Capítulo 10

Conclusiones.

Todo proceso científico nos permite generar conocimiento a la vez que deja siempre abiertos un gran número de interrogantes que permitirán nuevamente el inicio de otras investigaciones y supondrán, sin duda, nuevos conocimientos. En la Introducción comentábamos que en la Estadística se relacionan íntimamente los desarrollos teóricos y su aplicación empírica: la experiencia práctica sugiere la necesidad de nuevos instrumentos teóricos que permitirán extraer más información de los datos. Esta doble caracterización de la Estadística la trasladamos a los objetivos de nuestra tesis, a la vez que también se hace presente en el diseño del proceso de investigación que nos conduce hasta ellos.

Ahora, una vez finalizada la investigación propuesta en el Capítulo 1, queremos explicitar, no sólo hasta que punto hemos logrado alcanzar los objetivos planteados, sino también, todo el conocimiento que el desarrollo de este trabajo ha generado, así como las nuevas áreas donde ampliar nuestra investigación. Siguiendo con la dualidad establecida al definir los objetivos vamos a comentar, en primer lugar, los resultados teóricos obtenidos con nuestra tesis, para después pasar a enumerar los resultados empíricos; por último mencionaremos las líneas de investigación que este trabajo deja abiertas y que serán objeto de investigaciones futuras.

10.1 Resultados obtenidos a nivel teórico:

Antes de comentar los logros obtenidos con el desarrollo de esta tesis, recordaremos los objetivos que nos planteábamos a nivel teórico:

> "Simplificar el proceso de identificación y estimación del modelo SETAR (*"Self-Exciting AutoRegressive model"*). Obtener e implementar un esquema algorítmico que permita estimar los órdenes de los procesos autoregresivos que componen el modelo, a la vez que determine para cada uno de los procesos autoregresivos cuales son los retardos significativos. El algoritmo propuesto, a diferencia del de Thanoon (1990), debe permitir trabajar con modelos de más de dos regímenes y no presentar limitación sobre el número de variables regresoras en cada proceso autoregresivo."

La investigación realizada ha permitido ir más allá del objetivo planteado, de manera tal que, el esquema algorítmico para la estimación de los órdenes de los procesos autoregresivos inicialmente propuesto, ha sido incluido dentro de una nueva metodología MIEC (Metodología para la Identificación y Estimación de Coeficientes) diseñada para la identificación y estimación de los modelos SETAR.

Vamos a enumerar y detallar a continuación los principales resultados obtenidos:

1. **Elaboración de un nuevo algoritmo AIEC de selección de regresores** (Algoritmo para la Identificación y Estimación de Coeficientes). Este algoritmo nos permite seleccionar para cada uno de los regímenes del modelo el conjunto de regresores consiguiendo así, determinar el orden de cada uno de los procesos. Su diseño podía haberse planteado considerando todos los subconjuntos de posibles regresores, estimando para cada uno de ellos los coeficientes del modelo y, finalmente, utilizando un criterio de información que nos permitiera elegir entre todos los modelos posibles el que supone la optimización de dicho criterio. Pero, gracias al estudio de las características del criterio de información de Akaike y de otros criterios de información realizado en el Capítulo 4, hemos obtenido un resultado (Teorema 24) que permite simplificar enormemente el número de iteraciones del algoritmo.

2. Relación entre el MAICE y la selección de variables autoregresoras en un modelo SETAR. El resultado obtenido (presentado y demostrado en el Capítulo 6), nos asegura que el AIC del modelo SETAR es mínimo en cada régimen cuando consideramos un conjunto de regresores de retardos consecutivos.

Teorema 24:

Sea $\{X_t\}$ $t = 1, \cdots, N$ *una serie temporal que sigue un modelo autoregresivo* $SETAR$ *con* l *regímenes cada uno de ellos de orden máximo* p*. El* $MAICE$ *(mínimo valor del AIC estimado) se obtiene para un modelo* $SETAR(l; k_1, k_2, \cdots, k_l)$ *donde los órdenes* k_i $(k_i \leq p)$ *de cada uno de los procesos autoregresivos coincide con el número de variables autoregresoras de cada proceso.*

Como ya hemos comentado en el punto anterior el teorema 24 nos permite reducir ostensiblemente el número de casos a considerar en el algoritmo AIEC, obteniéndose así un proceso computacionalmente más sencillo.

