

Universidad de Valparaíso
Facultad de Ciencias
Departamento de Estadística



DIAGNÓSTICO DE INFLUENCIA EN
SISTEMAS ECONOMETRICOS
COINTEGRADOS

Marcelo Torres Trujillo

Tesis para optar al Grado Académico de

Magíster en Estadística

Profesores Guías

Dr. Manuel Galea Rojas

Dr. Felipe Osorio Salgado

Valparaíso, Junio de 2011.

Tesis financiada por CONICYT y el Gobierno Regional de Valparaíso

AGRADECIMIENTOS

Mis más sinceros agradecimientos al Dr. Manuel Galea por su desinteresada guía, al Dr. Felipe Osorio por su dedicación y entera disposición a ayudarme, al Dr. Ronny Vallejos por sus acertados comentarios y a la memoria del Profesor Renato Allende.

A ellos y a mi familia, dedico con estima y cariño el presente trabajo.

Marcelo Torres.

CONTENIDO

ABREVIATURAS	5
RESUMEN	6
ESTRUCTURA PROYECTO	7
REVISIÓN BIBLIOGRÁFICA	10
1 NOCIONES DE PROCESOS ESTOCÁSTICOS Y REGRESIÓN ESPURIA	11
1.1 Introducción	11
1.2 Regresión espuria	19
2 SERIES DE TIEMPO MULTIVARIADAS Y COINTEGRACIÓN	22
2.1 Procesos VAR	22
2.1.1 Procesos VAR estables y estacionarios	24
2.1.2 Estimación máximo verosímil de un modelo VAR	26
2.2 Cointegración	27
2.2.1 Modelo de corrección de errores	30
2.2.2 Estimación máximo verosímil en MCE	31
2.2.3 Test de Johansen para el rango de cointegración	34
3 DIAGNÓSTICO DE INFLUENCIA	36
3.1 Introducción	36
3.2 Influencia local en sistemas cointegrados	37

4	APLICACIÓN	42
4.1	Estudio de caso	42
4.2	Conclusiones estudio de caso	51
5	CONCLUSIONES	53
	BIBLIOGRAFÍA	55
	APENDICE A	58
	APENDICE B	64

ABREVIATURAS

A continuación se presentan las abreviaturas utilizadas a lo largo de esta tesis.

$AR(p)$	autorregresivo de orden p .
Cov	covarianza.
\mathbb{E}	esperanza.
EML	estimador máximo verosímil.
EMPOB	exportación al mundo de paltas en dólares FOB.
ERV	estadístico de razón de verosimilitudes.
FOB	free on board.
$I(d)$	integrado de orden d .
L	función de verosimilitud.
ℓ	función de log-verosimilitud.
MAPE	mean absolute percentage error.
MCE	modelo de corrección de errores.
ML	máximo verosímil (maximum likelihood).
N_k	Normal k -variado.
ODEPA	oficina de estudios y políticas agrarias.
OLS	mínimo cuadrado ordinario (ordinary least square).
TCR	tipo de cambio real peso/dólar.
TRV	test de razón de verosimilitudes.
$VAR(p)$	vectorial autorregresivo de orden p .
Var	varianza.
VBPC	valor en dólares por barril de petróleo crudo.

RESUMEN

El presente trabajo desarrolla la técnica de diagnóstico de influencia local en modelos de corrección de errores (MCE) de sistemas econométricos cointegrados con el fin evaluar la sensibilidad de observaciones potencialmente influyentes sobre las estimaciones de verosimilitud máxima o inferencias hechas sobre el modelo, perturbando el parámetro de escala el cual conectamos con la técnica de diagnóstico de eliminación de casos y que nos permite evaluar analíticamente el supuesto de homocedasticidad. Los resultados son ilustrados mediante el uso de una serie temporal multivariada de datos aplicada al área de la fruticultura primaria y cuyas componentes contienen información del nivel de exportación al mundo en dólares FOB de paltas en estado fresco, el tipo de cambio real peso/dólar y el valor por barril del petróleo crudo. La principal conclusión, es que el MCE propuesto es robusto a cambios de escala y que la metodología desarrollada nos permite identificar de manera eficiente los puntos influyentes en la estimación máximo verosímil del MCE.

Palabras clave: diagnóstico de influencia, modelo de corrección de errores, sistema cointegrado.

ESTRUCTURA Y OBJETIVOS DEL PROYECTO

Hoy en día, el análisis de modelos VAR (vectorial autorregresivo) ha asumido un rol importante en econometría como una herramienta estándar para el análisis de series de tiempo, debido principalmente a que las pruebas estadísticas son frecuentemente usadas para determinar la interdependencia y las posibles relaciones dinámicas entre variables o componentes de una serie temporal multivariada.

Si pensamos en una serie de tiempo univariada no estacionaria, generalmente ésta la podemos hacer estacionaria aplicando una diferenciación de forma repetida si es necesario. Entonces haciendo d (con d entero positivo) veces una diferenciación hasta que el proceso resultante sea estacionario, diremos que el proceso subyacente es integrado de orden d .

En la teoría económica generalmente podemos observar series temporales con un comportamiento no estacionario donde la media y la varianza dependen del tiempo. Ahora, si pensamos en una serie temporal K -variada donde cada una de las K -componentes es integrada de orden d , diremos a grandes rasgos (debido a que se explica con mayor detalle en los próximos capítulos) que éstas son cointegradas si existe una combinación lineal de éstas componentes de manera que la serie temporal K -variada sea estacionaria. Esta idea fue introducida por Granger (1981) y luego desarrollada por Engle y Granger, (1987) para un proceso bivariado. Sin embargo, una generalización del trabajo de Engle y Granger fue propuesto por Johansen, (1991). En el documento, Johansen presenta el método de máxima verosimilitud (denotado como ML) para el análisis de cointegración introduciendo relaciones de equi-

librio en el largo plazo en los modelos VAR con errores gaussianos, también discute el test de razón de verosimilitudes (denotado como TRV) para encontrar el rango de cointegración y la distribución asintótica del estadístico de prueba. Basándonos en este trabajo, proponemos estudiar la sensibilidad del modelo de corrección de errores (MCE) mediante la inclusión de pequeñas perturbaciones sobre la estructura de covarianza, con el fin de determinar observaciones influyentes que tienen un significativo impacto en las inferencias o predicciones del modelo. A esto se le denomina Diagnóstico de Influencia. Este tipo de modelamiento estadístico es útil en análisis de información que implique más de una serie temporal que creemos están relacionadas entre sí (como por ejemplo, la serie del precio del cobre y tipo de cambio peso-dólar). Sin embargo, consideramos que este análisis puede ser aplicable en áreas de prioridad para la región de Valparaíso como lo es la fruticultura y horticultura primaria. A la fecha, no existen aplicaciones actualizadas dedicadas a estudiar el efecto de variables macroeconómicas en las actividades relacionadas con la fruticultura y horticultura en la economía de cualquier región o a nivel nacional mediante técnicas de análisis de cointegración adecuadas y por ello, el presente proyecto busca ser un aporte a la discusión referente al tema como también entregar información relevante sobre el comportamiento de las variables en estudio que incluya un modelo econométrico que permita medir y estudiar las posibles relaciones existentes entre las componentes de la serie temporal multivariada en el largo plazo.

De esta manera, el objetivo general del presente proyecto está orientado a:

- Realizar análisis de influencia en MCE obtenidos previamente mediante un análisis de cointegración y aplicar esta técnica en áreas de interés regional de Valparaíso para contribuir con su desarrollo tecnológico e industrial.

Mientras que los objetivos específicos están orientados a:

1. Desarrollar un esquema de diagnóstico equivalente al de eliminación de casos en el MCE.
2. Proveer una aplicación de un modelo econométrico que permita medir el efecto de variables macroeconómicas en la comercialización de productos frutícolas y hortícolas mediante el uso de técnicas de series de tiempo multivariadas y cointegración.

Esta tesis está organizada de la siguiente forma: el *Capítulo 1* es de carácter introductorio, donde se entregan nociones de los procesos no estacionarios y regresión espuria. El *Capítulo 2* entrega alguna teoría de las series de tiempo multivariada concentrándose principalmente en los procesos vectoriales autorregresivos de orden p denotados por $\text{VAR}(p)$, como también en el análisis de cointegración y la estimación de los parámetros del MCE. El *Capítulo 3* entrega una introducción al diagnóstico de influencia y el desarrollo de la técnica de análisis de influencia local para el MCE en sistemas cointegrados. El *Capítulo 4* expone un estudio de casos aplicado a series de datos de un área de interés prioritario para la región de Valparaíso como lo es la fruticultura y horticultura primaria. Y el *Capítulo 5* entrega las conclusiones finales del trabajo de tesis.

REVISIÓN BIBLIOGRÁFICA

Granger y Newbold (1974) demostraron mediante simulaciones, que regresiones hechas sobre series temporales independientes y con tendencias tienden a mostrar relaciones estadísticamente significativas. A este fenómeno le denominaron regresión espuria. Para solucionar este problema, Granger (1981) publica el primer estudio sobre series de tiempo cointegradas. Posteriormente Engle y Granger (1987) demuestran los procedimientos estadísticos para probar la existencia de cointegración sobre una serie temporal bivariada. Luego Johansen (1988) entrega una generalización al trabajo presentado por Engle y Granger (1987) incorporando relaciones lineales estacionarias en modelos VAR (p).

