{pmatrix} \begin{pmatrix} \beta_{\theta 1}^* \\ \beta_{\theta 2}^* \end{pmatrix} = \begin{pmatrix} \gamma_{\theta 1} \\ \gamma_{\theta 2} \end{pmatrix}$$

que por (3.1.6) nos define el sistema recursivo de ecuaciones:

$$\begin{aligned} \Pi_{11}^* \beta_{\theta 1}^* &= \gamma_{\theta 1} \\ \Pi_{21}^* \beta_{\theta 1}^* &= 0 \end{aligned} \quad (3.1.9)$$

la matriz  $\Pi_{21}^*$  es  $(K_2, G_1 + 1)$  por lo que el sistema estará identificado (sujeto a una regla de normalización) si y sólo si

$$\lambda(\Pi_{21}^*) = G_1 \quad (3.1.10)$$

estamos por supuesto también suponiendo que  $K_2 > G_1$ , para que la ecuación estructural pueda estar identificada.

La matriz  $\Pi_{21}^*$  es una matriz rectangular (salvo en el caso justamente identificado); para la que se cumple

$$\lambda(\Pi_{21}^*) = \lambda(\Pi_{21}^* R \Pi_{21}^*) \quad (3.1.11)$$

siendo  $R$  una matriz de  $(K_2, K_2)$  positiva definida. De tal suerte - que como  $\Pi_{21}^{i*} R \Pi_{21}^*$  es una matriz cuadrada, y ya que estamos interesados en conocer su rango, podemos entonces preguntarnos sobre las raíces de la ecuación polinomial:

$$|\Pi_{21}^{i*} R \Pi_{21}^* - \eta H| = 0 \quad (3.1.12)$$

donde  $H$  será una matriz real positiva definida.

En virtud de que la matriz  $R$  es positiva definida y por él se gundo bloque de ecuaciones en (3.1.9), se sigue que la matriz ----  $\Pi_{21}^{i*} R \Pi_{21}^*$  es positiva semidefinida, por lo que las raíces características de esta matriz en la métrica inducida por  $H$  serán no negativas. Es claro que si el rango de  $\Pi_{21}^*$  es  $q$ ,  $q \leq G_1 - 1$ , entonces existirán  $q$  raíces de (3.1.12) diferentes de cero. Este argumento sería aplicable a nuestro problema si la matriz  $\Pi_{21}^*$  fuera conocida; sin embargo debemos recordar que los valores reales de los elementos de esta matriz (salvo las restricciones) son desconocidos, y -- por tanto también las raíces de (3.1.12). Por lo que nuestro problema se reduce entonces a crear una estadística de prueba que nos permita constatar si la expresión (3.1.10) es válida bajo ciertas - medidas de probabilidad. Para eso necesitamos las matrices  $R$  y  $H$  - definidas de tal forma específica que nos permitan construir dichas pruebas.

La parte de la forma reducida que nos interesa es aquella que se refiere a las variables endógenas incluidas (en este caso) en la

primera ecuación estructural:

$$\begin{aligned} Y_1^* &= X\Pi_1^* + V_1 = (X_1 \ X_2) \begin{pmatrix} \Pi_{11}^* \\ \Pi_{21}^* \end{pmatrix} + V_1 \\ &= X_1 \Pi_{11}^* + X_2 \Pi_{21}^* + V_1 \end{aligned} \quad (3.1.13)$$

donde  $V_1 \sim N(0, \Omega_{11})$ ,  $\Omega_{11}$  es una submatriz de  $\Omega$  no singular. El estimador M.C.O. de  $\Pi_1^*$  está dado por

$$\hat{\Pi}_1^* = (X'X)^{-1}X'Y_1^* \quad (3.1.14)$$

y el de  $\Omega_{11}$  por

$$\hat{\Omega}_{11} = \frac{1}{T} W_{11} \quad (3.1.15)$$

donde

$$\begin{aligned} W_{11} &= (Y_1^* - X\hat{\Pi}_1^*)' (Y_1^* - X\hat{\Pi}_1^*) \\ &= Y_1^{*'} (I - X(X'X)^{-1}X') Y_1^* = Y_1^{*'} M Y_1^* \end{aligned} \quad (3.1.16)$$

i.e.,  $W_{11}$  es la matriz de productos cruzados de residuos del ajuste por M.C.O. de  $Y_1^*$  sobre  $X$ , donde:

$$M = I_T - N; \quad N = X(X'X)^{-1}X' \quad (3.1.17)$$

Se puede verificar fácilmente que  $N$  y  $M$  son simétricas, idempotentes y con rangos iguales a sus trazas.

Consideremos ahora la forma reducida (3.1.13) y transforme--

mosla premultiplicando ambos lados de la igualdad por -----

$$M_1 = (I_T - X_1 (X_1' X_1)^{-1} X_1) :$$

$$M_1 V_1^* = M_1 X_1 \Pi_{11}^* + M_1 X_2 \Pi_{21}^* + M_1 V_1 \quad (3.1.18)$$

nótese que  $M_1 X_1 = 0$ , por lo que

$$M_1 V_1^* = M_1 X_2 \Pi_{21}^* + V_1^* \quad (3.1.19)$$

donde  $V_1^* = M_1 V_1$ .

La forma reducida (3.1.19) se puede interpretar como la explicación de la variabilidad de las  $V_1^*$  variables dependientes a partir de las  $X_2$  variables exógenas, pero permaneciendo  $X_1$  incluida en el modelo; y esto se logra al transformar el modelo reducido (3.1.13) por un conjunto de variables ortogonales a  $X_1$ . Ignorando que los errores han sido transformados, podemos obtener directamente el estimador M.C.O. de  $\Pi_{21}^*$  en la forma

$$\Pi_{21}^* = (X_2' M_1 X_2)^{-1} X_2' M_1 V_1^* \quad (3.1.20)$$

Sea  $\omega_{11}^*$  la matriz de productos cruzados de residuos de la regresión de  $V_1^*$  sobre  $X_1$ , i.e

$$\begin{aligned} \omega_{11}^* &= (V_1^* - X_1 \Pi_{11}^*)' (V_1^* - X_1 \Pi_{11}^*) \\ &= V_1^{*'} (I_T - X_1 (X_1' X_1)^{-1} X_1) V_1^* = V_1^{*'} M_1 V_1^* \end{aligned} \quad (3.1.21)$$

por (3.1.16) y (3.1.21) tenemos que:

$$\begin{aligned} \omega_{11}^* - \omega_{11} &= Y_1^* M_1 Y_1^* - Y_1^* M Y_1^* = Y_1^* (M_1 - M) Y_1^* \\ &= Y_1^* (N - N_1) Y_1^* \end{aligned} \quad (3.1.22)$$

Como una consecuencia directa de la inversión particionada de matrices (ver por ejemplo Goldberger 1964, p.p. 27, 28) obtenemos el siguiente resultado

$$\begin{aligned} N &= X(X'X)^{-1}X' = \\ &= (X_1 X_2) \begin{pmatrix} (X_1'X_1)^{-1} + (X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & -(X_1'X_1)^{-1}X_1'X_2(X_2'M_1X_2)^{-1} \\ -(X_2'M_1X_2)^{-1}X_2'X_1(X_1'X_1)^{-1} & (X_2'M_1X_2)^{-1} \end{pmatrix} \\ &\quad \cdot \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \end{aligned}$$

de donde

$$\begin{aligned} N &= N_1 + N_1 X_2 (X_2' M_1 X_2)^{-1} X_2' N_1 - N_1 X_2 (X_2' M_1 X_2)^{-1} X_2' \\ &\quad - X_2 (X_2' M_1 X_2)^{-1} X_2' N_1 + X_2 (X_2' N_1 X_2)^{-1} X_2' \\ &= N_1 + M_1 X_2 (X_2' M_1 X_2)^{-1} X_2' M_1 \\ \Rightarrow N - N_1 &= M_1 X_2 (X_2' M_1 X_2)^{-1} X_2' M_1 \end{aligned} \quad (3.1.23)$$

por lo que sutituyendo en (3.1.22) se sigue que:

$$\begin{aligned} \omega_{11}^* - \omega_{11} &= Y_1^* M_1 X_2 (X_2' M_1 X_2)^{-1} X_2' M_1 Y_1^* = \\ &= Y_1^* M_1 X_2 (X_2' M_1 X_2)^{-1} (X_2' M_1 X_2) (X_2' M_1 X_2)^{-1} X_2' M_1 Y_1^* \end{aligned}$$

y por (3.1.20)

$$\omega_{11}^* - \omega_{11} = \hat{\pi}_{21}^{1*} X_2^* M_1 X_2 \pi_{21}^* = \hat{\pi}_{21}^{1*} C \hat{\pi}_{21}^* \quad (3.1.24)$$

donde  $C = X_2^* M_1 X_2$ . Nótese que  $C$  es la matriz de productos cruzados de residuos de la regresión de  $X_2$  sobre  $X_1$ . Es por tanto  $C$  una matriz positiva definida y además  $\text{plim}_{T \rightarrow \infty} \frac{1}{T} C = \bar{C}$  existe, siendo  $\bar{C}$  también una matriz positiva definida.