3. El teorema 24 sigue siendo válido si consideramos el BIC (Criterio de Información Bayesiano**) o el AICc (**Criterio de información de Akaike corregido**).**

4. Estimación automática del parámetro de retardo d : Nuestra propuesta metodológica MIEC incorpora el test TAR-F de Tsay para determinar de forma automática la estimación del parámetro de retardo.

5. Estimación del valor umbral de los modelos SETAR de dos regímenes. Para modelos SETAR de dos regímenes hemos automatizado también el proceso de estimación del umbral, el algoritmo diseñado se basa nuevamente en la optimización del criterio de información elegido. Su implementación se ha llevado a cabo mediante un código que puede incorporarse como subrutina al programa que permite la estimación de los órdenes de los procesos autoregresivos, la selección de los regresores y la estimación de los coeficientes.

6. Unicidad de criterio en todos los procesos: La metodología que hemos elaborado realiza la estimación e identificación del modelo de forma conjunta a partir de un único proceso algorítmico gobernado en todas sus etapas por una misma ley, la minimización del criterio de información elegido. Esta unicidad de criterio supone una mejora conceptual del diseño algorítmico.

10.2 Resultados obtenidos a nivel empírico:

Nuestra trabajo se iniciaba motivado por el estudio de la volatilidad asociada a la rentabilidad de las acciones y, más concretamente, por nuestro interés en modelizar esta característica de los mercados financieros a partir del modelo SETAR. En el área de la investigación empírica planteábamos en el Capítulo 1 los siguientes objetivos:

> "Elegir el estimador de la volatilidad que proporcione la mejor estimación y predicción de la volatilidad en un mercado determinado".
>
> "Seleccionar un modelo, fácilmente identificable, que explique el comportamiento histórico de la volatilidad y permita predecirla en períodos futuros".

Ambos objetivos han sido alcanzados, pero además el desarrollo de la investigación nos ha llevado a considerar otros aspectos relacionados con la naturaleza de los datos como son, su frecuencia y la conveniencia de considerar el exceso de rentabilidad en el cálculo de la volatilidad.

Los resultados obtenidos para cada uno de los objetivos, así como las conclusiones a las que hemos llegado sobre la naturaleza de los datos se presentan a continuación:

1. **Definición de la medida de volatilidad:** La volatilidad no es una medida directamente observable, y es por tanto necesario estimarla a partir de la rentabilidad; pero el estudio de la volatilidad nos llevó a comprobar que no existe una única definición, es más, para los operadores financieros este término adquiere diversos significados según sea el papel que desempeñe en el mercado. Desde un punto estadístico, estas maneras diversas de entender la volatilidad suponen la existencia de diferentes estimadores de la volatilidad con características distintas tanto a nivel de estructura como de capacidad predictiva. En esta tesis entendemos la volatilidad como sigue:

"La volatilidad es una medida de la intensidad de los cambios aleatorios o impredecibles en la rentabilidad o en el precio de un título; en la representación gráfica de una serie histórica de rendimientos se asocia la volatilidad con la amplitud de las fluctuaciones del rendimiento, tanto es que se consideren en valor absoluto, como en desviaciones alrededor de un valor medio " (Definición 1).

2. Elección de la desviación absoluta respecto a la media como estimador de la volatilidad: Después de un análisis exhaustivo de cada uno de los estimadores propuestos para medir la volatilidad, hemos optado por una medida que sea consistente con el marco teórico del problema y permita entender la volatilidad como un proceso que evoluciona en el tiempo. El estimador finalmente elegido es el propuesto por Schwert (1990) y Schwert y Seguin (1990) que se define como sigue (Definición 11, capítulo 2).

Desviación absoluta respecto a la media: *Estimamos la volatilidad σ_t a partir de la siguiente expresión:*

$$\widehat{\sigma}_t = \sqrt{\frac{\pi}{2}}\,|R_t - \mu_t|$$

donde R_t es el exceso de rentabilidad y μ_t es tratado como un parámetro y estimado a partir de la media muestral de R_t.

- El comportamiento de la serie de volatilidades construida a partir de $\widehat{\sigma}_t$ es no lineal.
- La dinámica del proceso $\{\widehat{\sigma}_t\}$ puede ser explicada a partir de un modelo SETAR.