Desde el punto de vista del diagnóstico de influencia, Cook (1986) desarrolló el método de influencia local en base al desplazamiento de verosimilitudes para medir la sensibilidad de pequeñas perturbaciones hechas sobre el modelo estadístico. Los primeros vínculos entre series de tiempo y diagnóstico de influencia datan de la década de los 80. Sin embargo, es Zhang et al (2001) quienes examinan la influencia local en el MCE en base a la metodología presentada por Engle y Granger (1987). Más recientemente Nielsen (2008), emplea herramientas de diagnóstico para medir la influencia local en un modelo VAR cointegrado basado en el principio de leave- k -out, el cual deja k observaciones fuera del análisis. Sin embargo, a la fecha no existen desarrollos que consideren procesos VAR cointegrados cuando queremos medir la sensibilidad de las observaciones en las estimaciones de verosimilitud máxima asumiendo el supuesto de heterocedasticidad.

NOCIONES DE PROCESOS ESTOCÁSTICOS Y REGRESIÓN ESPURIA

El objetivo del presente capítulo es introducir brevemente el concepto de proceso estocástico y principalmente el de proceso estocástico no estacionario, como también exponer de manera simple el problema de la regresión espuria mediante el ejemplo de dos series temporales simuladas por medio de una rutina en código R que tienen como principal característica una tendencia lineal estocástica.

1.1 Introducción

El punto de partida en el análisis de series temporales es la selección de un posible modelo matemático (o clase de modelos) para los datos. Al considerar la incertidumbre de las observaciones de un proceso, es natural suponer que cada observación y_t es un valor obtenido de cierta variable aleatoria Y_t . Por lo tanto, diremos que la serie de tiempo $\{y_t, t \in T_0\}$ es una realización de la familia de variables aleatorias $\{Y_t, t \in T_0\}$, en consecuencia, esta con-

sideración sugiere que ahora el modelamiento de los datos se tome como una realización (o parte de una realización) de un proceso estocástico $\{Y_t, t \in T\}$ donde $T \supseteq T_0$. Con el fin de ilustrar esta idea se define lo siguiente:

Definición 1. *Un proceso estocástico es una familia de variables aleatorias*

$$\{Y_t, t \in T\} = \{Y_t(\omega), t \in T, \omega \in \Omega\} \quad (1.1)$$

definidas sobre un mismo espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ que están indexadas ya sea en el tiempo y/o en el espacio.

Comúnmente, en el análisis de series de tiempo T es un conjunto de puntos en el tiempo. Luego T puede ser continuo si es un intervalo o puede ser discreto si T es un conjunto finito o infinito numerable. Es importante notar que un proceso estocástico es una función de dos variables. Para un instante de tiempo fijo t , es una variable aleatoria $\{Y_t = Y_t(\omega), \omega \in \Omega\}$, mientras que para un valor fijo de $\omega \in \Omega$, es una función $\{Y_t = Y_t(\omega), t \in T\}$ que depende del tiempo la cual define una *realización o trayectoria de un proceso estocástico*.

Ejemplo 1. *Sea Q y M dos variables aleatorias independientes con $Q \geq 0$ y $M \sim N(\mu, \sigma^2)$. Un proceso estocástico $\{Y_t, t \in T\}$ puede ser definido en términos de Q y M para cualquier $a \geq 0$ y $b > 0$ como;*

$$Y_t = b^{-1}Q \cos(at + M),$$

o más explícitamente

$$Y_t(\omega) = b^{-1}Q(\omega) \cos(at + M(\omega)), \quad (1.2)$$

donde ω es un elemento del espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$ y en el cual Q y M están bien definidos.

Las realizaciones definidas por el proceso (1.2) son las funciones obtenidas cuando ω es fijo; es decir, funciones del tipo

$$y_t = b^{-1}q \cos(at + \theta). \quad (1.3)$$

La Figura 1.1 muestra un posible comportamiento de un proceso estocástico cuando fijamos ω .

Cabe hacer notar que un proceso estocástico también puede ser descrito por la distribución conjunta de todas las sub-colecciones finitas

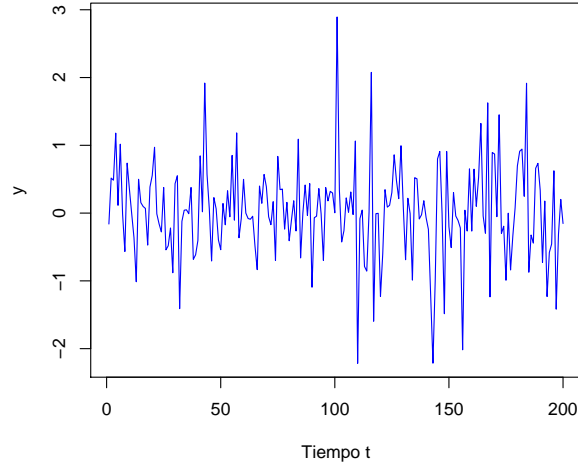


Figura 1.1: proceso estocástico para ω fijo.

de y_t . Sin embargo, en la práctica el sistema completo de distribuciones es desconocido. Es por ello es que generalmente en el análisis estadístico buscamos trabajar con procesos estacionarios, es decir procesos cuya $\mathbb{E}(y_t)$, $\text{Var}(y_t)$ y $\text{Cov}(y_t, y_s)$ son independientes del tiempo $t, \forall t, s \in \mathbb{Z}^+$.

Definición 2. *Un proceso estocástico vectorial o proceso estocástico multivariado es una función*

$$\begin{aligned} y : T \times \Omega &\rightarrow \mathbb{R}^K \\ (t, \omega) &\rightarrow \mathbf{y}(t, \omega), \end{aligned} \quad (1.4)$$

donde $\mathbf{y}(t, \omega)$ es un vector aleatorio K -dimensional.

En general una realización de un proceso estocástico vectorial (o univariado cuando $K = 1$) es una secuencia de $\mathbf{y}_t(\omega)$, $t \in Z$ para un ω fijo. De ahí, podemos decir que una serie temporal multivariada \mathbf{y} es un conjunto finito de tales realizaciones o bien, un vector de K diferentes procesos temporales medidos al mismo tiempo.

En un principio, cuando hablamos de procesos de tiempo discreto o más popularmente conocidos como procesos temporales o series de tiempo,

nuestro inconsciente casi de una manera automática asocia que el análisis de datos será realizado sobre procesos estacionarios el cual asume media y varianzas independientes del tiempo. Sin embargo, en muchas ocasiones este supuesto es inviable y un ejemplo claro en esa línea son las series de tiempo financieras y económicas las cuales intrínsecamente evolucionan, crecen y cambian a medida que avanza el tiempo, debido a que van de la mano con cambios en la tecnología de recolección de datos, cambios sociales, como también cambios en las políticas económicas y financieras de un determinado país, entre muchas otras razones. De esta manera hay veces que surgen interrogantes interesantes como las que se presentan a continuación:

1. ¿Qué tan importante es el supuesto de estacionariedad para modelar e inferir el comportamiento de los datos bajo un contexto económico o financiero?
2. ¿Cuales son los efectos de asumir erróneamente estacionariedad en los datos?
3. Dentro de la teoría estadística, ¿cuáles son los recursos existentes para afrontar el problema de no estacionariedad?
4. En el contexto económico o financiero, ¿Es válido transformar los datos para obtener una serie estacionaria?

Las respuestas a estas interrogantes según pueden ser muchas y muy variadas Hendry, (1995). Sin embargo a continuación se entrega una visión general para responder a cada una de ellas.

1. Cuando la media y la varianza de los datos no son constantes en el tiempo, una posible interpretación es que las observaciones provienen de distintas distribuciones y ello crea un problema de difícil resolución desde el punto de vista del modelamiento. Por lo tanto el supuesto de estacionariedad de los datos es un supuesto fuerte que el analista realiza al momento de modelar.
2. Cuando asumimos falsamente que la media y la varianza de los datos es constante en el tiempo, al realizar inferencias hechas en base a un modelo estacionario se incurrirá en estimaciones y predicciones de poca confiabilidad e indudablemente afectará gravemente en la toma de decisiones.

3. Como esbozamos anteriormente, la no estacionariedad puede estar explicada en gran medida por la evolución de la economía, desorden de políticas económicas y también en cambios tecnológicos.
4. Algunas formas de no estacionariedad pueden ser eliminadas por medio de alguna transformación, sin embargo, esto soluciona un problema estadístico pero crea un problema del punto de vista económico ya que se pierde información que puede ser relevante en el estudio al transformar los datos. En economía y finanzas muchas veces no es válido realizar transformaciones sobre los datos originales.