Considérese entonces la ecuación polinomial

$$|\omega_{11}^* - \ell \omega_{11}| = 0 \quad (3.1.25)$$

y hagamos una pequeña transformación:

$$\begin{aligned} |\omega_{11}^* - \ell \omega_{11}| &= |\omega_{11}^* - \omega_{11} - \ell \omega_{11} + \omega_{11}| = |\omega_{11}^* - \omega_{11} - (\ell - 1)\omega_{11}| \\ &= |\hat{\pi}_{21}^{1*} C \hat{\pi}_{21}^* - h \omega_{11}| = 0 \end{aligned} \quad (3.1.26)$$

donde  $h = \ell - 1$

Dado que  $\text{plim}_{T \rightarrow \infty} \hat{\pi}_{21}^{1*} C \hat{\pi}_{21}^* = \pi_{21}^{1*} \bar{C} \pi_{21}^*$  y  $\text{plim}_{T \rightarrow \infty} \frac{1}{T} \omega_{11} = \Omega_{11}$ , la distribución de las raíces  $h_{G_{1+1}} \geq, \dots, \geq h_1 \geq 0$  de (3.1.26) dependerá de las raíces  $\eta_{G_{1+1}} \geq, \dots, \geq \eta_1 \geq 0$  de

$$|\pi_{21}^{1*} \bar{C} \pi_{21}^* - \eta \Omega_{11}| = 0 \quad (3.1.27)$$

Por lo que

$$\text{plim}_{T \rightarrow \infty} (\eta_i - h_i) = 0$$

Las pruebas de identificación discutidas por Koopmans y Hood y las propuestas por Anderson y Rubin, se basan en la distribución -

de las raíces de (3.1.26). El procedimiento en ambos casos se sustenta en el principio de la razón de verosimilitud, que como se citará más adelante proporciona la apropiada estadística de prueba. En particular el estimador M.V.I.L. de  $\beta_1^*$  es el vector característico normalizado asociado con la menor raíz de (3.1.25).

Historicamente la posibilidad de elaborar pruebas de hipótesis de identificabilidad (Anderson y Rubin 1949) no fueron sino un resultado de la aplicación de algunas propiedades de los estimadores M.V.I.L. para una sola ecuación en particular. Por tal motivo creemos conveniente presentar de la manera más general los principios y resultados básicos de la estimación M.V.I.L., para con esto lograr una mayor claridad en la comprensión de las pruebas y enriquecer por ende, su significado con esta interpretación.

## 2. EL METODO DE MAXIMA VEROSIMILITUD CON INFORMACION LIMITADA

El método de estimación por máxima verosimilitud, como es bien sabido, hace uso explícito de supuesto de normalidad de las perturbaciones para derivar las estimaciones de los parámetros estructurales vía la maximización de la función de distribución conjunta. Para esto puede disponer de toda la información a priori de que disponga sobre un modelo dado, para estimar así todos sus parámetros conjuntamente (tal como es el caso de información completa); o bien, si está interesado en estimar sólo un subconjunto de relaciones del modelo, utilizar entonces aquella información a priori perteneciente a la ecuación o ecuaciones cuyos parámetros desea estimar, de tal forma que la información de aquellos que pertenezcan a las restantes ecuaciones no sea considerada. Para este fin, transforma el modelo con la particularidad de que los dos subsistemas resultantes sean mutuamente independientes y con la condición de que los parámetros del subsistema de interés no se alteren. Este último es el método de máxima verosimilitud con información limitada (M.V.I.L.) y es del que nos ocuparemos brevemente en esta sección.

Supongamos entonces que estamos interesados en las primeras  $G^*$  ecuaciones de un modelo con un total de  $G$ ,  $G^* < G$ , de tal manera que el sistema que nos interesa sea de la forma

$$YB_1 + X\Gamma_1 = V_1 \quad (3.2.1)$$

donde  $B_1$  y  $\Gamma_1$  son respectivamente las matrices de coeficientes de las variables endógenas y exógenas que aparecen en las primeras  $G^*$  ecuaciones,  $V_1$  es la matriz de perturbaciones correspondientes a las  $G^*$  ecuaciones y cuyos vectores tienen la propiedad:

$$u_t \sim N(0, \Sigma); \text{cov}(u_t, u'_s) = \Sigma \delta_{t,s}; E(u'_t X) = 0 \quad (3.2.2)$$

donde  $\delta_{t,s}$  es la delta de Kronecker y para  $t, s = 1, \dots, G^*$

El estimador M.V.I.L. de los parámetros estructurales para las primeras  $G^*$  ecuaciones se obtiene maximizando la función de log-verosimilitud "concentrada".

$$L(\alpha_0, \Sigma_{11}; Y, X) = C - \frac{T}{2} \ln |\Sigma_{11}| + \frac{T}{2} \ln (B'_1 W B_1) - \frac{T}{2} t_n (\Sigma_{11}^{-1} \alpha_0 S \alpha_0) \quad (3.2.3)$$

donde  $C = -\frac{GT}{2} (\ln(2\pi) + 1) + \frac{TG^*}{2} - \frac{T}{2} \ln |W|$

$$S = \frac{1}{T} \begin{pmatrix} Y'Y & Y'X \\ X'Y & X'X \end{pmatrix} = \begin{pmatrix} S_{yy} & S_{yx} \\ S_{xy} & S_{xx} \end{pmatrix}$$

y  $W$  es la matriz de productos cruzados de los residuos de la regresión de  $Y$  sobre  $X$ , i.e.,  $W = S_{yy} - S_{yx} S_{xx}^{-1} S_{xy} = Y'(I - X(X'X)^{-1}X')Y$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}; \quad \alpha_0 = \begin{pmatrix} B_1 \\ \Gamma_1 \end{pmatrix}$$

$\Sigma_{11}$  es la matriz de varianzas y covarianzas de los errores de las primeras  $G^*$  ecuaciones.

Si el subsistema en el que estamos interesados consistiera de una sola ecuación digamos la primera, entonces siguiendo la misma notación que en la sección anterior tendríamos

$$G^* = 1, \quad \alpha_0 = \begin{pmatrix} \beta_{01}^* \\ 0 \\ \gamma_{01} \end{pmatrix}, \quad B_1 = \begin{pmatrix} \beta_{01}^* \\ 0 \end{pmatrix}, \quad \Gamma_1 = \begin{pmatrix} \gamma_{01} \\ 0 \end{pmatrix}, \quad \Sigma_{11} = \sigma_{11} \quad (3.2.4)$$

De tal suerte que particionando  $W$  y  $S$  conformablemente a  $\alpha_0$  y  $B_1$  como

$$S = \begin{pmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{pmatrix}, \quad W = \begin{pmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{pmatrix} \quad (3.2.5)$$

entonces el estimador de  $\sigma_{11}$  y  $\delta_{01} = \begin{pmatrix} \beta_{01}^* \\ \gamma_{01} \end{pmatrix}$  se obtiene maximizando la función de log-verosimilitud concentrada

$$L(\alpha_0, \sigma_{11}; Y, X) = C_0 - \frac{T}{2} \ln \sigma_{11} (\beta_{01}^* W_{11} \beta_{01}^*) \quad (3.2.6) \\ - \frac{T}{2} \frac{1}{\sigma_{11}} (\beta_{01}^* S_{11} \beta_{01}^* + 2 \gamma_{01} S_{31} \beta_{01}^* + \gamma_{01}^2 S_{33})$$

donde  $C_0 = \frac{GT}{2} (\ln 2\pi + 1) + \frac{T}{2} (1 - \ln |W|)$  y  $W_{11}$  es la matriz de -- productos cruzados de los residuos de la regresión de  $Y_1^*$  en  $X$ . -- Maximizando (3.2.6) con respecto a  $\gamma_{01}$  se obtiene

$$\hat{\gamma}_{01} = S_{33}^{-1} S_{31} \beta_{01}^* \quad (3.2.7)$$

sustituyendo (3.2.7) en (3.2.6) y maximizando con respecto a  $\sigma_{11}$  resulta

$$\hat{\sigma}_{11} = \beta_{01}^* W_{11}^* \beta_{01}^* \quad (3.2.8)$$

donde  $\omega_{11}^* = S_{11} - S_{13} S_{33}^{-1} S_{31}$ , esto es,  $\omega_{11}^*$  es la matriz de productos cruzados de los residuos de la regresión de  $y_1^*$  en  $X_1$ . Es fácil verificar que tanto (3.2.7) como (3.2.8) representan un máximo para un valor dado de  $\beta_{01}^*$ . Sustituyendo ahora estas expresiones en (3.2.6) obtenemos la función de log-verosimilitud concentrada:

$$L(\beta_{01}^*; Y, X) = C_0 \frac{T}{2} - \frac{T}{2} \ln \left( \frac{\beta_{01}^{*'} \omega_{11}^* \beta_{01}^*}{\beta_{01}^{*'} \omega_{11} \beta_{01}^*} \right) \quad (3.2.9)$$

Al maximizar (3.2.9) con respecto a  $\beta_{01}^*$  obtendríamos su estimador M.V.I.L., que sustituyendo en (3.2.7) y (3.2.8) nos completaría la estimación de todos los parámetros de interés. Para este efecto nótese que maximizar (3.2.9) es equivalente a minimizar la expresión

$$\ell(\beta_{01}^*) = \frac{\beta_{01}^{*'} \omega_{11}^* \beta_{01}^*}{\beta_{01}^{*'} \omega_{11} \beta_{01}^*} \quad (3.2.10)$$

en donde  $\beta_{01}^{*'} \omega_{11}^* \beta_{01}^*$  representa la varianza para un vector dado  $\beta_{01}^*$ , de los residuos de la regresión de la variable "compuesta"  $\bar{y}_1^* = y_1^* \beta_{01}^*$  sobre  $X_1$ . Similarmente  $\beta_{01}^{*'} \omega_{11} \beta_{01}^*$  se interpreta como la varianza de los residuos de la regresión de  $\bar{y}_1^*$  sobre todo  $X$ . Es por tal motivo que este procedimiento de obtener estimaciones también se le denomina principio de la razón de mínima varianza.