3. Frecuencia mensual de los datos. Los sistemas de interconexión bursatil permiten la generación de datos de manera instantánea, este hecho que permite disponer de gran cantidad de observaciones, no debe condicionar la unidad temporal considerada al definir el proceso temporal. En este trabajo hemos decidido utilizar datos mensuales por diversas razones:

- Existen modelos específicos para trabajar con datos minuto a minuto ("tick by tick") o datos observados cada hora (modelos intradía). Los modelos SETAR no han sido utilizados habitualmente para frecuencias inferiores a un día, aunque McCulloch y Tsay (2001) introducen modelos TAR para ajustar el tiempo entre la variación de los precios intradía.

- Si tomamos como unidad la volatilidad diaria (Márquez y Villazón, 1999), debemos tener en cuenta que, debido a las características propias de negociación en los mercados, estos permanecen cerrados el sábado y domingo, por lo que el lunes es un día especial pues tiende a recoger toda la información que se ha generado en los dos días anteriores; en este sentido no es correcto considerar el lunes como cualquier otro día. Una situación similar es la que ocurre después de un día festivo.

- Cuando se utiliza la semana como unidad (ver trabajo de Márquez y Villazón, 1997) es necesario un proceso de interpolación en aquellos casos en que aparece un día festivo o más. Una situación excepcional puede provocar el cierre del mercado por durante varios días y esto supone un poblema para medir la volatilidad del período considerado. También es bien conocido el efecto que en este tipo de series puede tener el período correspondiente a Semana Santa y Navidades.

- Debemos tener también en cuenta que si trabajamos con datos de elevada frecuencia (por ejemplo diaria) las estimaciones presentan una elevada correlación serial negativa (Figlewski 1997)

- Finalmente, es importante considerar que el horizonte de predicción queda condicionado por la frecuencia de los datos, así la utilización de datos mensuales nos va a permitir realizar predicciones a 1 año. Si la frecuencia de los datos es mayor, el horizonte de predicción debe ser más próximo.

Estimar la volatilidad para un período mensual conlleva, desde el punto de vista estadístico, una mayor homogeneidad en la naturaleza de los datos. Las incidencias que hemos comentado pueden tener una gran influencia si trabajamos con datos de frecuencia elevada, pero quedan diluidas al considerar un período mensual.

4. Cálculo de la volatilidad a partir del exceso de rentabilidad.

Otra particularidad de nuestro trabajo es que el estimador de la volatilidad elegido, desviación absoluta respecto a la media, no se calcula directamente a partir de la rentabilidad obtenida por la variación de los precios, r_t, sino con el denominado "exceso de rentabilidad" R_t. En el mercado financiero español no es práctica habitual trabajar con el exceso de rentabilidad pero, aunque no existe una gran diferencia entre el comportamiento de ambas series $\{r_t\}$ y $\{R_t\}$, conceptualmente nos parece más adecuado la utilización de la última pues permite obtener una medida más precisa de la volatilidad al recoger la intensidad de los cambios debidos al riesgo.

5. Elección como activo libre de riesgo del tipo LIBOR sobre el Euro a 1 mes. Como nuestros datos de partida son rentabilidades mensuales, es necesario que el activo libre de riesgo elegido calcule la rentabilidad para el mismo período. En el mercado americano es habitual calcular el exceso de rentabilidad a partir del tipo de las Letras del Tesoro a 1 mes, al no emitirse en el mercado monetario español con este vencimiento, hemos elegido como tipo de referencia el del LIBOR sobre el Euro a 1 mes.

6. Estimación e identificación de un SETAR (2; 2,8) para la volatilidad asociada a la rentabilidad del IBEX 35 (1990-2000).

Utilizando nuestra propuesta metodológica (MIEC) hemos identificado y estimado un modelo SETAR (2; 2,8) para la serie $y_t = \ln(\frac{w_t}{w_{t-1}})$, obtenida como la variación relativa mensual de la volatilidad w_t asociada a la rentabilidad del IBEX 35 en el período enero de 1990 a diciembre del 2000, la expresión del cual es la siguiente:

$$\mathbf{y}_t = \underset{(0,4035)}{-0,2195} - \underset{(0,1899)}{0,6773}\, y_{t-1} - \underset{(0,1958)}{0,2403}\, y_{t-2} + \varepsilon_t^{(1)} \qquad si \qquad y_{t-6} \leq -0,3005$$

$$\mathbf{y}_t = \underset{(0,1734)}{-0,3215} - \underset{(0,0879)}{0,6705}\, y_{t-1} - \underset{(0,1155)}{0,7869}\, y_{t-2} - \underset{(0,1333)}{0,7988}\, y_{t-3} - \underset{(0,1437)}{0,7275}\, y_{t-4} - \underset{(0,1438)}{0,6893}\, y_{t-5}$$

$$- \underset{(0,1621)}{0,1736}\, y_{t-6} - \underset{(0,1375)}{0,4505}\, y_{t-7} - \underset{(0,1147)}{0,2884}\, y_{t-8} + \varepsilon_t^{(2)} \qquad si \qquad y_{t-6} > -0,3005$$