Es así que en las series de tiempo financieras y económicas es habitual presenciar comportamientos con tendencias como los que muestra la Figura 1.2. Tales tendencias necesitan ser incorporadas en el análisis estadístico y

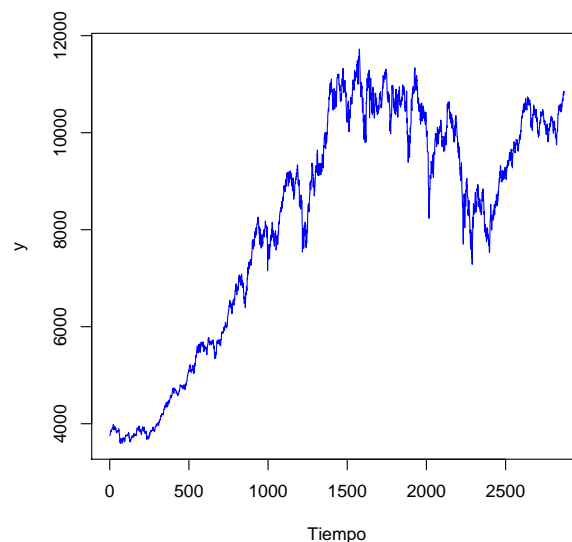


Figura 1.2: serie temporal con tendencia.

para ello existen varias técnicas, una de las más populares es la enfocada sobre un tipo de no estacionariedad estocástica inducida por la acumulación de efectos de errores anteriores llamada procesos de raíz unitaria. Tales procesos pueden ser interpretados de manera que permitan diferenciar las

tendencias en todos los puntos en el tiempo, así se dice de ellos que tienen tendencia estocástica.

Para introducir los conceptos básicos de econometría primero debemos tener presente que el punto de partida de un modelo econométrico es un modelo económico y éste se transformará en modelo econométrico cuando se añadan las especificaciones necesarias para su implementación empírica, de esta forma entenderemos por modelo econométrico a todo modelo que sigue los conceptos de la estadística a la teoría económica para su verificación y la solución de problemas económicos.

Consideremos un modelo de regresión para una variable y_t la cual contiene una tendencia lineal determinística con una pendiente β generada desde un valor inicial y_0 tal que

$$y_t = y_0 + \beta t + u_t, \quad t = 1, 2, \dots, T. \quad (1.5)$$

donde u_t es un proceso autorregresivo de primer orden AR(1) tal que:

$$u_t = \rho u_{t-1} + \epsilon_t, \quad (1.6)$$

donde el shock estocástico $\epsilon_t \sim N(0, \sigma_\epsilon^2)$. Un proceso como $\{\epsilon_t\}$ es llamado *white noise* o *ruido blanco* y de sus tres propiedades, llámese independencia, idénticamente distribuidos y normalidad, las dos primeras son las más deseables.

En adelante utilizaremos L como operador de rezagos. De esto se desprende que $Lx_t = x_{t-1}$. Entonces, si operamos sobre la ecuación (1.6) obtenemos:

$$\begin{aligned} u_t &= \rho L u_t + \epsilon_t \\ 1 &= \frac{u_t - \epsilon_t}{\rho L u_t} \\ \epsilon_t &= u_t - \rho L u_t \\ u_t &= \frac{\epsilon_t}{1 - \rho L}. \end{aligned} \quad (1.7)$$

Si $|\rho| < 1$, el término $(1 - \rho L)$ se puede expandir como

$$1 + \rho L + \rho^2 L^2 + \dots \quad (1.8)$$

Entonces,

$$u_t = \epsilon_t + \rho \epsilon_{t-1} + \rho^2 \epsilon_{t-2} + \dots \quad (1.9)$$

se puede interpretar como la suma de todas las innovaciones aleatorias en los instantes anteriores, los cuales tienen menor importancia a medida que están más alejadas del instante t . Sin embargo, si $|\rho| = 1$ es fácil notar que las innovaciones aleatorias persistirán de manera indefinida en el tiempo lo que repercute directamente en un efecto constante sobre u_t , y que comúnmente en la literatura se puede encontrar con el nombre de tendencia estocástica de tal manera que $\sum_{i=1}^t \epsilon_t$ es la suma de los errores o innovaciones. En procesos de tiempo continuo esta suma es reemplazada por una integral. De allí el nombre de procesos integrados que denotamos como $u_t \sim I(1)$ cuando $|\rho| = 1$ llamado comúnmente como proceso de raíz unitaria y $u_t \sim I(0)$ cuando $|\rho| < 1$. En general, los procesos univariados con d raíces unitarias (d raíces iguales a 1) en su operador autorregresivo son llamados integrados de orden d ($I(d)$). Para diferenciar entre tendencia estocástica y tendencia determinística decimos que en la primera, los incrementos del proceso son aleatorios, mientras que en la segunda, los incrementos del proceso son constantes en el tiempo. En la Figura 1.3 se observa el comportamiento característico de una serie de tiempo simulada con tendencia determinística, una serie con tendencia estocástica y otra cuyo comportamiento sigue un camino aleatorio.

Si tomamos

$$y_t = \beta t + \frac{\epsilon_t}{1 - \rho L} + y_0 \quad (1.10)$$

$$(1 - \rho L) y_t = (1 - \rho L) \beta t + (1 - \rho L) y_0 + \epsilon_t, \quad (1.11)$$

podemos observar que si $|\rho| = 1$, entonces $y_t - y_{t-1} = \beta t - \beta(t-1) + y_0 - y_{-1} + \epsilon_t$ ello hace que la esperanza de este proceso, $\mathbb{E}[y_t - y_{t-1}]$, dependa del tiempo t y por lo tanto no será estacionaria.

Si en vez de lo anterior, escribimos (1.11) usando $Lx_t = x_{t-1}$ se tiene que:

$$y_t = \rho y_{t-1} + \beta(1 - \rho)t + \rho\beta + (1 - \rho)y_0 + \epsilon_t, \quad (1.12)$$

podemos decir entonces que el modelo (1.5) es un modelo de regresión con residuos autocorrelacionados, el cual es equivalente a un modelo dinámico con residuos ruidos blancos.

En este tipo de modelos podemos considerar cuatro tipos de casos diferentes. Para facilitar la comprensión de estos casos escribamos el modelo (1.12) de una manera levemente diferente

$$y_t = b_1 y_{t-1} + b_2 t + b_0 + \epsilon_t, \quad (1.13)$$

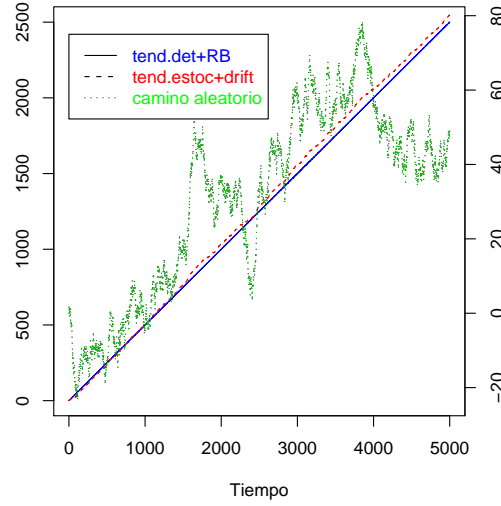


Figura 1.3: series con tendencia determinística, estocástica y camino aleatorio.

donde

$$\begin{aligned} b_1 &= \rho, \\ b_t &= \beta(1 - \rho), \\ b_0 &= \rho\beta + (1 - \rho)y_0. \end{aligned}$$

Así tenemos los siguientes casos:

Caso 1: $\rho = 1$ y $\beta \neq 0$. Desde (1.12) deducimos que tomando una diferenciación del proceso, $\Delta y_t = y_t - y_{t-1}$, se tiene que $\Delta y_t = \beta + \epsilon_t$. De allí que $\mathbb{E}[\Delta y_t] = \beta \neq 0$ es equivalente a decir que y_t tiene una tendencia lineal.

Caso 2: $\rho = 1$ y $\beta = 0$. Del caso 1 desprendemos de manera inmediata que $\Delta y_t = \epsilon_t$ y por lo tanto $\mathbb{E}[\Delta y_t] = 0$, esto es equivalente a decir que y_t no tiene tendencia.

Caso 3: $|\rho| < 1$ y $\beta \neq 0$ dado el modelo con tendencia lineal en (1.13), la interpretación de los coeficientes b_0 , b_1 y b_2 se debe hacer en función que la tendencia del proceso es $\beta = b_2 / (1 - \rho)$ y no en b_2 como se pudiera pensar en un principio.

Caso 4: $|\rho| < 1$ y $\beta = 0$ tenemos $y_t = \rho y_{t-1} + (1 - \rho)y_0 + \epsilon_t$ el cual es un

modelo AR(2) con un término constante.

- En el modelo de regresión (1.5), el término constante es esencialmente un contador del valor inicial de la serie y .
- En el modelo dinámico (1.13), el término constante es una ponderación de la razón de crecimiento de β y el valor inicial y_0 .
- En el modelo diferenciado $\rho = 1$, el término constante es solamente una medida de la razón de crecimiento β .

1.2 Regresión espuria

El problema de la regresión espuria o también llamada regresión sin sentido comienza a ser estudiado por vez primera en 1926 gracias a Undy Yule, quien analiza los riesgos en que se incurre cuando realizamos una regresión entre variables con tendencias, proponiendo formalmente un análisis para correlaciones espurias que son correlaciones extremadamente altas entre variables que no tienen explicación causal, como por ejemplo la razón de partos humanos con el número de cigüeñas avistadas en una determinada ciudad cuyo comportamiento se caracteriza por un R^2 alto.

Este efecto es causado porque la variable endógena (o variable respuesta) contiene una tendencia estocástica y la variabilidad total de ella está dada por $\sum_{t=1}^T (y_t - \bar{y})^2$, es decir, se asume erróneamente que la serie tiene una media fija. Además debemos tener en cuenta que

$$R^2 = 1 - \frac{\sum_{t=1}^T \hat{\epsilon}_t^2}{\sum_{t=1}^T (y_t - \bar{y})^2}, \quad (1.14)$$

mide la calidad del ajuste, y en estos casos, éste tenderá a valores cercanos a 1 debido a los altos valores del denominador.