La razón de mínima varianza satisface que

$$\ell = \ell(\beta_{01}^*) \geq 1 \quad (3.2.11)$$

esto se debe a que el denominador debe ser siempre igual o menor que el numerador, porque el uso de variables independientes adicionales en una regresión no puede incrementar la suma de los cuadrados de los residuos. En última instancia las variables adicionales no evidenciarán alguna relación respecto de la variable dependiente, por lo que  $\ell(\beta_{01}^*)$  será igual a 1.

La función  $\ell(\beta_{01}^*)$  es homogénea de grado cero, lo que implica que si un valor minimizante de  $\beta_{01}^*$  existe, digamos  $\beta_{01}^* = b_{01}$ , -- también se satisfará para  $\beta_{01}^* = kb_{01}$ , i.e.  $b_{01}$  no será único. Para asignar un único valor a  $k$ , es necesario imponer alguna normalización, que podría ser por ejemplo, que el primer elemento de  $\beta_{01}^*$  sea -1 (como se requirió en el capítulo II) ó  $\beta_{01}^* \beta_{01}^* = 1$ , o bien  $\beta_{01}^* w_{11} \beta_{01}^* = 1$ , etc. Se destaca pues que la normalización aquí requerida no es una en particular, esto es, no es una parte integrante del procedimiento estadístico de estimación, sino más bien un requisito matemático para lograr una solución única.

Para minimizar  $\ell(\beta_{01}^*)$  con respecto a  $\beta_{01}^*$ , hay que diferenciar  $\ell$  con respecto a cada componente de  $\beta_{01}^*$ , igualar los resultados a cero y resolver para los componentes de  $\beta_{01}^*$  i.e.:

$$\begin{aligned} \frac{\partial \ell}{(\beta_{01}^*)_i} &= \frac{\partial}{(\beta_{01}^*)_i} \frac{\beta_{01}^* w_{11} \beta_{01}^*}{\beta_{01}^* w_{11} \beta_{01}^*} \\ &= 2 \frac{\beta_{01}^* w_{11} \beta_{01}^* (w_{11} \beta_{01}^*)_i - \beta_{01}^* w_{11} \beta_{01}^* (w_{11} \beta_{01}^*)_i}{(\beta_{01}^* w_{11} \beta_{01}^*)^2} = 0 \end{aligned}$$

$$\Rightarrow (w_{11} \beta_{01}^*)_i = \ell(w_{11} \beta_{01}^*)_i \quad (\text{por (3.2.10)})$$

dato que esto debe valer para cada uno de los componentes de los vectores  $W_{11}^* \beta_{11}^*$  y  $W_{11} \beta_{11}^*$ , puede entonces formularse en términos de los vectores mismos:

$$W_{11}^* \beta_{11}^* = \ell W_{11} \beta_{11}^* \quad (3.2.12)$$

$$\Rightarrow (W_{11}^* - \ell W_{11}) \beta_{11}^* = 0 \quad (3.2.13)$$

el sistema lineal homogéneo de  $G_1+1$  ecuaciones (3.2.13) tendrá solución no nula si y sólo si

$$\chi(W_{11}^* - \ell W_{11}) \leq G_1 \quad (3.2.14)$$

es decir, si

$$|W_{11}^* - \ell W_{11}| = 0 \quad (3.2.15)$$

que es una ecuación polinomial en  $\ell$  de grado  $G_1 + 1$ .

En virtud de que  $W_{11}^*$  y  $W_{11}$  son matrices positivas definidas, existe una matriz  $F$  no singular tal que

$$W_{11} = F'F \quad \text{y} \quad W_{11}^* = F' \Lambda F$$

donde  $\Lambda$  es la matriz diagonal cuyos elementos son las raíces características de la solución a (3.2.15). De ahí que (3.2.16) pueda escribirse como

$$\ell(\beta_{11}^*) = \frac{\beta_{11}^* F' \Lambda F \beta_{11}^*}{\beta_{11}^* F' F \beta_{11}^*} = \frac{\xi' \Lambda \xi}{\xi' \xi} = \sum_{i=1}^{G_1+1} \ell_i \left( \frac{\xi_i^2}{\sum_{j=1}^{G_1+1} \xi_j^2} \right) \geq \ell_1 \quad (3.2.16)$$

donde  $\xi = F \beta_{11}^*$  y  $\ell_1 = \min_i \ell_i$

la última desigualdad se da por ser los coeficientes de  $\ell_i$  positivos y dado que  $\ell_i \geq 0$ .

Por lo tanto el estimador M.V.I.L. de  $\beta_{11}^*$  será el vector característico de  $\omega_{11}^*$  en la métrica inducida por  $\omega_{11}$ , correspondiente a la menor raíz característica  $\ell_1$ , esto es:

$$\omega_{11}^* \hat{\beta}_{11}^* = \ell_1 \omega_{11} \hat{\beta}_{11}^* \quad (3.2.17)$$

Es claro entonces, con la serie de resultados que hemos aquí enunciado sobre la estimación M.V.I.L., el porqué este procedimiento nos enmarca la elaboración de pruebas de identificabilidad: basta comprobar que las expresiones (3.1.26) y (3.2.15) son la misma. De hecho como lo apuntamos en la sección anterior, las pruebas de Koopmans y Hood, y las de Anderson y Rubin, se basan en la distribución de la menor raíz de

$$|\hat{\pi}_{21}^* C \hat{\pi}_{21}^* - h \omega_{11}| = |\omega_{11}^* - \ell \omega_{11}| = 0$$

donde  $h = \ell - 1$ , es decir, en la raíz asociada al estimador  $\hat{\beta}_{11}^*$  de M.V.I.L. Antes de enunciar las pruebas, las siguientes observaciones serán importantes.

Primero, tenemos por (3.2.14) que

$$\chi(\omega_{11}^* - \ell \omega_{11}) \leq G_1$$

Koopmans y Hood demuestran que (ver Koopmans y Hood 1953 apéndice F, p.p. 195);  $\chi(w_{11}^* - \ell_i w_{11}) = G_1$  siempre y cuando  $\ell_i$  sea una raíz simple y  $\chi(w_{11}^* - \ell_i w_{11}) < G_1$  cuando  $\ell_i$  sea una raíz múltiple. En particular este resultado se relaciona con la menor raíz  $\ell_i$  de (3.2.16). Se sigue por tanto que si  $\ell_1$  es una raíz simple, entonces existirá una única solución (después de la normalización) para  $\beta_{01}^*$ .

:

Segundo, hemos visto que  $\ell(\beta_{01}^*) \geq 1$  y por (3.2.16) se sigue -- también que  $\ell_1 \geq 1$ . Es interesante citar aquí bajo que condiciones se tiene  $\ell_1 = 1$  y bajo cuales  $\ell_1 > 1$ .

Para esto citaremos un teorema (ver Fisher 1966, p.p. 186---187) que rephraseándolo en términos de nuestra notación nos será de gran utilidad. Además vendrá a reforzar lo que ya hemos planteado.

Teorema.- Considérese la ecuación polinomial

$$|w_{11}^* - \ell w_{11}| = 0 \quad (3.2.18)$$

donde  $w_{11}$  es positiva definida, y sea  $\ell_1 = \min_i \ell_i$  la menor raíz de la ecuación (3.2.18). Sea  $\beta_{01}^*$  un vector de  $(G_1+1, 1)$  no nulo, entonces

$$\ell_1 = \min \frac{\beta_{01}^* w_{11}^* \beta_{01}^*}{\beta_{01}^* w_{11} \beta_{01}^*} \quad (3.2.19)$$

"Prueba.- Dado que el lado derecho de la ecuación (3.2.19) -

es homogénea de grado cero en  $\beta_{01}^*$ , podemos imponer la normalización

$$\beta_{01}^* \omega_{11} \beta_{01}^* = 1 \quad (3.2.20)$$

y minimizar  $\beta_{01}^* \omega_{11} \beta_{01}^*$  sujeto a esta restricción.

Para lograrlo utilizamos el Lagrangeano

$$L = \beta_{01}^* \omega_{11} \beta_{01}^* - \ell (\beta_{01}^* \omega_{11} \beta_{01}^* - 1) \quad (3.2.21)$$

Por condiciones de primer orden para un mínimo, tenemos además de (3.2.20) que

$$(\omega_{11}^* - \ell \omega_{11}) \beta_{01}^* = 0 \quad (3.2.22)$$

que tendrá solución en  $\beta_{01}^*$  si y solo si  $(\omega_{11}^* - \ell \omega_{11})$  es singular, es decir, el multiplicador de Lagrange  $\ell$  es una raíz de la ecuación (3.2.18).

Suponga entonces que  $\beta_{01}^*$  satisface las ecuaciones (3.2.20) y (3.2.22). Premultipliquemos (3.2.22) por  $\beta_{01}^*$  y reordenando términos obtenemos

$$\beta_{01}^* \omega_{11} \beta_{01}^* = \ell (\beta_{01}^* \omega_{11} \beta_{01}^*) = \ell \quad (3.2.23)$$

de tal forma que, como  $\beta_{01}^*$  satisface las condiciones de primer orden,  $\ell$  es el valor minimizante; es decir,  $\ell_1$ .

Una consecuencia directa de este teorema es el siguiente corolario.

Corolario.-  $\ell_1 > 1$  si y sólo si  $(W_{11}^* - W_{11})$  es positiva definida y  $\ell_1 = 1$  si y sólo si  $(W_{11}^* - W_{11})$  es positiva semidefinida.