$\mathrm{var}(\varepsilon_t^{(1)}) = 1,673184$; $\mathrm{var}(\varepsilon_t^{(2)}) = 1,089787$;

AIC global $= 55,24158$; varianza residual global $= 1,315618$.

El análisis del modelo $SETAR(2;2,8)$ nos permite llegar a las siguientes conclusiones:

- El modelo estimado es un modelo robusto cuyos parámetros estructurales se mantienen constantes aunque cambie el tamaño de la muestra.
- La dinámica del proceso está gobernada por dos ecuaciones distintas, la elección de una u otra depende del valor de la variable 6 periodos atrás, y_{t-6}.
- El primer régimen explica, mediante un proceso $AR(2)$, aquellas observaciones y_t tal que 6 períodos atrás se produjo un descenso importante de la volatilidad, $y_{t-6} \leq -0,3005$.
- En el segundo régimen un proceso autoregresivo de orden 8, $AR(8)$, determina el comportamiento de la variable y_t. Este régimen explica el comportamiento de aquellas observaciones y_t tal que 6 períodos atrás se produjo un descenso suave de la volatilidad si $-0,3005 < y_{t-6} \leq 0$, o un aumento de la volatilidad si $y_{t-6} > 0$.
- Los órdenes de los procesos autoregresivos $\widehat{k}_1 = 2$ y $\widehat{k}_2 = 8$ no són muy elevados lo que conduce a un modelo parsimonioso y de gran simplicidad.

- Podemos reinterpretar el modelo en términos de la medida de volatilidad w_t,

$$w_t = \exp(y_t) \cdot w_{t-1}.$$

si disponemos del valor de la volatilidad en el período anterior, w_{t-1}, y conocemos a partir del modelo el valor de y_t, es facil obtener el valor de la volatilidad para el período t, w_t.

- Los cambios de régimen se pueden interpretar también en términos de w_t, así cuando

$$y_{t-6} \leq -0,3005 \Rightarrow w_{t-6} \leq 0,74045 w_{t-7} = 74,045\% \ w_{t-7}$$

, es decir, hay un decrecimiento de la volatilidad 6 meses atrás importante. En cambio, el segundo régimen explica la volatilidad cuando consideramos que hace medio año se produjo un aumento de volatilidad o un descenso suave $w_{t-6} > 0,74045 w_{t-7}$.

- El "skeleton" asegura la estabilidad a largo plazo del sistema: Si no existiesen perturbaciones ($\varepsilon_t = 0, \forall t$) el sistema evolucionaría hacia un punto límite $-0,0570$ lo que supone una variación relativa negativa, y por tanto que la volatilidad de un período t decrece, aunque muy ligeramente, respecto al período anterior:

$$Si \quad y_t \longrightarrow -0,0570 \Rightarrow \ln(\frac{w_t}{w_{t-1}}) \longrightarrow -0,0570 \Rightarrow w_t \longrightarrow 0,9446 w_{t-1}$$

Si el punto límite tuviese signo positivo, el crecimiento de la volatilidad sería constante y esto llevaría a una situación inestable.

- El ajuste en muestra es bueno, y el ajuste fuera de muestra, "pedicción genuina", resulta muy acertado en la predicción de la tendencia futura, pero no lo es tanto en la predicción del valor. Este particular no es inconveniente para que los operadores financieros obtengan de este modelo información muy útil ya que, conocer con un elevado grado de confianza si la volatilidad aumentará o disminuirá en los próximos meses, va a permitir elaborar estrategias para la valoración de opciones de compra y venta.

10.3 Nuevas líneas de investigación.

Al avanzar en los capítulos de este trabajo hemos ido comentando como la investigación en algunas de las áreas relacionadas con los modelos SETAR no está todavía completada, esta situación es debida tanto a la juventud del modelo, como también al interés que su estudio y aplicación despierta.