Granger y Newbold (1974) destacaron que un buen ajuste entre series con tendencias cuyas correlaciones resultaban ser significativas en los residuos, era un resultado asociado a la regresión espuria. Estos autores, además verifican a través de simulaciones que dadas dos series completamente no correlacionadas y con un orden de integración igual a 1, es decir $y_i \sim I(1)$ con $i = 1, 2$, una regresión entre ellas tenderá a producir relaciones aparentemente significativas. Como una regla, ellos sugieren que se debiera dudar de la regresión hecha si se asume estacionariedad cuando el valor de R^2 es

mayor que el estadístico Durbin-Watson.

A continuación se muestra un ejemplo en código R para la regresión espuria simulando dos caminos aleatorios no relacionados y realizando una regresión sobre ellos.

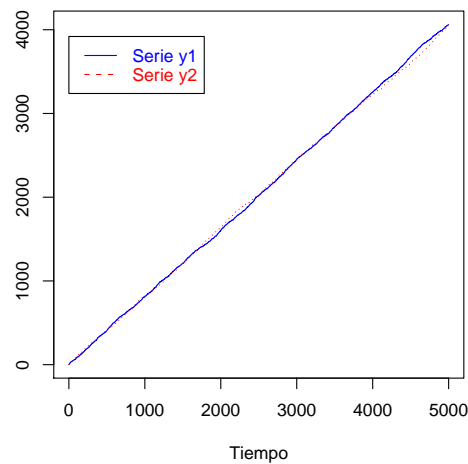
```
library(lmtest)
set.seed(123456)
error1 <- rnorm(5000)
error2 <- rnorm(5000)
tendet <- 1:5000
y1 <- 0.8 * tendet + cumsum(error1)
y2 <- 0.6 * tendet + cumsum(error2)
regre <- lm(y1~y2)
estdw <- dwtest(regre)$statistic
```

La salida de este código en R es la que se muestra en la Tabla 1.1, mientras que la Figura 1.4 muestra el comportamiento de las series creadas para la regresión espuria. El valor de $R^2 = 0,9995$; muy cercano a uno, mientras que el valor del estadístico de Durbin-Watson es $DW = 0,004$ por lo que se deduce que la regresión hecha representa una regresión espuria tal como se esperaba.

Tabla 1.1: tabla resumen de la regresión espuria.

Coefficientes	estimación	error estándar	valor t	$\mathbb{P}(> t)$
Intercepto	-0,216	0,829	-26,10	0
y_2	1,323	0,0005	2853,47	0

Con todo esto, es válido preguntarse: *¿tiene sentido hacer una regresión de variables $I(1)$?* La respuesta es sí, si los residuos de la regresión establecida son $I(0)$, por lo tanto la pregunta ahora es *¿puede ser esto posible?* La respuesta es si, cuando las variables de la regresión establecida co-mueven en el largo plazo es decir, están cointegradas. De ahí surge la necesidad de desarrollar técnicas o herramientas útiles para analizar relaciones de equilibrio entre series de tiempo no estacionarias en el largo plazo.

Figura 1.4: series y_1 y y_2 de la regresión espuria

Aunque se muestra más detalladamente en el próximo capítulo, es importante acotar que un *análisis de cointegración* busca estudiar la relación de equilibrio en el largo plazo existente entre componentes de una serie temporal multivariada. Para ello se realizan combinaciones lineales de componentes que tengan un comportamiento no estacionario, es decir $y_{it} \sim I(d)$ con $d \geq 1$ y $j = 1, 2, \dots$, de manera tal que la combinación lineal resultante tenga un orden de integración menor que el de las componentes. Este trabajo fue presentado por primera vez en 1987 por Engle y Granger (1987) y luego por Johansen (1988) quien presenta una generalización del trabajo de Engle y Granger introduciendo combinaciones lineales de componentes no estacionarias en un modelo $\text{VAR}(p)$. Los modelos resultantes de introducir esas combinaciones lineales son comúnmente denominados modelos de corrección de errores y es sobre este tipo de modelos a los cuales se les introducirá el esquema de perturbación que se muestra en el Capítulo 3.

SERIES DE TIEMPO MULTIVARIADAS Y COINTEGRACIÓN

En el presente capítulo discutimos los modelos vectoriales autorregresivos de orden p y los modelos de corrección de errores vectoriales que están estrechamente vinculados con cointegración, algunas de sus propiedades estadísticas y la estimación de sus parámetros mediante el método de Máxima verosimilitud.

2.1 Procesos VAR

Los modelos vectoriales autorregresivos de orden p ($\text{VAR}(p)$) son modelos vectoriales de variables provenientes de procesos autorregresivos, donde cada variable depende linealmente de sus propios rezagos y de aquellos rezagos de las otras variables del vector (Rachev et al., (2007)). En otras palabras, los modelos VAR pueden ser interpretados como modelos que predicen valores futuros del proceso como una suma ponderada de valores pasados y presentes mas un término de ruido aleatorio. Es importante tener

presente que un modelo VAR es simplemente una extensión multivariada de los procesos AR univariados.

Definición 3. Diremos que un proceso vectorial es autorregresivo de orden p , denotado por $\text{VAR}(p)$, si lo podemos escribir como

$$\mathbf{y}_t = \boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-2} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \boldsymbol{\epsilon}_t, \quad (2.1)$$

donde

\mathbf{y}_t es un vector aleatorio ($K \times 1$).

$\boldsymbol{\nu}$ es un vector de intercepto fijo ($K \times 1$).

\mathbf{A}_i es una matriz de coeficientes autorregresivos ($K \times K$).

$\boldsymbol{\epsilon}_t$ es un vector de innovaciones o ruidos blancos K -dimensional.

con:

$$\mathbb{E}(\boldsymbol{\epsilon}_t) = \mathbf{0}, \mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma} \text{ y } \mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_s') = \mathbf{0} \quad \forall s \neq t.$$

Cualquier proceso VAR con $p > 1$ puede ser escrito como un proceso VAR(1). Más precisamente, si \mathbf{y}_t es un proceso VAR(p), una representación correspondiente a un proceso VAR(1) Kp -dimensional está dada por

$$\mathbf{Y}_t = \mathbf{V} + \mathbf{QY}_{t-1} + \mathbf{U}_t$$

donde

$$\mathbf{Y}_t = \begin{bmatrix} \mathbf{y}_t \\ \mathbf{y}_{t-1} \\ \vdots \\ \mathbf{y}_{t-p+1} \end{bmatrix}_{Kp \times 1}, \quad \mathbf{V} = \begin{bmatrix} \boldsymbol{\nu} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}_{Kp \times 1},$$

$$\mathbf{Q} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2 & \dots & \mathbf{A}_{p-1} & \mathbf{A}_p \\ \mathbf{I}_K & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_K & \dots & \mathbf{0} & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{I}_K & \mathbf{0} \end{bmatrix}_{Kp \times Kp}, \quad \mathbf{U}_t = \begin{bmatrix} \boldsymbol{\epsilon}_t \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{bmatrix}_{Kp \times 1}.$$

2.1.1 Procesos VAR estables y estacionarios

En esta sección discutiremos una propiedad deseable en los procesos VAR denominada estacionariedad, la cual está ligada estrechamente con los procesos denominados estables.

Siguiendo el desarrollo de Lütkepohl, (2006), consideremos que desde un modelo VAR(1)

$$\mathbf{y}_t = \boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t, \quad (2.2)$$

podemos generar un conjunto de series temporales que comiencen en el tiempo $t = 1$, de esta manera tenemos

$$\begin{aligned} \mathbf{y}_1 &= \boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_0 + \boldsymbol{\epsilon}_1, \\ \mathbf{y}_2 &= \boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_1 + \boldsymbol{\epsilon}_2, \\ &= \boldsymbol{\nu} + \mathbf{A}_1 (\boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_0 + \boldsymbol{\epsilon}_1) + \boldsymbol{\epsilon}_2, \\ &= (\mathbf{I}_K + \mathbf{A}_1) \boldsymbol{\nu} + \mathbf{A}_1^2 \mathbf{y}_0 + \mathbf{A}_1 \boldsymbol{\epsilon}_1 + \boldsymbol{\epsilon}_2, \\ &\vdots \\ \mathbf{y}_t &= (\mathbf{I}_K + \mathbf{A}_1 + \dots + \mathbf{A}_1^{t-1}) \boldsymbol{\nu} + \mathbf{A}_1^t \mathbf{y}_0 + \sum_{i=0}^{t-1} \mathbf{A}_1^i \boldsymbol{\epsilon}_{t-i}, \\ &\vdots \end{aligned}$$

De este modo, los vectores $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t$ están únicamente determinados por $\mathbf{y}_0, \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_t$. También la distribución conjunta de $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_t$ es determinada por la distribución conjunta de $\mathbf{y}_0, \boldsymbol{\epsilon}_1, \dots, \boldsymbol{\epsilon}_t$.