Prueba.- Si  $(W_{11}^* - W_{11})$  es positiva definida, entonces -----  
 $\beta_{01}^* W_{11}^* \beta_{01}^* > \beta_{01}^* W_{11} \beta_{01}^*$  y si  $(W_{11}^* - W_{11})$  es positiva semidefinida entonces  $\beta_{01}^* W_{11}^* \beta_{01}^* \geq \beta_{01}^* W_{11} \beta_{01}^*$ . Dado que  $W_{11}$  es positiva definida,  $\beta_{01}^* W_{11} \beta_{01}^* > 0$  y dividiendo las desigualdades por esta expresión nos da el resultado deseado, como una consecuencia del teorema anterior.

Dado que en nuestro caso la aplicación del corolario corresponde a la diferencia de matrices de varianzas y covarianzas para dos conjuntos de residuales, se sigue que  $(W_{11}^* - W_{11})$  es al menos una matriz positiva semidefinida y será positiva definida sólo si la matriz  $(W_{11}^* - W_{11})$  tiene todos sus renglones linealmente independientes. Lo que si podemos asegurar es que asintóticamente -----  $(W_{11}^* - W_{11})$  tendrán renglones linealmente independientes, es decir

$$\text{plim } \ell_1 = 1. \quad (3.2.24)$$

Conviene recordar (Capítulo II, sección 1) que el estimador de mínimos cuadrados indirectos (M.C.I.) para la primera ecuación estructural, se puede obtener en forma única si

$$\kappa(\Pi_{21}^*) = G_1 \text{ con } K_2 = G_1 \quad (3.2.25)$$

pues bien, si (3.2.25) se cumple, entonces los estimadores M.C.I. y M.V.I.L. coincidirán y se tendrá en este caso que  $\ell_1 = 1$ .

Supongamos ahora que  $K_2 = G_1 + 1$ , entonces por la condición de orden la ecuación está sobreidentificada y con probabilidad no cero podríamos tener

$$h(\pi_{21}^*) = G_1 + 1$$

en cuyo caso la condición necesaria y suficiente para identificación se viola y solamente la solución trivial existirá para ---- (5.2.17). En cuyo caso se tendrá

$$l(\beta_{21}^*) > 1, \quad l_1 > 1$$

por lo tanto podemos ver nuestro problema de probar hipótesis de identificación, como la prueba a un cierto nivel de significancia del exceso de  $l_1$  sobre 1.

### 3. PRUEBAS DE HIPOTESIS DE LA CONDICION DE RANGO PARA IDENTIFICACION

El criterio de la razón de verosimilitud es una opción al problema de construir pruebas de hipótesis, en particular cuando queremos probar una hipótesis compuesta contra otra alternativa compuesta, o bien una hipótesis simple contra una alternativa compuesta. "... El criterio de la razón de verosimilitud, como es sabido, no es necesariamente la prueba uniformemente más potente, pero se ha comprobado en la literatura que esta prueba posee frecuentemente propiedades deseables". (Hogg y Craig 1978 p.p. 258).

La razón de verosimilitud se define como

$$\lambda = \frac{\max_{H_0} L^* (\alpha_0, \sigma_{11}, Y, X)}{\max_{H_1} L^* (\alpha_0, \sigma_{11}, Y, X)} \quad (3.3.1)$$

es decir,  $\lambda$  es la razón de las dos funciones de verosimilitud cuando se sustituyen sus parámetros por sus estimadores máximo-verosímiles. Se espera que si  $H_0$  es cierta, entonces  $\lambda$  tenderá a 1 y si por el contrario  $H_0$  no es cierta,  $\lambda$  tenderá a ser menor que la unidad. Por lo tanto para verificar una hipótesis  $H_0$ , simple o compuesta, se utiliza el estadístico  $\lambda$  y se rechaza  $H_0$  siempre y cuando el valor muestral de  $\lambda$  satisfaga  $\lambda < \lambda_\alpha$ ; donde  $\lambda_\alpha$  es alguna cantidad definida a un nivel de significancia  $\alpha$ .

Para nuestros propósitos transcribiremos y refrasearemos (sin

demostración) dos teoremas debidos a Anderson (1953 p.p. 336-337) - sobre el criterio de la razón de la verosimilitud. (Recuérdese -- que  $\Pi_{21}^*$  es  $(K_2, G_1 + 1)$ ).

Teorema 1.- El criterio de la razón de la verosimilitud para probar la hipótesis de que el rango de  $\Pi_{21}^*$  es  $\kappa_1$  contra la alternativa de que es  $\kappa_2$  ( $\kappa_2 > \kappa_1$ ) es:

$$\lambda = \frac{\prod_{i=\kappa_1+1}^{G_1+1} (1+h_i)^{-1T}}{\prod_{i=\kappa_2+1}^{G_1+1} (1+h_i)^{-1T}} = \frac{\kappa_2}{\prod_{i=\kappa_1+1}^{\kappa_2} (1+h_i)^{-1T}} \quad (3.3.2)$$

donde  $\{h_i\}$  son las raíces ordenadas en forma descendente de

$$|\hat{\Pi}_{21}^* C \hat{\Pi}_{21}^* - h W_{11}| = 0 \quad (3.3.3)$$

En particular el criterio para probar  $\kappa(\Pi_{21}^*) = \kappa_1$ , contra cualquier otra posible alternativa requiere del producto sobre todas las  $(G_1 - 1) - \kappa_1$  raíces más pequeñas. En este caso

$$-2 \ln \lambda = T \sum_{i=\kappa_1+1}^{G_1+1} \ln (1+h_i) = T \sum_{i=\kappa_1+1}^{G_1+1} h_i \quad (3.3.4)$$

para muestras suficientemente grandes.

Teorema 2.- Sea  $p \lim \frac{1}{T} X'X = Q$ ,  $Q$  simétrica y positiva definida, y supóngase que  $\kappa(\Pi_{21}^*) = \kappa_1$ . Entonces  $-2 \ln \lambda$ , donde  $\lambda$  es el criterio de la razón de verosimilitud definida en el teorema 1 para probar la hipótesis de que el rango de  $\Pi_{21}$  es  $\kappa_1$  contra la alternativa de que no es mayor que  $G_1 + 1$ , se distribuye asintóticamente co-

mo  $\chi^2$  con  $(G_1+1-k_1) \cdot (K_2-k_1)$  grados de libertad.

Suponga entonces que estamos interesados en probar la hipótesis

$$H_0 = \kappa(\Pi_{21}^*) = G_1$$

contra la alternativa

(3.3.5)

$$H_1 = \kappa(\Pi_{21}^*) = G_1 + 1$$

El criterio de la razón de verosimilitud para probar (3.3.5) es, por (3.3.2)

$$\lambda = (1+h_1)^{-1T} \quad (3.3.6)$$

donde  $h_1$  es la raíz más pequeña de (3.3.3). Si  $\Pi_{21}^*$  tiene rango  $G_1$ , entonces  $-2\ell n \lambda$  y su equivalente asintótico  $Th_1$  tendrán asintóticamente una distribución  $\chi^2$  con  $(K_2 - G_1)$  grados de libertad. Por lo tanto la región crítica asintótica para probar  $H_0$  contra  $H_1$  a un nivel de significancia  $\alpha$  es

$$T \ell n (1+h_1) \geq \chi_{\alpha}^2(K_2 - G_1)$$

se espera que para valores de  $h_1$  cercanos a cero, i.e., valores de  $\ell_1$  cercanos a 1 (véase 3.1.26) no se rechace la hipótesis nula. Es evidente que la prueba es innecesaria cuando los grados de libertad sean cero o negativos, es decir, la prueba no sirve para ecuaciones exactamente identificadas porque para ellas  $K_2 - G_1 = 0$ , de modo que no hay grados de libertad para la prueba y  $\ell_1$  resulta ser siempre exactamente 1. Más resulta inoperable si la condición de orden no se satisface. Nótese además que si rechazamos la hipóte-

sis nula, esto es, si no rechazamos que  $\kappa(\Pi_{21}^*) = G_1 + 1$ , entonces no existirá una solución no nula para el estimador M.V.I.L., es decir  $\beta_{01}^* = 0$ . (véase 3.1.9).

La anterior prueba puede generalizarse. Supongamos que tenemos alguna información que nos sugiere que  $\kappa(\Pi_{21}^*) = q$ ,  $q < G_1$ . Podemos probar entonces, de acuerdo al teorema 1, la hipótesis:

$$H_0: \kappa(\Pi_{21}^*) = q \quad (3.3.8)$$

contra la alternativa

$$H_1: \kappa(\Pi_{21}^*) = G_1 + 1$$

La región crítica asintótica para probar (3.3.8) a un nivel de significancia  $\alpha$  es

$$T \sum_{i=q+1}^{G_1+1} \ln(1+h_i) \geq \chi_{\alpha}^2 (G_1+1-q) \cdot (K_2-q) \quad (3.3.9)$$

se espera que esta prueba rechace la hipótesis nula si al menos -- una de las raíces  $h_2, \dots, h_q$  es suficientemente grande ( $h_1=0$ ). Por lo tanto si la hipótesis nula se rechaza podemos hacer pruebas secuenciales para poder decidir si la condición de rango para identificabilidad es satisfecha. Probaríamos entonces:

$$H_0: \kappa(\Pi_{21}^*) = q + 1$$

contra la alternativa

$$H_1: \kappa(\Pi_{21}^*) = G_1 + 1$$

y si la hipótesis nula es rechazada probaríamos entonces

$$H_0: \kappa(\Pi_{21}^*) = q+2$$

contra la alternativa

$$H_1: \kappa(\Pi_{21}^*) = G_1+1$$

y así sucesivamente.

La región crítica para estas pruebas secuenciales a un nivel  $\alpha$  de significancia es:

$$T \sum_{i=q+j+1}^{G_1+1} \ln(1+h_i) \geq \chi_{\alpha}^2 (G_1+1-q-j) \cdot (K_2-q-j) \quad (3.3.10)$$

donde  $0 \leq j \leq G_1 - K_2 - 2$ .