Entre las diferentes direciones hacia donde apuntan las nuevas investigaciones podemos señalar:

- Utilización de técnicas bayesianas para la estimación del umbral (Pfann, Schotman and Tschernig, 1996) y la determinación automática del número de regímenes.
- En cuanto a la distribución del parámetro umbral, destacamos la propuesta de Hansen (2000) que elabora, utilizando técnicas boostrap, una teoría basada en que la distribución del umbral es no estandar; esta teoría le permite testar la existencia de uno o varios umbrales, elaborar intervalos o regiones de confianza para el umbral estimado, y también determinar la distribución asintótica de los coeficientes. Gonzalo y Wolf (2001) abordan este tema utilizando técnicas de submuestreo en lugar de técnicas boostrap.
- Regiones de confianza para la predicción (Hansen, 2000, Gonzalo y Pitarakis, 2002)
- Tratamiento de observaciones atípicas y su efecto en la estimación por mínimos cuadrados (Chan y Cheung, 1994).
- Modelos SETAR multivariantes (Tiao y Tsay, 1993, Tsay 1998).

El interés que nos lleva al estudio de los modelos SETAR se enmarca en el campo de su aplicación a las series financieras y, más en concreto, en el estudio de la volatilidad, por este motivo nuestras investigaciones futuras se van a centrar principalmente en los modelos SETAR multivariantes, y también en el desarrollo de una interficie que permita realizar el proceso de identificación y estimación de los modelos SETAR de una manera sencilla, rápida y fiable.

Los modelos multivariantes para series temporales nos parecen muy útiles pues permiten el análisis de fenómenos complejos que evolucionan en el tiempo, así para determinar de forma más precisa el comportamiento de la volatilidad en los mercados financieros es necesario tener en cuenta otras características del mercado que se relacionan con la volatilidad pero que no podemos aislar cuando realizamos el estudio a partir de su evolución histórica. Los modelos SETAR multivariantes nos van a permitir analizar la serie de volatilidades en diferentes mercados y relacionar la evolución de estas con otros procesos como son la serie temporal de cotizaciones de determinados valores bursátiles, la serie temporal de tipos de interés e incluso la serie que nos informa de la evolución del volumen negociado. Todas estos procesos temporales son una valiosa fuente de información exógena que va a permitir mejorar el conocimiento de la serie de volatilidades.

Otro área en el que nos parece muy interesante desarrollar nuestro trabajo es completando la automatización del proceso de identificación y estimación de los modelos SETAR, para poder finalmente lograr conseguir un código que, a través de una interficie desarrollada a partir de menús despleglabes, permita realizar de manera rápida y sencilla el proceso de modelización. En esta interficie se incorporarán además de los códigos de estimación de los parámetros estructurales propuestos en este trabajo, salidas gráficas que permitirán una inspección visual del proceso y de sus características; también se incluirán diferentes tests de linealidad, así como tests para detectar la ireversibilidad del proceso. Un punto importante a considerar es la validación del modelo, para ello deben incorporarse pruebas de hipótesis (tests) y gráficos; finalmente deberemos tener en cuenta los últimos avances en la predicción a m-pasos para poder ofrecer así regiones de confianza para las predicciones obtenidas. El diseño de este instrumento informático debe tener en cuenta la necesidad de ir incorporando los nuevos datos que se van generando, esto va a suponer la reestimación del modelo y la actualización de las predicciones a m-pasos que ya han sido realizadas.

Apéndice A

Caracterización de $X'X$ y $X_u'X_u$.

A.1 Definición del problema.

Supongamos un proceso temporal $\{\mathbf{X}_t\}$ que sigue un modelo autoregresivo de orden p, $AR(p)$.

$$\mathbf{X}_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + \cdots + a_p X_{t-p} + \varepsilon_t$$

la estimación de los coeficientes supone la resolución del siguiente sistema de ecuaciones:

$$Y = X\beta + \varepsilon \tag{A.1}$$

donde

$$\beta = \begin{pmatrix} a_0, & a_1, & a_2, & \cdots, & a_p \end{pmatrix}'$$

$$X = \begin{pmatrix} 1 & x_{t-1} & x_{t-2} & \cdots & x_{t-p} \\ 1 & x_{t+1-1} & x_{t+1-2} & \cdots & x_{t+1-p} \\ 1 & x_{t+2-1} & x_{t+2-2} & \cdots & x_{t+2-p} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{pmatrix}$$

$$Y = \begin{pmatrix} x_t, & x_{t+1}, & x_{t+2}, & \cdots, & x_n \end{pmatrix}'$$

$$\varepsilon = \begin{pmatrix} \varepsilon_t, & \varepsilon_{t+1}, & \varepsilon_{t+2}, & \cdots, & \varepsilon_n \end{pmatrix}'$$

x_t son las observaciones del proceso $\{X_t\}$.