$$\begin{aligned} \mathbf{y}_t &= \boldsymbol{\nu} + \mathbf{A}_1 \mathbf{y}_{t-1} + \boldsymbol{\epsilon}_t \\ &= (\mathbf{I}_K + \mathbf{A}_1 + \dots + \mathbf{A}_1^j) \boldsymbol{\nu} + \mathbf{A}_1^{j+1} \mathbf{y}_{t-j-1} + \sum_{i=0}^{t-1} \mathbf{A}_1^i \boldsymbol{\epsilon}_{t-i}. \end{aligned}$$

Si todos los valores propios de \mathbf{A}_1 tienen módulo menor que 1, se puede demostrar que la secuencia \mathbf{A}_1^j con $j = 0, 1, \dots$ es absolutamente sumable. Por lo tanto, la suma infinita $\sum_{i=1}^{\infty} \mathbf{A}_1^i \boldsymbol{\epsilon}_{t-i}$ converge en media cuadrática tal que

$$(\mathbf{I}_K + \mathbf{A}_1 + \dots + \mathbf{A}_1^j) \boldsymbol{\nu} \xrightarrow{j \uparrow \infty} (\mathbf{I}_n - \mathbf{A}_1)^{-1} \boldsymbol{\nu}. \quad (2.3)$$

Así, si todos los valores propios de \mathbf{A}_1 tienen módulo menor que 1, el proceso VAR(1) \mathbf{y}_t , es un proceso estocástico bien definido.

$$\mathbf{y}_t = \boldsymbol{\mu} + \sum_{i=0}^{\infty} \mathbf{A}_1^i \boldsymbol{\epsilon}_{t-i}, \quad t = 0, \pm 1, \pm 2, \dots,$$

de esta manera, es fácil ver que la media de este proceso está dada por:

$$\begin{aligned} \mathbb{E}(\mathbf{y}_t) &= \mathbf{A}_1 \mathbb{E}(\mathbf{y}_t) + \boldsymbol{\nu} \\ \boldsymbol{\mu} &= \mathbf{A}_1 \boldsymbol{\mu} + \boldsymbol{\nu} \\ \boldsymbol{\mu} &= (\mathbf{I}_K - \mathbf{A}_1)^{-1} \boldsymbol{\nu}. \end{aligned} \quad (2.4)$$

El primer y segundo momento del proceso \mathbf{y}_t están dados por

$$\mathbb{E}(\mathbf{y}_t) = \boldsymbol{\mu}, \quad \forall t,$$

y

$$\boldsymbol{\Gamma}(h) = \mathbb{E}(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-h} - \boldsymbol{\mu})',$$

con $\mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_s') = 0 \quad \forall s \neq t$ y $\mathbb{E}(\boldsymbol{\epsilon}_t \boldsymbol{\epsilon}_t') = \boldsymbol{\Sigma}, \quad \forall t$. Se puede demostrar que:

$$\boldsymbol{\Gamma}(h) = \sum_{i=0}^{\infty} \mathbf{A}_1^{h+i} \boldsymbol{\Sigma} \mathbf{A}_1^{i'}. \quad (2.5)$$

Ahora estamos en condiciones de dar la siguiente definición

Definición 4. *Un proceso VAR(1) es estable si todos los valores propios de \mathbf{A}_1 tienen módulo menor que 1. Esto es*

$$\det(\mathbf{I}_K - \mathbf{A}_1 z) \neq 0, \quad \text{para } |z| \leq 1 \quad (2.6)$$

Vale la pena destacar que un proceso \mathbf{y}_t para $t = 0, \pm 1, \pm 2, \dots$ puede también ser definido si la condición de estabilidad no se satisface.

Dada la equivalencia entre un modelo VAR(p) y un modelo VAR(1) se puede también considerar procesos VAR estacionarios de una manera general. Para obtener mayores detalles de esta equivalencia revise Lütkepohl, (2006) (Parte I, sección 2).

De lo anterior, extraemos que el proceso \mathbf{y}_t es *estable* si

$$\det(\mathbf{I}_K - \mathbf{A}_1 z - \dots - \mathbf{A}_p z^p) \neq 0, \quad \text{para } |z| \leq 1, \quad (2.7)$$

esta condición es llamada *condición de estabilidad*.

A continuación se presentan dos definiciones de un proceso estocástico estacionario que son equivalentes.

Definición 5. *Un proceso estocástico es estacionario si su primer y segundo momento son invariantes en el tiempo.*

En otras palabras un proceso estocástico \mathbf{y}_t es estacionario si se cumple que $\mathbb{E}(\mathbf{y}_t) = \boldsymbol{\mu}, \quad \forall t$ y $\mathbb{E}[(\mathbf{y}_t - \boldsymbol{\mu})(\mathbf{y}_{t-h} - \boldsymbol{\mu})']$ depende sólo de $h \quad \forall t, h = 0, 1, 2, \dots$

Definición 6. Un proceso \mathbf{y}_t que puede ser modelado por un proceso VAR será estacionario si todas las soluciones de

$$\det(\mathbf{I}_K - \mathbf{A}_1 z) = 0 \quad (2.8)$$

están fuera del círculo unitario.

Como las soluciones de (2.8) son inversas a los valores propios de \mathbf{A}_1 , una condición equivalente es que el módulo de todos los valores propios de \mathbf{A}_1 sean menores que 1. Es decir, $\det(\mathbf{I}_K - \mathbf{A}_1 z) \neq 0$, $|z| \leq 1$. Las definiciones 5 y 6 descritas anteriormente dan paso a la siguiente proposición:

Proposición 1. Si un proceso \mathbf{y}_t sigue un comportamiento VAR estable, entonces \mathbf{y}_t es estacionario e invertible.

Nota: Debido a que la proposición 1 no siempre se cumple, un proceso *no estable* no necesariamente es un proceso *no estacionario*.

2.1.2 Estimación máximo verosímil de un modelo VAR

En esta sección centramos la atención en la estimación máximo verosímil (ML) (por sus siglas en inglés *maximun likelihood*) de los modelos VAR. Para ello asumimos que en una serie temporal \mathbf{y}_t , tenemos un tamaño muestral T para cada una de las K variables y además p condiciones y_{-p+1}, \dots, y_0 pre-muestreadas.

Si consideramos el modelo (2.1) y asumimos media conocida tenemos que

$$(\mathbf{y}_t - \boldsymbol{\mu}) = \mathbf{A}_1 (\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \mathbf{A}_2 (\mathbf{y}_{t-2} - \boldsymbol{\mu}) + \dots + \mathbf{A}_p (\mathbf{y}_{t-p} - \boldsymbol{\mu}) + \boldsymbol{\epsilon}_t, \quad (2.9)$$

además $\mathbf{Y}^* = (\mathbf{y}_1 - \boldsymbol{\mu}, \mathbf{y}_2 - \boldsymbol{\mu}, \dots, \mathbf{y}_T - \boldsymbol{\mu})$ que tiene orden $(K \times T)$, $\mathbf{A} = (\mathbf{A}_1, \dots, \mathbf{A}_p)$ de orden $(K \times Kp)$, $\mathbf{Y}_t^* = (\mathbf{y}_t - \boldsymbol{\mu}, \dots, \mathbf{y}_{t-p+1} - \boldsymbol{\mu})$ de orden $(Kp \times 1)$, $\mathbf{B} = (\mathbf{y}_1^*, \mathbf{y}_2^*, \dots, \mathbf{y}_T^*)$ de orden $(Kp \times T)$ y finalmente $\mathbf{U} = (\boldsymbol{\epsilon}_1, \boldsymbol{\epsilon}_2, \dots, \boldsymbol{\epsilon}_T)$ de orden $(K \times T)$ podemos escribir (2.9) de una manera reducida dada por

$$\mathbf{Y}^* = \mathbf{A}\mathbf{X} + \mathbf{U}, \quad (2.10)$$

donde $\mathbf{U} \sim N_{KT}(\mathbf{0}, \boldsymbol{\Sigma}_u)$. Cabe hacer notar que

$$\boldsymbol{\Sigma}_u = \mathbf{I}_T \otimes \boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma} & 0 & \dots & 0 \\ 0 & \boldsymbol{\Sigma} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \boldsymbol{\Sigma} \end{pmatrix}, \quad (2.11)$$

mantiene la estructura de covarianza del modelo original.