Sin embargo, nos dice Savin (1973 p.p. 12), este método secuencial no es completamente satisfactorio, porque si se utiliza sólo una muestra para efectuar la prueba secuencial, entonces las esta-

dísticas  $T \sum_{i=q-j-1}^{G_1+1} \ln(1+h_i)$  no serán estadísticamente independientes, y por lo tanto los resultados presentarán sesgos.

#### Prueba de doble raíz de Koopmans y Hood.-

Hemos visto que la distribución de las raíces  $h_{G_1+1} \geq \dots \geq h_1 > 0$  de (3.3.3) dependen de las raíces  $\eta_{G_1+1} \geq \dots \geq \eta_1 \geq 0$  de

$$|\Pi_{21}^* \bar{C} \Pi_{21}^* - \eta \Omega_{11}| = 0 \quad (3.3.11)$$

y dado que  $\hat{\Pi}_{21}^*$ ,  $C$  y  $w_{11}$  son estimadores consistentes de  $\Pi_{21}^*$ ,  $\bar{C}$  y  $\Omega_{11}$  respectivamente, se sigue que

$$plim (h_{\lambda} - \eta_{\lambda}) = 0 \quad (3.3.12)$$

Por otra parte  $\kappa(\Pi_{21}^*) = G_1$  si y sólo si  $\eta_1 = 0$  y  $\eta_2 \neq 0$ . Por lo tanto resultará intuitivamente razonable rechazar la hipótesis de que  $\kappa(\Pi_{21}^*) = G_1$  si la menor raíz (3.3.3) es suficientemente grande.

Koopmans y Hood sugieren que para obtener una prueba de la condición de rango para identificación, es conveniente (en el sentido de seguir la práctica general de reservar el término de hipótesis nula para el subconjunto de menor dimensión en el espacio de parámetros) escoger como hipótesis nula:

$$H_0: \kappa(\Pi_{21}^*) \leq G_1 - 1 \quad (3.3.13)$$

contra la alternativa

$$H_1: \kappa(\Pi_{21}^*) = G_1$$

Estrictamente hablando, esta es una prueba de no identificabilidad en vez de una prueba de identificabilidad, porque prueba la hipótesis de que la ecuación en consideración (en este caso la primera) esté no identificada.

Si la hipótesis nula se satisface, entonces  $h_1 = 0$  será por lo menos una raíz doble de (3.3.3), esto es, la segunda raíz más pequeña  $h_2$  de (3.3.3), será también 0. Si la hipótesis alternati---

va no se rechaza se tendrá que  $h_1=0$  es una raíz simple, es decir,  $h_2>0$ . Esto por supuesto sugiere el uso de la segunda raíz  $h_2$  de (3.3.3), como el criterio de prueba.

El criterio de la razón de verosimilitud para probar (3.3.13) si  $H_0$  es cierta es

$$\lambda_2 = (1+h_2)^{-\frac{1}{2}T} \quad (3.3.14)$$

sin embargo, la distribución asintótica de  $T \ln(1+h_2)$  no es  $\chi^2$  bajo la hipótesis nula, debido a la multiplicidad ( $h_1=h_2$ ). Los autores sugieren por este motivo usar la estadística

$$T \ln(1+h_1)(1+h_2) \quad (3.3.15)$$

o su equivalente asintótico  $T(h_1+h_2+2)$ , los cuales se distribuyen asintóticamente como  $\chi^2_\alpha$  con  $K_2-G_1-1$  grados de libertad. Por lo tanto la región crítica asintótica para probar (3.3.13) a un nivel de significancia  $\alpha$  es

$$T \ln(1+h_2)(1+h_1) \geq \chi^2_\alpha (K_2-G_1-1) \quad (3.3.16)$$

siempre y cuando  $h_3>h_2$ .

La prueba nuevamente será innecesaria, si los grados de libertad son cero o negativos, porque en ese caso la condición de orden para identificación no se satisface.

Savin (1973, p.p. 10) basado en el trabajo de Fisk (1967, p. 52) indica que los grados de libertad reportados por Koopmans y

Hood en su prueba de doble raíz son erróneos. La verdadera región crítica nos dice, para probar que el rango de  $\Gamma_{21}^*$  es a lo más  $G_1 - 1$  contra la alternativa de que es  $G_1$  a un nivel de significancia  $\alpha$  - es

$$T \ln(1+h_2)(1+h_1) \geq \chi_{\alpha}^2 2(K_2 - G_1 - 1) \quad (3.3.17)$$

además propone una;

Prueba de doble raíz para pequeña muestra:

La región crítica para pequeña muestra para probar (3.3.13) - es

$$[(1+h_2)(1+h_1)]^{-1} \leq u_{\alpha}(2, K_2, T-K) \quad (3.3.18)$$

donde  $u_{\alpha}(2, K_2, T-K)$  denota una distribución Wilk, a un nivel de significancia  $\alpha$  con 2,  $K_2$  y  $T-K$  grados de libertad (ver Anderson 1958 p.p. 193-202).

Para hacer operativa la prueba se hace uso del hecho que (teorema 8.5.4 p. 196 Anderson 1958)

$$\frac{1 - \sqrt{u_{2, K_2, T-K}}}{\sqrt{u_{2, K_2, T-K}}} \sim F_{2K_2, 2(T-K-1)} \quad (3.3.19)$$

donde  $F_{2K_2, 2(T-K-1)}$  denota una distribución F con  $2K_2$  y  $2(T-K-1)$  grados de libertad.

Aunque esta prueba concluye Savin, es conservadora en el sen-

tido de que si  $H_0$  es cierta, será aceptada en al menos  $100(1-\alpha)\%$  de las muestras.

### Pruebas de las restricciones cero

Hemos apuntado en secciones anteriores que una ecuación estará identificada si el número de restricciones que existen sobre esta, son las suficientes como para poder determinar sus parámetros estructurales en forma única. Las restricciones a las que nos hemos referido son solamente hacia el interior de las ecuaciones, es decir, exclusiones y normalizaciones que en particular hemos denotado (las exclusiones) como:

$$\begin{pmatrix} \beta_2 \\ \gamma_2 \end{pmatrix} = 0$$

Si una ecuación, digamos la primera, está realmente identificada, es decir, si  $\lambda(\pi_{21}^*) = G_1$ , entonces podemos suponer que las exclusiones que hemos considerado son realmente correctas. Esto se puede traducir en que las pruebas que anteriormente se han enunciado sobre la condición de rango para identificación son válidas para probar las restricciones cero, esto es, la prueba:

$$H_0: \lambda(\pi_{21}^*) = G_1 \quad (3.3.20)$$

contra la alternativa

$$H_1: \lambda(\pi_{21}^*) = G_1 + 1$$

es equivalente a probar

$$H_0: \begin{pmatrix} \beta_2 \\ \gamma_2 \end{pmatrix} = 0 \quad (3.3.21)$$

contra la alternativa

$$H_1: \begin{pmatrix} \beta_2 \\ \gamma_2 \end{pmatrix} \neq 0$$

En particular la prueba (3.3.20) se puede interpretar como la prueba de que hay  $K_1$  variables exógenas en la primera ecuación estructural, contra la alternativa de que hay más de  $K_1$ .

Como ya sabemos, el criterio de la razón de verosimilitud para probar  $H_0$  contra la alternativa  $H_1$  es  $\lambda = (1+h_1)^{-1T}$  y la región crítica a un nivel de significancia  $\alpha$  será

$$T \ln(1+h_1) \geq \chi_\alpha^2(K_2 - G_1) \quad (3.3.22)$$

La prueba basada en la región crítica (3.3.22) para probar -- (3.3.21) puede producir resultados erráticos para muestras pequeñas. Basman (1960) basado en experimentos Monte Carlo, indica que esta prueba tiene un marcado sesgo contra  $H_0$ . Por tal motivo propone una prueba alternativa basada en la región crítica

$$h_1 \frac{T-K}{K_2} \geq F_\alpha(K_2, T-K)$$

donde  $F_\alpha(K_2, T-K)$  denota una distribución  $F$  con  $K_2$  y  $T-K$  grados de libertad. Esta prueba se encontró que es sesgada en favor de  $H_0$ , por lo que finalmente sugiere el uso de la región crítica

$$h_1 \frac{T-K}{K_2-G_1} \geq F_\alpha (K_2-G_1, T-K)$$

como una aproximación para muestras pequeñas de la prueba de las restricciones cero.

Los detalles de las anteriores consideraciones no los abordaremos en el presente trabajo ya que superan los objetivos del mismo.

En resumen, las pruebas de hipótesis de la condición de rango para identificación, pueden garantizarnos resultados favorables si nuestro objetivo es hacer pruebas estadísticas sobre las exclusiones de variables en nuestro modelo. Dada la naturaleza de las pruebas, la prueba de Anderson y Rubin procederá cuando no se considere explícitamente la normalización  $\beta_{ii} = -1$ , por otra parte si hacemos uso explícito de la normalización entonces procederá la prueba de Kóopmans y Hood.

Los aquí descritos no son por supuesto los únicos métodos para elaborar pruebas de hipótesis de la condición de rango para identificación. Gill (1978) utilizando distribuciones matriciales prueba directamente la condición (1.2.12) para identificabilidad; Byron (1972) y Wald proponen algunas variantes a la prueba de Anderson y Rubin en términos de estimadores clase- $k$ . Sin embargo, la heterogeneidad de la notación y terminología utilizada dificultan su comprensión. Esto se refleja también en que los diferentes estadísticos que utilizan la menor raíz de (3.3.3) como criterio -

de prueba, aparentemente no utilizan criterios homogéneos.