Así los valores estimados para β, por el método de mínimos cuadrados condicionados, se obtienen como sigue

$$\widehat{\beta} = (X'X)^{-1}(X'Y)$$

A partir de la ecuación A.1 podemos obtener todas las posibles autoregresiones construidas a partir de los subconjuntos de $\{X_{t-1}, X_{t-2}, \cdots, X_{t-p}\}$. El número de posibles subconjuntos es $2^p - 1$ y por tanto podemos plantearnos $2^p - 1$ autoregresiones.

De esta manera para cada régimen debemos resolver $2^p - 1$ sistemas de ecuaciones de la forma

$$Y_u = X_u \beta_u + \varepsilon \tag{A.2}$$

donde el subíndice $u = 1, \cdots, 2^p - 1$ nos informa del sistema de ecuaciones en que nos encontramos, es decir, nos permite identificar el subconjunto de variables regresoras que estamos considerando. Para un valor concreto u las variables regresoras son el subconjunto X_{t-p_1}, $X_{t-p_2}, \cdots, X_{t-p_k},$ donde $p_1,\ p_2, \cdots,\ p_k$ estan incluidos en el conjunto de los enteros $\{1, 2, \cdots, p\}$

con $1 \leq p_1 < p_2 < \cdots < p_k \leq p$. La matriz X_u, que es una submatriz obtenida a partir de la matriz X, se construye como sigue:

$$X_u = \begin{pmatrix} 1 & x_{t-p_1} & x_{t-p_2} & \cdots & x_{t-p_k} \\ 1 & x_{t+1-p_1} & x_{t+1-p_2} & \cdots & x_{t+1-p_k} \\ 1 & x_{t+2-p_1} & x_{t+2-p_2} & \cdots & x_{t+2-p_k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n-p_1} & x_{n-p_2} & \cdots & x_{n-p_k} \end{pmatrix}$$

Y_u es un subvector del vector Y

$$Y_u = (\ x_{t-p_1+1}, \quad x_{t+1-p_1+1}, \quad x_{t+2-p_1+1}, \quad \cdots, \quad x_{n-p_1+1} \ )'$$

y β_u el vector de coeficientes se puede expresar como

$$\beta_u = (\ a_{p_0}, \quad a_{p_1}, \quad a_{p_2}, \quad \cdots, \quad a_{p_k} \ )'$$

La solución del sistema de ecuaciones tiene ahora la forma:

$$\beta_u = (X_u' X_u)^{-1} (X_u' Y_u)$$

Conscientes de que nuestro interés se centra en obtener los $\widehat{\beta}_u$, vamos a construir para cada valor de u la matriz $X_u'X_u$ y el vector $X_u'Y_u$. Tanto la matriz como el vector pueden obtenerse de la matriz $X'X$ y el vector $X'Y$, ya que los primeros son, respectivamente, una submatriz y un subvector de estos últimos. La propia descomposición binaria del valor u nos permite conocer qué coeficientes de la matriz $X'X$ y el vector $X'Y$, formarán la correspondiente submatriz y subvector.

Para la resolución de los sistemas de ecuaciones de la forma A.2 es importante conocer las características de las matrices de la forma $X_u'X_u$; en primer lugar analizamos las características de la matriz $X'X$, se puede demostrar a partir de su relación con la matriz de autocovarianzas Γ_p definida para el modelo autoregresivo $AR(p)$, que $X'X$ es simétrica y semidefinida positiva.

Utilizando las propiedades de las matrices Γ_p y $X'X$ demostraremos que $X_u'X_u$ es también simétrica y semidefinida positiva.

A.1.1 Definiciones previas.