Vectorizando el modelo (2.10) obtenemos

$$\begin{aligned}\text{vec}(\mathbf{Y}^*) &= \text{vec}(\mathbf{A}\mathbf{X}) + \text{vec}(\mathbf{U}) \\ \text{vec}(\mathbf{Y}^*) &= \left(\mathbf{X}' \otimes \mathbf{I}_K\right) \text{vec}(\mathbf{A}) + \text{vec}(\mathbf{U}) \\ \mathbf{y}^* &= \mathbf{ba} + \mathbf{u},\end{aligned}\tag{2.12}$$

donde $\mathbf{y}^* = \text{vec}(\mathbf{Y}^*)$, $\mathbf{a} = \text{vec}(\mathbf{A})$, $\mathbf{b} = \left(\mathbf{X}' \otimes \mathbf{I}_K\right)$ y $\mathbf{u} = \text{vec}(\mathbf{U})$ y cuya función de densidad ahora queda determinada por

$$f_{\mathbf{y}^*}(\mathbf{y}^*) = \det\left(\frac{\partial \mathbf{u}}{\partial \mathbf{y}^*}\right) f_{\mathbf{u}}(\mathbf{u}),$$

y su función de verosimilitud por

$$L(\mathbf{a}, \boldsymbol{\Sigma}_u) = (2\pi)^{-KT/2} \det(\boldsymbol{\Sigma}_u)^{T/2} \exp\left\{\frac{1}{2}(\mathbf{y}^* - \mathbf{ba})' \boldsymbol{\Sigma}_u^{-1} (\mathbf{y}^* - \mathbf{ba})\right\}.$$

La función de log-verosimilitud del modelo es descrita por

$$\ell(\mathbf{a}, \boldsymbol{\Sigma}_u) = -\frac{KT}{2} \log(2\pi) - \frac{1}{2} \log(\det(\boldsymbol{\Sigma}_u)) - \frac{1}{2} (\mathbf{y}^* - \mathbf{ba})' \boldsymbol{\Sigma}_u^{-1} (\mathbf{y}^* - \mathbf{ba}).$$

Minimizamos esta función respecto de \mathbf{a} manteniendo $\boldsymbol{\Sigma}_u$ fijo, obtenemos que el estimador ML de \mathbf{a} , $\hat{\mathbf{a}}$ está dado por

$$\hat{\mathbf{a}} = \left(\mathbf{b}' \boldsymbol{\Sigma}_u^{-1} \mathbf{b}\right)^{-1} \left(\mathbf{b}' \boldsymbol{\Sigma}_u^{-1} \mathbf{y}^*\right).\tag{2.13}$$

Es bien sabido que en el caso de innovaciones con distribución normal las estimaciones ML y OLS son equivalentes. Para el estudio de otras técnicas de estimación de los parámetros de interés en un modelo VAR y el estudio de criterios de información revise Hamilton (1994), Lütkepohl (2006) y Moretin (2006).

2.2 Cointegración

En la presente sección nos enfocamos al estudio de los sistemas cointegrados; concepto que describe relaciones de equilibrio entre un conjunto de variables medidas a través del tiempo, y sus relaciones de cointegración que pueden ser caracterizadas por medio de un modelo de corrección de errores el cual

representa variaciones de corto plazo desde el estado de equilibrio del sistema cointegrado (Zhang et al., (2001)). Para realizar un análisis de cointegración existen dos pasos cruciales; primero debemos establecer alguna relación de equilibrio en el largo plazo entre las componentes de la serie temporal multivariada y luego, estimar un modelo de correlación dinámico. Este modelo, que revisamos en la siguiente sección, es llamado en la literatura como Modelo de Corrección de Errores (MCE) debido a que las desviaciones de equilibrio de corto plazo son corregidas revelando la causalidad en el sentido de Granger que ésta puede presentar en un sistema cointegrado (Alexander, (2001)). Esto implica que si el resultado de una variable a afecta el resultado de otra variable que llamamos b , ayudaría a predecir el resultado de una de ellas conociendo la información del pasado de la otra. A continuación entregamos la definición dada por Engle y Granger (1987) para procesos cointegrados.

Definición 7. *Sea x_t e y_t procesos $I(d)$. Entonces una combinación lineal que llamamos $z_t = y_t - ax_t$ es en general, también $I(d)$. Además es posible que z_t sea integrado de orden menor, digamos $I(d - b)$, con $b > 0$. Si $d = b = 1$, entonces x_t e y_t serán $I(1)$ y z_t será $I(0)$. En este caso se dice que x_t e y_t son cointegrados.*

Una generalización del trabajo de Engle y Granger (1987) es dada por Johansen (1988).

Definición 8. *Las componentes del vector \mathbf{y}_t se dicen ser cointegradas de orden (d, b) , denotado como $\mathbf{y}_t \sim I(d, b)$ si existe al menos un vector $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_K)'$, no nulo tal que*

$$\boldsymbol{\beta}' \mathbf{y}_t = \beta_1 y_{1t} + \dots + \beta_K y_{Kt},$$

sea estacionaria, es decir $\boldsymbol{\beta}' \mathbf{y}_t \sim I(d, b)$, $d \geq b > 0$.

En el caso que no exista cointegración entre las componentes, usualmente el mejor modelo para estudiar las tendencias en el largo plazo es un modelo VAR.

Note que el vector de cointegración $\boldsymbol{\beta}$ no es único. Si tomamos un escalar $\lambda \neq 0$, entonces $\lambda \boldsymbol{\beta}$ también es un vector de cointegración. Usualmente una de las variables es usada para normalizar $\boldsymbol{\beta}$, fijando su coeficiente en 1, es común tomar $\boldsymbol{\beta} = (1, -\beta_2, \dots, -\beta_K)'$ de este modo tenemos que $\boldsymbol{\beta}' \mathbf{y}_t = y_{1t} - \beta_2 y_{2t} - \dots - \beta_K y_{Kt}$. Por último, si \mathbf{y}_t tiene $K \geq 2$ componentes, pueden existir varios vectores de cointegración. Si existen exactamente r

vectores de cointegración linealmente independientes con $0 < r \leq K - 1$, entonces estos se pueden agrupar en una matriz \mathbf{B} de orden $K \times r$, con rango r , llamado rango de cointegración de manera que

$$\mathbf{B}'\mathbf{y}_t = \begin{bmatrix} \beta'_1\mathbf{y}_t \\ \beta'_2\mathbf{y}_t \\ \vdots \\ \beta'_r\mathbf{y}_t \end{bmatrix} = \begin{bmatrix} u_{1t} \\ \vdots \\ u_{rt} \end{bmatrix}$$

sea estacionaria, donde u_{it} con $i = 1, 2, \dots, r$ es un paseo aleatorio.

De modo general, si el vector \mathbf{y}_t es cointegrado con r vectores de cointegración, $0 < r < K$, entonces existirán $K - r$ tendencias estocásticas comunes que no serán estacionarias.

Ejemplos gráficos del comportamiento de sistemas cointegrados se muestran en la Figura 2.1. Un dato importante a considerar es que el hecho que exista

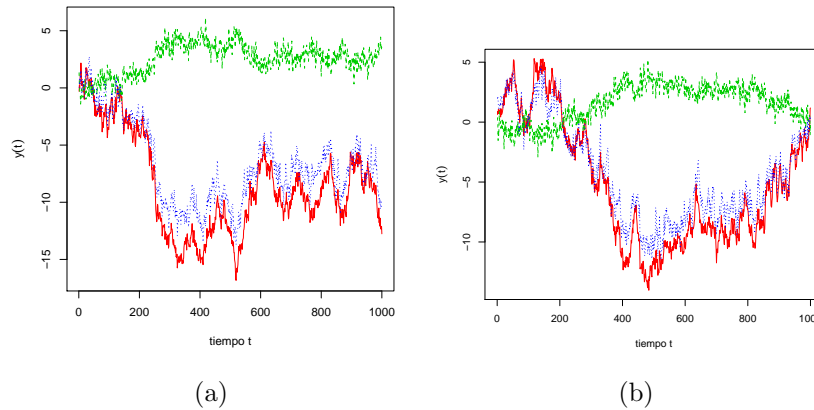


Figura 2.1: ilustración gráfica de sistemas cointegrados.

cointegración entre las componentes de una serie temporal multivariada, no significa que éstas presenten alta correlación, por ello, las gráficas de estos procesos generalmente no proporcionan información sobre la existencia o no existencia de esta característica en el largo plazo entre las componentes. Se recomienda siempre probar la existencia de cointegración mediante la aplicación de pruebas estadísticas.

2.2.1 Modelo de corrección de errores

Cuando hablamos de sistemas cointegrados, en un MCE los cambios que se producen en una variable siempre dependen de una relación de equilibrio existente con alguna otra variable, esta relación es llamada *causalidad de granger*, así un MCE es un modelo dinámico para la primera diferencia de variables $I(1)$ el cual elimina las tendencias, ya sea determinística o estocástica presentes en la serie temporal y como consecuencia soluciona el problema de la regresión espuria.

Para ilustrar los MCE, consideremos un modelo VAR con los supuestos usuales tal como (2.1). Este modelo se dice cointegrado de rango r si $\mathbf{\Pi} = \mathbf{I}_K - \mathbf{A}_1 - \dots - \mathbf{A}_p$ tiene rango r con $r < K$, de esta manera $\mathbf{\Pi}$ puede ser descompuesta como $\mathbf{\Pi} = \boldsymbol{\alpha}\boldsymbol{\beta}'$, con $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ de orden $K \times r$ de rango r . Denominamos a $\boldsymbol{\beta}$ como la *matriz de cointegración* y a $\boldsymbol{\alpha}$ como la *matriz de cargas*. Si $r = 0$, entonces $\Delta\mathbf{y}_t$ tiene una representación VAR($p-1$) estacionaria y si $r = K$, entonces \mathbf{y}_t tiene una representación VAR(p) estacionaria.

Engle y Granger (1987) demostraron que el modelo (2.1) tendrá una representación de MCE dada por

$$\Delta\mathbf{y}_t = \boldsymbol{\nu} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-1} + \mathbf{F}_1\Delta\mathbf{y}_{t-1} + \mathbf{F}_2\Delta\mathbf{y}_{t-2} + \dots + \mathbf{F}_{p-1}\Delta\mathbf{y}_{t-p+1} + \boldsymbol{\epsilon}_t, \quad (2.14)$$

donde $\mathbf{F}_i = -\sum_{j=1}^p \mathbf{A}_j$.