Para finalizar esta sección, citaremos algunas de las conclusiones a las que llegan Farebrother y Savin (1973) al probar la -- condición de rango para identificación, con los criterios aquí descritos, en las ecuaciones del modelo Klein I y Klein II. "... una inspección a los resultados sugiere que la prueba de doble raíz es sesgada en favor del rechazo de la hipótesis nula, *ie*, no rechazar la condición de rango, para identificación. ... Los datos también son consistentes con el hecho de que la prueba de Anderson y Rubin es sesgada en favor de la hipótesis nula, esto es, en no rechazar la condición de rango. ... las pruebas serán más potentes si incorporan la información a priori de que el coeficiente de una de las variables endógenas es -1. Finalmente debe ser enfatizado, que las pruebas aquí empleadas han supuesto independencia serial - en las perturbaciones, un supuesto que puede no ser satisfecho."

#### 4. UNA INTERPRETACION ALTERNATIVA SOBRE LAS PRUEBAS DE IDENTIFICABILIDAD

En esta sección pretendemos dar una interpretación alternativa y - sobre todo una alternativa de cómputo a las pruebas de identifica- bilidad.

Como se ha señalado, las pruebas de la condición de rango pa- ra identificación dependen de las raíces  $h_{G_1+1} \geq \dots \geq h_1 \geq 0$  del poli- nomio

$$|\hat{\Pi}_{21}^* C \hat{\Pi}_{21}^* - h \omega_{11}| = 0 \quad (3.4.1)$$

que asintóticamente correponderá a las raíces  $\eta_{G_1+1} \geq \dots \geq \eta_1 \geq 0$  de

$$|\Pi_{21}^* \bar{C} \Pi_{21}^* - \eta \Omega_{11}| = 0 \quad (3.4.2)$$

Así que los resultados de la sección anterior requerirán primero - la obtención de las estimaciones de las matrices  $\Pi_{21}$ ,  $C$  y  $\omega_{11}$  por mínimos cuadrados ordinarios y despues utilizar estos resultados - para dar solución a (3.4.1). Existe un método que nos proporciona rá directamente los resultados requeridos y que además es parte in- tegrande de muchos paquetes estadísticos y econométricos de evalua- ción. El procedimiento a que nos referimos esta basado en los as- pectos elementales de la teoría de la correlaciones canónicas.

El método de correlaciones canónicas, creado por Hotelling en 1935, como una generalización del concepto usual de regresión, con- siste en analizar la dependencia entre dos conjuntos de variables

aleatorias y en determinar su máxima correlación. Adoptando el desarrollo de Dhrymes, las generalidades del método las consideraremos a continuación.

Sea  $V_1 = (v_1, \dots, v_{m_1})'$  y  $V_2 = (v_{m_1+1}, \dots, v_{m_1+m_2})'$  dos vectores de variables aleatorias, donde

$$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = V \sim N(\mu, T), \quad T \text{ simétrica y positiva definida, } V \text{ es } (m, 1),$$

con  $m = m_1 + m_2$ .

Particionando  $T$  conformablemente tenemos:

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

$T_{11}$  es  $(m_1, m_1)$ ,  $T_{12}$  es  $(m_1, m_2)$ ,  $T_{21}$  es  $(m_2, m_1)$  y  $T_{22}$  es  $(m_2, m_2)$

Sea

$$Z = \alpha' V_1 \quad \text{y} \quad W = \beta' V_2 \tag{3.4.1}$$

y defínase

$$\rho = \frac{\text{Cov}(\alpha' V_1, \beta' V_2)}{[\text{Var}(\alpha' V_1) \text{Var}(\beta' V_2)]^{1/2}} \tag{3.4.2}$$

donde los coeficientes  $\alpha'$  y  $\beta'$  serán determinados de forma tal que maximicen  $\rho$ .  $\rho$  se denomina la correlación canónica entre  $V_1$  y  $V_2$  y  $Z$  y  $W$  son llamadas la variables canónicas asociadas a  $V_1$  y  $V_2$ .

Dado que  $\text{Var}(\alpha' V_1) = \alpha' T_{11} \alpha$ ,  $\text{Var}(\beta' V_2) = \beta' T_{22} \beta$  y  $\text{Cov}(\alpha' V_1, \beta' V_2) = \alpha' T_{12} \beta$ , el problema de correlaciones canónicas puede expresarse como la solución a

$$\max_{\alpha, \beta} \frac{\alpha' T_{12} \beta}{[(\alpha' T_{11} \alpha) (\beta' T_{22} \beta)]^{1/2}} \quad (3.4.3)$$

La solución a (3.4.3) no es única dado que esta expresión representa una función homogénea de grado cero. Para resolver esa indeterminación es necesario imponer una regla de normalización.

El método impone las restricciones

$$\alpha' T_{11} \alpha = 1 \quad \text{y} \quad \beta' T_{22} \beta = 1 \quad (3.4.4)$$

El anterior es entonces un problema de maximización restringida -- que puede resolverse por el método de multiplicadores de Lagrange. Esto es, buscamos la solución a

$$L = \alpha' T_{12} \beta + \frac{1}{2} \theta_1 (1 - \alpha' T_{11} \alpha) + \frac{1}{2} \theta_2 (1 - \beta' T_{22} \beta) \quad (3.4.5)$$

Derivando  $L$  con respecto a  $\alpha$ ,  $\beta$ ,  $\theta_1$ , y  $\theta_2$ , igualando las parciales a cero y resolviendo para  $\theta_1$  y  $\theta_2$  obtenemos

$$\theta = \theta_1 = \theta_2 = \alpha' T_{12} \beta = \beta' T_{21} \alpha \quad (3.4.6)$$

Resolviendo las parciales para  $\alpha$  y  $\beta$ ; sustituyendo (3.4.6) y reordenando términos, resulta que estamos buscando una solución al siguiente sistema de ecuaciones:

$$\begin{pmatrix} -\theta T_{11} & T_{12} \\ T_{21} & -\theta T_{22} \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \quad (3.4.7)$$

que por tratarse de un sistema de ecuaciones lineal homogéneo tendrá solución no trivial si y sólo si

$$\begin{vmatrix} -\theta T_{11} & T_{12} \\ T_{21} & -\theta T_{22} \end{vmatrix} = 0 \quad (3.4.8)$$

que representa una ecuación polinomial en  $\theta$  de grado  $m$ . A cada raíz de (3.4.8) corresponderá una solución del sistema (3.4.7). -- Por tanto las correlaciones canónicas se obtienen como las raices de (3.4.8), mientras las variables canónicas se obtienen como

$$z_i^* = \alpha_i^* v_1 \quad w_i^* = \beta_i^* v_2$$

donde  $\alpha_i^*$  y  $\beta_i^*$  son los vectores solución de (3.4.7) asociadas a las raices  $\theta_1, \dots, \theta_m$  de (3.4.8);  $\theta_i = \rho_i \quad i=1, \dots, m$ .

Con esta información obtendremos ahora el resultado que particularmente nos interesa. Para esto nótese que

$$\begin{vmatrix} -\theta T_{11} & T_{12} \\ T_{21} & -\theta T_{22} \end{vmatrix} = \begin{vmatrix} -\theta T_{11} + \frac{1}{\theta} T_{12} T_{22}^{-1} T_{21} & 0 \\ T_{21} & -\theta T_{22} \end{vmatrix} =$$

$$= (-\theta)^{-m_1} |-\theta T_{22}| |\theta^2 T_{11} - T_{12} T_{22}^{-1} T_{21}| =$$

$$= (-1)^m \theta^{m_2 - m_1} |T_{22}| |\theta^2 T_{11} - T_{12} T_{22}^{-1} T_{21}| = 0 \quad (3.4.9)$$

Por lo que las raíces diferentes de cero de (3.4.8) son exactamente las mismas raíces diferentes de cero de:

$$|\theta^2 T_{11} - T_{12} T_{22}^{-1} T_{21}| = 0 \quad (3.4.10)$$

Definamos ahora la matriz de varianzas y covarianzas como:

$$\begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \begin{bmatrix} Y_1' M_1 Y_1 & Y_1' M_1 X_2 \\ X_1' M_1 Y_1 & X_2' M_1 X_2 \end{bmatrix} \quad (3.4.11)$$

Aplicando directamente el resultado (3.4.10) a la matriz (3.4.11) se sigue

$$|\theta^2 Y_1' M_1 Y_1 - Y_1' M_1 X_2 (X_2' M_1 X_2)^{-1} X_2' M_1 Y_1| = 0 \quad (3.4.12)$$

recuérdese ahora que de (3.1.21) y (3.1.24)

$$w_1^* = Y_1' M_1 Y_1 \text{ y } w_{11}^* - w_{11} = Y_1' M_1 X_2 (X_2' M_1 X_2)^{-1} X_2' M_1 Y_1$$

de ahí que (3.4.12) pueda escribirse como

$$|\theta^2 w_{11}^* - (w_{11}^* - w_{11})| = 0 \quad (3.4.13)$$

$$|\theta^2 w_{11}^* - (w_{11}^* - w_{11})| = |w_{11} - (1 - \theta^2) w_{11}^*| =$$

$$= (-1 + \theta^2)^{m_1} |w_{11}^* - \frac{1}{1 - \theta^2} w_{11}| = 0$$

$$\Rightarrow |w_{11}^* - \frac{1}{1 - \theta^2} w_{11}| = 0 \quad (3.4.14)$$

que coincide precisamente con el resultado (3.1.25) que a noso---

tros nos interesa, donde

$$\ell = \frac{1}{1-\theta^2} \quad (3.4.15)$$

$$\Rightarrow 0 \leq \theta^2 = \frac{\ell-1}{\ell} \leq 1 \quad (3.4.16)$$

Bastará entonces encontrar las raíces  $\theta_i$  por correlaciones canónicas y utilizar después la relación (3.4.15) para con estos resultados obtener las pruebas de identificabilidad expuestas en la sección 3 del presente capítulo.