Definition 1 *Sea A una matriz simétrica y real, diremos que es definida positiva si y sólo si, sus valores propios son positivos.*

Definition 2 *Sea A una matriz simétrica y real, diremos que es definida positiva si y sólo si, existe una matriz Q no singular tal que $A = QQ'$.*

Definition 3 *Si A es una matriz simétrica tal que $\forall x$ vector $x \neq 0, \quad x'Ax > 0$, se denomina matriz definida positiva.*

Definition 4 *Si A es una matriz simétrica tal que $\forall x$ vector $x \neq 0, \quad x'Ax \geq 0$, se denomina matriz semidefinida positiva.*

Propiedad*: Cualquier submatriz cuya diagonal principal sea un subconjunto de la diagonal principal de una matriz definida positiva (semidefinida positiva) es definida positiva (semidefinida positiva). En particular, cualquier submatriz cuadrada principal de una matriz definida positiva (semidefinida positiva) es definida positiva (semidefinida positiva).*

A.2 Resultados:

A.2.1 $X'X$ es simétrica y semidefinida positiva.

Con objeto de simplificar la demostración y sin perder la generalidad, podemos considerar que las variables $\{X_{t-1}, X_{t-2}, \cdots, X_{t-p}\}$ estan centradas respecto a la media, por lo que desaparece el término independiente de la ecuación A.1, ahora

$$\beta = (\ a_1, \quad a_2, \quad \cdots, \quad a_p \ )'$$

$$X = \begin{pmatrix} x_{t-1} & x_{t-2} & \cdots & x_{t-p} \\ x_{t+1-1} & x_{t+1-2} & \cdots & x_{t+1-p} \\ x_{t+2-1} & x_{t+2-2} & \cdots & x_{t+2-p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{pmatrix}$$

por lo que

$$X'X = \begin{pmatrix} \sum_{i=t}^{n} (x_{i-1})^2 & \sum_{i=t}^{n} (x_{i-1} \cdot x_{i-2}) & \cdots & \sum_{i=t}^{n} (x_{i-1} \cdot x_{i-p}) \\ \sum_{i=t}^{n} (x_{i-2} \cdot x_{i-1}) & \sum_{i=t}^{n} (x_{i-2})^2 & \cdots & \sum_{i=t}^{n} (x_{i-2} \cdot x_{i-p}) \\ \sum_{i=t}^{n} (x_{i-3} \cdot x_{i-1}) & \sum_{i=t}^{n} (x_{i-3} \cdot x_{i-2}) & \cdots & \sum_{i=t}^{n} (x_{i-3} \cdot x_{i-p}) \\ \vdots & \vdots & \vdots & \vdots \\ \sum_{i=t}^{n} (x_{i-p} \cdot x_{i-1}) & \sum_{i=t}^{n} (x_{i-p} \cdot x_{i-2}) & \cdots & \sum_{i=t}^{n} (x_{i-p})^2 \end{pmatrix}$$

si dividimos por el número de observaciones n

$$(1/n) \sum_{i=t}^{n} (x_{i-r} \cdot x_{i-s}) = Cov(X_{t-r}, X_{t-s}) = \widehat{\gamma}(t-s, t-r) \quad donde \quad r, s \in \{1, 2, \cdots, p\}$$

obtenemos la estimación de la función de autocovarianza muestral $\gamma(\cdot, \cdot)$. Cuando el proceso $\{\mathbf{X}_t\}$ es estacionario entonces:

$$\widehat{\gamma}(t-s, t-r) = \widehat{\gamma}(r-s) = \widehat{\gamma}(s-r)$$

Defiimos la matriz de covarianzas $\widehat{\Gamma}_p$

$$\widehat{\Gamma}_p = \begin{pmatrix} \widehat{\gamma}(0) & \widehat{\gamma}(1) & \widehat{\gamma}(2) & \cdots & \widehat{\gamma}(p-1) \\ \widehat{\gamma}(-1) & \widehat{\gamma}(0) & \widehat{\gamma}(1) & \cdots & \widehat{\gamma}(p-2) \\ \widehat{\gamma}(-2) & \widehat{\gamma}(-1) & \widehat{\gamma}(0) & \cdots & \widehat{\gamma}(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\gamma}(-(p-1)) & \widehat{\gamma}(-(p-2)) & \widehat{\gamma}(-(p-3)) & \cdots & \widehat{\gamma}(0) \end{pmatrix}$$

la matriz $\widehat{\Gamma}_p$ es una matriz simétrica y semidefinida positiva[1]. También es sencillo demostrar que si

$$\widehat{\gamma}(0) > 0 \Rightarrow \det(\widehat{\Gamma}_p) > 0 \Rightarrow \widehat{\Gamma}_p \quad es \quad definida \quad positiva$$

A.2.2 $X_u'X_u$ es simétrica y semidefinida positiva.