Ahora como $\Delta\mathbf{y}_y \sim I(0)$ y $\boldsymbol{\beta}'\mathbf{y}_{t-1} \sim I(0)$, estos términos tienen una media constante, por ello, $\mathbb{E}(\Delta\mathbf{y}_y) = \boldsymbol{\gamma}$, es un vector $K \times 1$ el cual representa tasas de crecimiento, y $\mathbb{E}(\boldsymbol{\beta}'\mathbf{y}_{t-1}) = \boldsymbol{\mu}$ un vector $r \times 1$ que representa los interceptos en las relaciones de cointegración. De esta manera, (2.14) puede ser escrito como

$$\boldsymbol{\nu} = (\mathbf{I}_K - \mathbf{F}_1 - \dots - \mathbf{F}_{p-1})\boldsymbol{\gamma} + \boldsymbol{\alpha}\boldsymbol{\mu}, \quad (2.15)$$

o también

$$\Delta\mathbf{y}_t - \boldsymbol{\gamma} = \sum_{i=1}^{p-1} \mathbf{F}_i (\Delta\mathbf{y}_{t-i} - \boldsymbol{\gamma}) - \boldsymbol{\alpha} (\boldsymbol{\beta}'\mathbf{y}_{t-1} - \boldsymbol{\mu}) + \boldsymbol{\epsilon}_t. \quad (2.16)$$

Desde (2.16) observamos que existen dos maneras de corrección de equilibrio; una relacionada con el crecimiento lineal de los datos y otra, de vectores de cointegración en relación a su valor medio. Los dos valores medios $\boldsymbol{\mu}$ y $\boldsymbol{\gamma}$ juegan un rol importante en el modelo VAR cointegrado y por ello en el

análisis empírico de las series temporales es importante comprobar si son significativamente distintos de cero.

Existen cinco casos posibles que se pueden presentar al momento de estudiar la tendencia de una serie temporal y que en este caso se expresan como restricciones lineales de las componentes determinísticas del modelo VAR y que a su vez pueden ser testeadas estadísticamente. Para el estudio de estas restricciones revise Lütkepohl (2006).

Como se esbozó anteriormente, los MCE dependen de una relación de equilibrio la cual puede ser interpretada geoméricamente. La Figura 2.2 muestra un sistema cointegrado desde un punto de vista geométrico. En ella, observamos el modelo $\Delta \mathbf{y}_t = \alpha \beta' \mathbf{y}_{t-1} + \epsilon$; el punto $\mathbf{y}_t = (y_{1t}, y_{2t})$ se mueve hacia el valor en el largo plazo dado por y_∞ que yace sobre el espacio de las combinaciones lineales estacionarias (o conjunto atractor) por la fuerza que ejerce $-\alpha$ ó α , y empujado a lo largo de este espacio por el conjunto de $K - r$ tendencias estocásticas comunes dadas en el espacio de ajuste, $sp(\alpha)$.

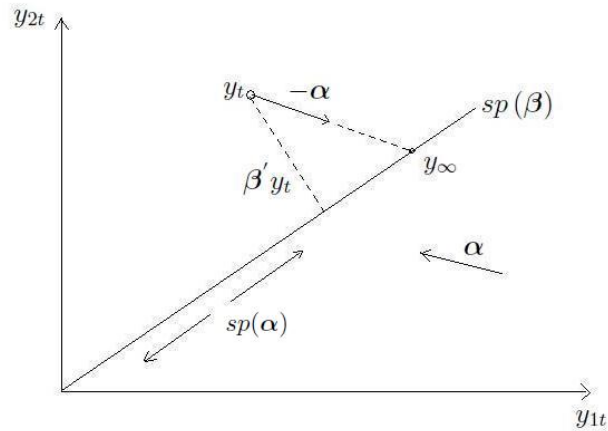


Figura 2.2: geometría de sistemas cointegrados.

2.2.2 Estimación máximo verosímil en MCE

La metodología de estimación ML para sistemas cointegrados fue desarrollada por Johansen (1988), a la que se hace referencia como metodología de Johansen, posteriormente Banerjee y Hendry (1992) proponen un método alternativo utilizando la idea de la función de verosimilitud concentrada para

llegar al mismo resultado y que es llamada comúnmente en estadística como la *verosimilitud perfilada*. Para ilustrar el procedimiento de ML presentado por Johansen, consideramos un modelo VAR(p) cointegrado tomando en cuenta la restricción $\mathbf{\Pi} = -\boldsymbol{\alpha}\boldsymbol{\beta}'$, de esta manera tenemos

$$\Delta \mathbf{y}_t = \boldsymbol{\nu} + \boldsymbol{\delta}t + \boldsymbol{\psi}\mathbf{d}_t + \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{y}_{t-p} + \sum_{i=1}^{p-1} \mathbf{F}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\epsilon}_t, \quad (2.17)$$

donde $\boldsymbol{\nu}$, $\boldsymbol{\delta}t$, $\boldsymbol{\psi}\mathbf{d}_t$ representan un término constante, una tendencia lineal y una variable dummies (este tipo de variables se incorpora cuando existe estacionalidad) respectivamente. Tomemos

$$\mathbf{z}_{0t} = \Delta \mathbf{y}_t, \quad (2.18)$$

$$\mathbf{z}_{2t} = (\mathbf{1}, t, \mathbf{d}_t, \Delta \mathbf{y}_{t-1}, \Delta \mathbf{y}_{t-2}, \dots, \Delta \mathbf{y}_{t-p+1}), \quad (2.19)$$

$$\mathbf{z}_{pt} = \mathbf{y}_{t-p}, \quad (2.20)$$

$$\boldsymbol{\Theta} = (\boldsymbol{\nu}, \boldsymbol{\delta}, \boldsymbol{\psi}, \mathbf{F}_1, \dots, \mathbf{F}_{p-1}) \quad (2.21)$$

y re-escribamos (2.17) como

$$\mathbf{z}_{0t} = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{z}_{pt} + \boldsymbol{\Theta}\mathbf{z}_{2t} + \boldsymbol{\epsilon}_t. \quad (2.22)$$

La metodología de Johansen elimina los términos de \mathbf{F} haciendo una regresión de $\Delta \mathbf{y}_t$ y \mathbf{y}_{t-p} sobre los parámetros contenidos en $\boldsymbol{\Theta}$. Separando los efectos en el modelo dados por $\boldsymbol{\Theta}\mathbf{z}_{2t}$, podemos obtener un modelo más simple por medio de las ecuaciones

$$\mathbf{z}_{0t} = \hat{\mathbf{D}}_1' \mathbf{z}_{2t} + \mathbf{R}_{0t} \quad (2.23)$$

$$\mathbf{z}_{pt} = \hat{\mathbf{D}}_2' \mathbf{z}_{pt} + \mathbf{R}_{1t}, \quad (2.24)$$

donde la estimación por el método OLS entrega valores $\hat{\mathbf{D}}_1' = \mathbf{M}_{02}\mathbf{M}_{22}^{-1}$ y $\hat{\mathbf{D}}_2' = \mathbf{M}_{p2}\mathbf{M}_{22}^{-1}$ con matriz de producto-momento dada por $\mathbf{M}_{ij} = T^{-1} \sum_{t=1}^T \mathbf{z}_{it}\mathbf{z}_{jt}'$ con $i, j = 0, 2, p$. El modelo reducido ahora es escrito como

$$\mathbf{R}_{0t} = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} + \mathbf{e}_t, \quad (2.25)$$

con $\mathbf{e}_t \sim N_K(\mathbf{0}, \boldsymbol{\Sigma})$. Obviando el término residual, si post-multiplicamos por $\mathbf{R}_{1t}'\boldsymbol{\beta}$, tenemos que

$$\mathbf{R}_{0t}\mathbf{R}_{1t}'\boldsymbol{\beta} = \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t}\mathbf{R}_{1t}'\boldsymbol{\beta}.$$

Si sumamos en t y dividimos por T , encontramos que

$$\begin{aligned} \frac{1}{T} \sum_{t=1}^T \mathbf{R}_{0t} \mathbf{R}'_{1t} \boldsymbol{\beta} &= \frac{1}{T} \sum_{t=1}^T \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \boldsymbol{\beta}, \\ \mathbf{S}_{01} \boldsymbol{\beta} &= \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{S}_{11} \boldsymbol{\beta}, \end{aligned}$$

con $\mathbf{S}_{ij} = T^{-1} \sum_{t=1}^T \mathbf{R}_i \mathbf{R}'_j = \mathbf{M}_{ij} - \mathbf{M}_{i2} \mathbf{M}_{22}^{-1} \mathbf{M}_{2j}$. Finalmente, el estimador del vector de cargas $\boldsymbol{\alpha}$ dado por $\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta})$ con $\boldsymbol{\beta}$ fijo es

$$\hat{\boldsymbol{\alpha}}(\boldsymbol{\beta}) = \mathbf{S}_{01} \boldsymbol{\beta} \left(\boldsymbol{\beta}' \mathbf{S}_{11} \boldsymbol{\beta} \right)^{-1}. \quad (2.26)$$

La función de verosimilitud del modelo reducido bajo el supuesto de normalidad está dado por

$$L^{-2/T}(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = \det(\boldsymbol{\Sigma}) \exp \left\{ \frac{1}{T} \sum_{t=1}^T \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \right\}. \quad (2.27)$$