Recordando que el estimador M.C.I.K. necesitaba que la ecuación particular que se estuviese analizando fuese sobreidentificada; podemos garantizar entonces la existencia de este, si la prueba de la condición de rango para identificación se satisface. En el próximo capítulo, suponiendo que ya hemos efectuado estas pruebas, procederemos a enunciar las pruebas de hipótesis basadas en el estimador de M.C.I. de Khazzoom.

CAPITULO IV

## PRUEBAS DE HIPOTESIS BASADAS EN EL ESTIMADOR M.C.I.K.

El problema de la inferencia estadística en econometría puede ser dividido en dos áreas principales. La primera se refiere a la estimación de parámetros estructurales en modelos de interés, y la segunda a la comprobación de la validez de las hipótesis concernientes a esos parámetros o a sus funciones de distribución. Para el segundo caso las hipótesis más discutidas son el tipo i)  $H: \delta = 0$ , la hipótesis de que todos los elementos de  $\delta$  son cero, ii)  $H: \delta = \delta_0$ , la hipótesis de que cada  $\delta_i$  es igual a un valor específico  $\delta_{i0}$ . -- iii)  $H: \alpha' \delta = m$ , que algunas combinaciones lineales de algunos elementos de  $\delta$  sean iguales a una constante específica; y iv)  $H: \delta_j = 0$ , que algunos de los  $\delta_j$ ,  $j$  de ellos, sean cero. Todas estas hipótesis son casos particulares de la hipótesis lineal general:

$$H: Q' \delta = \ell \quad (4.1.1)$$

donde  $\delta = \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$  es un vector  $(G+K, 1)$  de parámetros estructurales,  $Q$  es una matriz  $(m, G+K)$  de constantes ( $m$  el número de restricciones) y  $\ell$  es un vector  $(m, \delta)$  de constantes. La única restricción que existe en la hipótesis lineal general es que  $Q$  tenga rango completo por renglones, es decir, que las funciones lineales de  $\delta$  derivadas de la hipótesis, sean linealmente independientes.

En secciones previas hemos enunciado hasta ahora algunas pruebas de hipótesis concernientes a la condición de rango para identificación. La razón de haberlas presentado previamente a este ca

pitulo es que, pruebas del tipo (4.1.1.) sobre parámetros estructurales necesitan que el modelo este identificado, tanto bajo la hipótesis nula, como bajo la alternativa. En este sentido las pruebas de identificación pueden también ahorrarnos trabajo extra.

En el capítulo anterior se destacó asimismo, que las pruebas de identificabilidad nos dan un criterio de prueba sobre las restricciones cero aún antes de haber estimado los parámetros estructurales. Esta interpretación no es sino un caso particular de la hipótesis lineal general.

En la literatura econométrica existen un buen número de trabajos que nos presentan algunas variedades de como construir pruebas de hipótesis del tipo (4.1.1) basados en las distribuciones asintóticas de los estimadores pertenecientes a la clase- $k$ . En nuestro caso presentaremos de forma análoga, la manera general de construir pruebas de hipótesis basadas en la distribución del estimador de mínimos cuadrados indirectos de Khazzoom (M.C.I.K.). - Dado que los resultados obtenidos sobre el estimador M.C.I.K. son asintóticos, no existe otra forma de construir pruebas de hipótesis más que las que están basadas en las distribuciones normal y  $\chi^2$  cuadrada. Las conocidas pruebas basadas en distribuciones  $t$  y  $F$  para el caso de parámetros con distribuciones exactas, no tienen justificación teórica en el caso asintótico.

El estimador consistente de M.C.I.K. para la primera ecuación estructural está dado por: (véase la 2a. sección del capítulo II).

$$\hat{\delta}_{o1} = [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'y_{o1} \quad (4.1.2)$$

o bien sustituyendo  $y_{o1}$  en (4.1.2), por

$$\hat{\delta}_{o1} = \delta_{o1} + [Z_1'X(X'X)^{-2}X'Z_1]^{-1}Z_1'X(X'X)^{-2}X'u_{o1} \quad (4.1.3)$$

donde

$$T(\hat{\delta}_{o1} - \delta_{o1}) \overset{a}{\sim} N(0, \sigma_{11} \text{plim } \Psi) \quad (4.1.4)$$

$$\Psi = T(X'Z_1)'X'X^{-1}(X'Z_1)'$$

Es claro entonces que

$$TQ(\hat{\delta}_{o1} - \delta_{o1}) \overset{a}{\sim} N(0, \sigma_{11} \text{plim } Q\Psi Q') \quad (4.1.5)$$

$\text{plim } Q\Psi Q'$  es no singular debido a que la matriz  $\Psi$  es positiva definida y  $Q$  tiene rango completo por columnas. Hemos obtenido también un estimador consistente para la varianza dado por:

$$\hat{\sigma}_{11} = \frac{1}{T} \hat{u}_{o1}' \hat{u}_{o1} = \frac{1}{T} u_{o1}' u_{o1} - 2 \frac{1}{T} u_{o1}' Z_1 (\hat{\delta}_{o1} - \delta_{o1}) + (\hat{\delta}_{o1} - \delta_{o1})' \frac{1}{T} Z_1' Z_1 (\hat{\delta}_{o1} - \delta_{o1}) \quad (4.1.6)$$

Para nuestros propósitos el siguiente resultado sobre distribuciones de formas cuadráticas nos será de gran utilidad (ver --- Searle 1971 teorema 2 p.p. 57)

**Teorema.** - Cuando un vector  $y \sim N(0, V)$  entonces  $y' A y \sim \chi^2 [r(A)]$  si y sólo si  $AV$  es idempotente.

Con la extensión de este resultado para distribuciones norma-

les asintóticas se sigue que, bajo la hipótesis nula  $H_0: Q'\delta = \ell$  -

$$\frac{1}{\sigma_{11}} (Q'\hat{\delta}_{0,1} - \ell)' [p \lim Q\Psi Q']^{-1} (Q'\hat{\delta}_{0,1} - \ell) \stackrel{a}{\sim} \chi^2_{\kappa(Q)} \quad (4.1.7)$$

Cabe hacer notar las siguientes consideraciones. Si  $\sigma_{11}$  es desconocido, podemos directamente sustituirlo por su estimador  $\hat{\sigma}_{11}$  y la distribución límite no se alterará debido a que este último es consistente. Por otra parte, en la práctica no es posible obtener  $p \lim Q\Psi Q'$ , pero podemos utilizar en su lugar  $Q\Psi Q'$ , el cual tendrá a su límite en probabilidad conforme  $T$  crezca. Por lo tanto la prueba de hipótesis lineal general a un nivel de significancia  $\alpha$  podemos expresarla como:

$$t(\hat{\delta}_{0,1}, \hat{\sigma}_{11}) = \frac{1}{\hat{\sigma}_{11}} (Q'\hat{\delta}_{0,1} - \ell)' (Q\Psi Q')^{-1} (Q'\hat{\delta}_{0,1} - \ell) \stackrel{a}{\sim} \chi^2_{\alpha, \kappa(Q)} \quad (4.1.8)$$

Se sigue que la región crítica asintótica para probar  $H_0: Q'\delta = \ell$  contra la alternativa  $H_1: Q'\delta \neq \ell$  será

$$t(\hat{\delta}_{0,1}, \hat{\sigma}_{11}) \geq \chi^2_{\alpha, \kappa(Q)} \quad (4.1.9)$$

En particular cuando  $\kappa(Q) = 1$ , la prueba puede expresarse utilizando la distribución normal, esto pasa dado que:

$$t(\hat{\delta}_{0,1}, \hat{\sigma}_{11}) \stackrel{a}{\sim} \chi^2_1 \implies \sqrt{t(\hat{\delta}_{0,1}, \hat{\sigma}_{11})} \stackrel{a}{\sim} N(0, 1) \quad (4.1.10)$$

Para muestras pequeñas es evidente que tenemos un problema, en estos casos se sugiere utilizar pruebas de hipótesis basadas en distribuciones  $t$  y  $F$  como variantes de las distribuciones normal y  $\chi^2$ , que aunque no tengan justificación teórica se construyen en analogía a las pruebas de hipótesis de parámetros cuyas distribuciones son exactas.

**CONSIDERACIONES FINALES**

La intención de este trabajo ha sido en parte, la de presentar de forma más o menos detallada, el método de mínimos cuadrados indirectos como una alternativa más para la estimación de modelos lineales en su forma estructural. Si bien la interpretación dada a estos estimadores por D. J. Khazzoom resulta novedosa y los resultados empíricos presentados por el autor satisfactorios, parece ser que el método no ha sido del todo aceptado. Reflejo de esto es el hecho de que en la literatura no se encuentren artículos -- que continúen, apliquen o comenten estos resultados. A pesar de que los estadísticos presentados sólo tienen propiedades asintóticas (como en la mayoría de los métodos de estimación de modelos estructurales) su comportamiento en pequeña muestra es aceptable, no obstante el autor reconoce sus excelencias en otro sentido: "... se observa que la atracción del método propuesto no es tanto su posible superior comportamiento para pequeña muestra, sino más bien el hecho de que trabaja solamente con el estimador de la forma reducida para derivar los estimadores de los parámetros estructurales. El estimador de la forma reducida es todo lo que el mundo real tiene que dar". Asimismo el autor sugiere la ampliación de su trabajo en el siguiente sentido: "...examinar extensivamente la sensibilidad de las estimaciones con especificaciones alternativas de las restricciones estructurales y tratar con otros aspectos de los experimentos Monte Carlo ... una segunda extensión se relaciona con el aspecto de las variables instrumentales. El uso de la inversa Moore-Penrose abre el camino de la familia de los estimadores instrumentales en el caso sobreidentificado, los cuales tienen la misma estructura pero difieren en la definición

de la norma". A este respecto quisiéramos hacer un comentario extra. Se ha definido un estimador único de M.C.I. para el caso sobreidentificado, dada la unicidad de la inversa Moore-Penrose. Sin embargo podemos obtener otro estimador M.C.I. para el caso sobreidentificado si utilizamos cualquier otra inversa generalizada. Ahora bien, aunque los estimadores obviamente serán diferentes, será interesante contrastar las estadísticas de prueba que de ésta resulten; dado que a este respecto en el artículo de Khazzoom no se menciona ningún criterio de optimalidad al utilizar la inversa Moore-Penrose.