Como ya hemos comentado el proceso de estimación utilizado se basa en la resolución de p sistemas de ecuaciones

$$Y_u = X_u\beta_u + \varepsilon \quad donde \quad u = 2^0, 2^0 + 2^1, \cdots, \sum_{i=0}^{p-1} 2^i$$

esto nos lleva a calcular

$$\beta_u = (X_u'X_u)^{-1}(X_u'Y_u)$$

Ya vimos en el Capítulo 7 que el valor u nos permite conocer las variables seleccionadas $\{X_{t-1}, X_{t-i_1}, X_{t-i_2}, \cdots, X_{t-i_k}\}$ dentro del conjunto $\{X_{t-1}, X_{t-2}, \cdots, X_{t-p}\}$, el teorema 25 nos asegura que los retardos de las variables seleccionadas serán consecutivos.

Las matrices $X_u'X_u$ obtenidas al plantear los diferentes sistemas de ecuaciones son submatrices de $X'X$, obtenidas al eliminar determinadas últimas filas y columnas. Al igual que la matriz $X'X$ se relacionaba íntimamente con la matriz $\widehat{\Gamma}_p$, cada matriz $X_u'X_u$ se corresponde con una matriz $\widehat{\Gamma}_{\{u\}}$ cuya diagonal principal es un subcojunto de la diagonal principal de $\widehat{\Gamma}_p$ de la siguiente forma:

Si consideramos la matriz $\widehat{\Gamma}_p$

$$\widehat{\Gamma}_p = \begin{pmatrix} \widehat{\gamma}(0) & \widehat{\gamma}(1) & \widehat{\gamma}(2) & \cdots & \widehat{\gamma}(p-1) \\ \widehat{\gamma}(1) & \widehat{\gamma}(0) & \widehat{\gamma}(1) & \cdots & \widehat{\gamma}(p-2) \\ \widehat{\gamma}(2) & \widehat{\gamma}(1) & \widehat{\gamma}(0) & \cdots & \widehat{\gamma}(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\gamma}((p-1)) & \widehat{\gamma}((p-2)) & \widehat{\gamma}((p-3)) & \cdots & \widehat{\gamma}(0) \end{pmatrix}$$

[1] Ver demostración en Brockwell y Davis (1997).

a partir de ella obtenemos las submatrices

$$\widehat{\Gamma}_{u=1} = \begin{pmatrix} \widehat{\gamma}(0) \end{pmatrix}$$

$$\widehat{\Gamma}_{u=3} = \begin{pmatrix} \widehat{\gamma}(0) & \widehat{\gamma}(1) \\ \widehat{\gamma}(1) & \widehat{\gamma}(0) \end{pmatrix}$$

$$\widehat{\Gamma}_{u=7} = \begin{pmatrix} \widehat{\gamma}(0) & \widehat{\gamma}(1) & \widehat{\gamma}(2) \\ \widehat{\gamma}(1) & \widehat{\gamma}(0) & \widehat{\gamma}(1) \\ \widehat{\gamma}(2) & \widehat{\gamma}(1) & \widehat{\gamma}(0) \end{pmatrix}$$

$$\vdots$$

$$\widehat{\Gamma}_{u=}\sum_{i=0}^{p-2} 2^i = \begin{pmatrix} \widehat{\gamma}(0) & \widehat{\gamma}(1) & \widehat{\gamma}(2) & \cdots & \widehat{\gamma}(p-2) \\ \widehat{\gamma}(1) & \widehat{\gamma}(0) & \widehat{\gamma}(1) & \cdots & \widehat{\gamma}(p-3) \\ \widehat{\gamma}(2) & \widehat{\gamma}(1) & \widehat{\gamma}(0) & \cdots & \widehat{\gamma}(p-4) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \widehat{\gamma}((p-2)) & \widehat{\gamma}((p-3)) & \widehat{\gamma}((p-4)) & \cdots & \widehat{\gamma}(0) \end{pmatrix}$$

Por la **propiedad** demostrada en A.1.1 podemos asegurar que todas las matrices $\widehat{\Gamma}_u$ son semidefinidas positivas y si $\widehat{\gamma}(0) > 0$ podemos asegurar que son definidas positivas.