Dado que el estimador M.C.I.K. se ha definido para ecuaciones sobreidentificadas, se ha destacado en este sentido la importancia de las pruebas de identificabilidad como una opción preliminar a la estimación. Si el investigador se decide por esta opción, será conveniente que haga las siguientes consideraciones. Primero, si en la especificación de la ecuación en cuestión se incluye la normalización, convendrá entonces utilizar la prueba de Koopmans y Hood, si por el contrario la normalización no se hace explícita, entonces procederá la prueba de Anderson y Rubin. Segundo, si se cuenta con un paquete estadístico o econométrico de evaluación que incluya correlaciones canónicas, los resultados de las pruebas de identificación podrán agilizarse sobremanera dada la interpretación que se ha dado a éstas en la sección 3.4. Tercero, debe considerarse que las pruebas para la condición de rango de identificación tienen una interpretación equivalente a probar las restricciones cero, lo que ilustra en este sentido, si la

especificación de nuestras ecuaciones ha sido correcta.

Hasta ahora en el trabajo se ha supuesto, o bien que el modelo está identificado, o bien se han expuesto las técnicas para corroborar que efectivamente se satisface esta propiedad. Lo que hasta ahora no se ha considerado es que sucederá si efectivamente no se rechaza la hipótesis de que el modelo está subidentificado. Bajo este enfoque quisiera hacer las últimas anotaciones:

Si el modelo está subidentificado, el camino natural parece -- ser retomar la especificación del modelo y con una nueva definición de las ecuaciones regresar a los procedimientos ya descritos. Ahora bien, ¿pueden las ecuaciones subidentificadas proporcionarnos alguna información sobre la nueva especificación de éstas?. La siguiente interpretación tal vez proporcione alguna idea al respecto.

Hemos ejemplificado hasta ahora con la primera ecuación estructural, que el subconjunto de relaciones que nos interesa para comprobar la identificabilidad de ésta, es el siguiente sistema de ecuaciones de parámetros reducidos.

$$\pi_{11} \beta_{01} + \gamma_{01} = \pi_{01} \quad (5.1.1)$$

$$\pi_{21} \beta_{01} = \pi_{02}$$

y la primera ecuación estará identificada si y sólo si

$$h(\pi_{21}) = G_1$$

si bien estas condiciones se deben satisfacer para los valores -- reales de los parámetros estructurales, sólo podemos conocer sus estimaciones, es decir, contamos con el siguiente sistema de ecuaciones

$$\begin{aligned} \hat{\pi}_{11} \hat{\beta}_{01} + \hat{\gamma}_{01} &= \hat{\pi}_{01} \\ \hat{\pi}_{21} \hat{\beta}_{01} &= \hat{\pi}_{02} \end{aligned} \quad (5.1.3)$$

supongamos entonces que

$$\lambda(\hat{\pi}_{21}) < G_1 \quad (5.1.4)$$

lo que significa que la ecuación en cuestión está subidentificada.

Recordemos la forma del modelo lineal clásico de regresión:

$$y = X\beta + u \quad (5.1.5)$$

para el que, bajo los supuestos usuales se sigue

$$E(y) = \hat{y} = X\hat{\beta} \quad (5.1.6)$$

donde  $\beta$  puede obtenerse por M.C.O., si  $X$  tiene rango completo por columnas. Interpretemos entonces al segundo bloque de ecuaciones en (5.1.3) como un modelo de regresión, i.e:

$$E(\pi_{02}) = \hat{\pi}_{02} = \hat{\pi}_{21} \hat{\beta}_{01} \quad (5.1.7)$$

Es importante hacer notar que (5.1.7) no es un modelo de regresión, dado que  $\hat{\pi}_{21}$  no es una matriz de observaciones en las variables exógenas, sino una matriz de estimaciones de la forma reducida. Sin embargo, esta interpretación nos puede conducir a re

sultados interesantes.

Dado que  $\hat{\Pi}_{21}$  es de rango incompleto podemos incluir a (5.1.7) dentro de la teoría de modelos de regresión de rango incompleto, en donde particularmente nos interesará la discusión sobre funciones estimables y lo concerniente a la prueba de hipótesis lineal general.

Sin entrar en detalles, mencionaremos solamente que una función lineal se define como estimable si es idénticamente igual a alguna función lineal del valor esperado del vector de observaciones de las variables endógenas. Esto significa, (para el modelo (5.1.5)) que  $q'\hat{\beta}$  es estimable si  $q'\hat{\beta} = t'E(y)$ , para algún vector  $t$ . Por tanto la hipótesis  $H: \hat{\beta} = 0$  no puede ser probada dado que  $\hat{\beta}$  no es una función estimable; sin embargo, la hipótesis  $H: \hat{\Pi}_{21}\hat{\beta} = 0$  si puede ser probada por ser una función estimable. En general, combinaciones lineales de funciones estimables son estimables.

Es en este sentido que la ecuación (5.1.7) puede proporcionarnos alguna información sobre las restricciones que podamos imponer en nuestra especificación. La teoría de modelos de regresión de rango incompleto es muy basta, por tal motivo solamente nos concretaremos a citar este enfoque como un tema más de futuras investigaciones. (Para la introducción a modelos de rango incompleto, véase por ejemplo Searle (1971) Cap. 5).

**BIBLIOGRAFIA**

- Anderson, T. W., and H. Rubin, (1949); "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations". Annals of Mathematical Statistics, vol. 20, 1949, p.p. 46-63
- Anderson, T. W., and H. Rubin, (1950); "The Asymptotic Properties of Estimates of the Parameters of a Single Equation in a Complete System of Stochastic Equations". Annals of Mathematical Statistics, vol. 20, 1950, p.p. 570-582.
- Anderson, T. W., (1951), "Estimating Linear Restrictions on Regression Coefficients for Multivariate Normal Distributions". Annals of Mathematical Statistics, vol. 22, p. p. 327-351.
- Anderson, T. W., (1958); An Introduction to Multivariate Statistical Analysis; New York, Wiley.
- Basmann, R.L., (1960); "On Finite Sample Distribution of GCL Identifiability Test Statistics". Journal of the American Statistical Association, vol. 55, 1960, p.p.650-659.
- Christ, C. F., (1966); Modelos y Métodos Económicos, Ed. Limusa, España.
- Dhrymes, P.J., (1970); Econometrics: Statistical Foundations and Applications; Evanston, Harper and Row. 1970.

- Farebrother, R. W., and N. E. Savin, (1973); "Testing the Rank Conditions in Klein's Model I", in Discussion Papers in Economics, Econometrics and Decision Theory, University of Manchester.
- Fisher, F. M., (1959); "Generalization of the Rank and Order Conditions for Identifiability". *Econometrica*, vol. 27 - 1959, p.p. 431-447.
- Fisher, F. M., (1966), *The Identification Problem in Econometrics*; New York, McGraw-Hill.
- Guill, L. (1978); "Testing for Identifiability"; Discussion Paper, Department of Econometrics, University of Manchester.
- Goldberger, A. S., (1964); *Econometric Theory*; New York, Wiley.
- Khazzoon, J. D., (1976); "An Indirect Least Squares Estimator for overidentified Equations". *Econometrica*, vol. 44, -- 1976, p.p. 741-750.
- Koopmans, T. C., and H. Rubin (1950); "The Estimation of Simultaneous Economic Relationships", in *Studies in Econometric Methods*, W.C. Hood and T.C. Koopmans (Eds), New York, Wiley.
- Malinvaud, E. (1963); *Métodos Estadísticos de la Econometría*, ed. Ariel, Barcelona.

- Rao, C. R., and S. K. Mitra, (1971); Generalized Inverse of -  
Matrices and its Applications, New York, Wiley.
- Sabau, H.C.L., (1979); "Test of Overidentifying Restriccions  
in the Simultaneous Equation Model: A Simulation Study";  
University of Manchester, un published.
- Sabau, H.C.L., (1981); "Notas Docentes"; curso de Economía ---  
Aplicada II: CIDE, no publicado.
- Savin, N.E., (1973); "On Testing the Zero Restriccions and --  
The Rank Conditions for Identification in Limited Infor-  
mation Maximun Likelihood Estimation". In Discussion Pa-  
pers in Economics, Econometrics and Decision Theory; Uni-  
versity of Manchester.
- Sawa, T. (1972); "Finite-Sample Properties of the k-class Es-  
timators". Econometrica, vol. 40, # 4, 1972, p.p.653-680.
- Searle, S. R., (1971), Linear Models, New York, John Wiley --  
and Sons, 1971.
- Theil, H. (1971), Principles of Econometrics, New York, John  
Wiley and Sons, 1971.
- Wallis, K. F., (1972); Introducción a la Econometría, Alianza  
Universidad, 1972.