Sin embargo, el problema de encontrar el estimador de $\boldsymbol{\beta}$ no es un problema lineal por lo que la función de verosimilitud del modelo reducido tiene su máximo en

$$L^{-2/T}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \det(\boldsymbol{\Sigma}(\boldsymbol{\alpha}, \boldsymbol{\beta})) + \text{const},$$

donde $\boldsymbol{\Sigma}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = T^{-1} \sum_{t=1}^T (\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t})(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t})'$ y cuyo estimador está dado por

$$\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \mathbf{S}_{00} - \mathbf{S}_{01} \boldsymbol{\beta} \left(\boldsymbol{\beta}' \mathbf{S}_{11} \boldsymbol{\beta} \right)^{-1} \boldsymbol{\beta}' \mathbf{S}_{10}. \quad (2.28)$$

Para ver el desarrollo del cálculo de $\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ diríjase a Juselius (2006). Minimizando el $\det(\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta}))$ obtenemos el estimador ML de $\boldsymbol{\beta}$, $\hat{\boldsymbol{\beta}}$. La función $\det(\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta}))$ es minimizada al resolver el problema de valores propios dado por la ecuación

$$\det(\lambda \mathbf{S}_{11} - \mathbf{S}_{11} + \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01}) = 0. \quad (2.29)$$

La solución de (2.29) entrega los valores propios $\hat{\lambda}_i$ asociados a los $\boldsymbol{\Lambda}_i$ vectores propios. De acuerdo a esto podemos escribir

$$\det(\hat{\boldsymbol{\Sigma}}(\boldsymbol{\alpha}, \boldsymbol{\beta})) = \det(\mathbf{S}_{00}) \prod_{i=1}^r \ln(1 - \lambda_i), \quad (2.30)$$

donde claramente se desprende que el máximo de $L^{-2/T}(\boldsymbol{\alpha}, \boldsymbol{\beta})$ está dado por

$$L_{max}^{-2/T}(\boldsymbol{\alpha}, \boldsymbol{\beta}) = \det(\mathbf{S}_{00}) \prod_{i=1}^r \ln(1 - \lambda_i) + \text{const.} \quad (2.31)$$

Los valores propios son ordenados de acuerdo a $\hat{\lambda}_1 > \dots > \hat{\lambda}_K > 0$. La magnitud de $\hat{\lambda}_i$ es una medida de la *estacionariedad* de la relación $\boldsymbol{\Lambda}'_i \mathbf{y}_t$, mientras mayor es la magnitud del valor propio asociado, más estacionaria es esta relación.

Para evitar el problema de no identificación de la matriz $\boldsymbol{\Pi}$, Johansen (1988) impuso las condiciones de normalización haciendo $\hat{\mathbf{v}}' \mathbf{S}_{11} \hat{\mathbf{v}} = \mathbf{I}$, con $\hat{\mathbf{v}} = (\hat{\boldsymbol{\Lambda}}_1, \dots, \hat{\boldsymbol{\Lambda}}_r)$. De esta manera los vectores de cointegración pueden ser estimados como

$$\hat{\boldsymbol{\beta}} = (\hat{\boldsymbol{\Lambda}}_1, \dots, \hat{\boldsymbol{\Lambda}}_r), \quad (2.32)$$

y en definitiva el estimador ML de la matriz $\boldsymbol{\Pi}$, $\hat{\boldsymbol{\Pi}}$ es dado por

$$\hat{\boldsymbol{\Pi}}(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}) = \mathbf{S}_{01} \hat{\boldsymbol{\beta}} (\hat{\boldsymbol{\beta}}' \mathbf{S}_{11} \hat{\boldsymbol{\beta}})^{-1} \hat{\boldsymbol{\beta}}'. \quad (2.33)$$

Cuando el problema se traduce en determinar el número de relaciones estacionarias, que es equivalente a encontrar el rango de la matriz $\boldsymbol{\Pi}$, Johansen, (1988) desarrolló una prueba que discrimina las $K - r$ relaciones no estacionarias y que revisamos brevemente a continuación.

2.2.3 Test de Johansen para el rango de cointegración

La metodología de Johansen para encontrar el rango r de cointegración es una generalización multivariada de la prueba de Dickey-Fuller. El problema consiste en derivar un procedimiento que discrimine entre los λ_i con $i = 1, \dots, r$ que son lo suficientemente grandes para entregar combinaciones lineales estacionarias y aquellos valores propios λ_i con $i = r + 1, \dots, K$ que son lo suficientemente pequeños y que corresponden a los vectores propios no estacionarios (Johansen, (1988)). Así el rango r se determina por medio de un test de razón de verosimilitudes (TRV) entre dos hipótesis:

$$H_0 : r = r_0 \quad v/s \quad H_1 : r_0 < r < r_1, \quad (2.34)$$

bajo los supuestos usuales del modelo (2.17), el máximo de la función de log-verosimilitud para un modelo con rango r de cointegración está dado por

$$\ell_{max}(r) = -\frac{T}{2} \left[\ln(\det(\mathbf{S}_{00})) + \sum_{i=1}^r \ln(1 - \lambda_i) \right] + \text{const.} \quad (2.35)$$

De este resultado el estadístico de razón de verosimilitudes (ERV) para probar (2.34) está dado por

$$\lambda_{LR}(r_0, r_1) = 2 [\ell_{max}(r_1) - \ell_{max}(r_0)] \quad (2.36)$$

$$= T \left[-\sum_{i=1}^{r_1} \ln(1 - \lambda_i) + \sum_{i=1}^{r_0} \ln(1 - \lambda_i) \right] \quad (2.37)$$

$$= -T \sum_{i=r_0+1}^{r_1} \ln(1 - \lambda_i), \quad (2.38)$$

entonces, si tenemos $\lambda_{LR}(r_0, K)$ para probar (2.34) el ERV es llamado *Estadístico de la Traza* dado por

$$\lambda_{LR}(r_0, K) = -T \sum_{i=r_0+1}^K \ln(1 - \lambda_i). \quad (2.39)$$

Si tenemos $\lambda_{LR}(r_0, r_0 + 1)$ para probar (2.34) el ERV es llamado estadístico del máximo valor propio dado por

$$\lambda_{LR}(r_0, r_0 + 1) = -T \ln(1 - \lambda_{r_0+1}). \quad (2.40)$$

Para el estudio de las distribuciones asintóticas y convergencias de los estadísticos aquí mencionados revise (Lütkepohl (2006), Cap. 7, Sección 7.2.)

DIAGNÓSTICO DE INFLUENCIA

El objetivo del presente capítulo es exponer el desarrollo del diagnóstico de influencia en sistemas cointegrados siguiendo el enfoque presentado por Cook (1986). En una primera instancia, se entrega una introducción del concepto de diagnóstico de influencia para luego centrar la atención en el análisis de influencia local y los resultados derivados de la aplicación en sistemas econométricos cointegrados.

3.1 Introducción

En Estadística es bien sabido que el modelamiento juega un rol preponderante al momento de describir el comportamiento de datos, realizar inferencias y también predicciones futuras, basados en el supuesto de normalidad generalmente usamos técnicas de estimación de verosimilitud máxima para estimar sus parámetros por las buenas propiedades estadísticas y bajo costo computacional que este método tiene. Sin embargo, también es posible encontrarse con problemas serios al momento de modelar, como por ejemplo, que el modelo seleccionado no represente fielmente el fenómeno práctico subyacente o que el modelo propuesto sea sensible a ciertos datos atípicos. Es común encontrarse con que unos pocos elementos grandes que

expliquen en mayor proporción la dimensión total del problema y muchos otros elementos pequeños que contribuyen muy poco al conjunto total. Esto se explica debido a cambios significativos en las conclusiones hechas por el modelo postulado cuando perturbamos algunos datos.

Con el ánimo de solucionar estos problemas algunos enfoques estadísticos han sido desarrollados durante las últimas décadas según Pan y Fang (2000). De la literatura revisada, se desprenden dos canales de investigación: el diagnóstico de influencia y los procedimientos robustos. El diagnóstico de influencia es un enfoque capaz de detectar observaciones influyentes estudiando el efecto de cada observación de la muestra en la variable respuesta del modelo (Pan y Fang, (2000)). Los Procedimientos Robustos, son un enfoque que se centra en las inferencias estadísticas basadas en pequeños cambios del modelo postulado cuando éste se perturba leve o moderadamente. En base a esto, se hace importante estudiar la información de cada observación en un modelo econométrico como el que estudiamos en este proyecto de tesis, debido principalmente a que unas pocas observaciones pueden ser responsables de la mayoría de las propiedades relevantes del MCE finalmente propuesto.

En la literatura estadística la metodología de identificación de observaciones influyentes es clasificada en dos categorías: la **Influencia Global** y la **Influencia Local**. Una de las principales tareas de la influencia global es elegir una métrica apropiada para medir la diferencia entre los resultados estadísticos obtenidas mediante el uso del conjunto completo de datos y mediante el uso de un conjunto reducido de esos datos. Mientras que la influencia local usa la curvatura geométrica de alguna métrica apropiada, como por ejemplo, el desplazamiento de verosimilitudes propuesto por Cook, (1986) para investigar el comportamiento de pequeñas perturbaciones cuando se introducen en el modelo o en los datos. A continuación desarrollamos la influencia local para MCE en sistemas cointegrados.

3.2 Influencia local en sistemas cointegrados

Durante la década de los 90, estudios de simulación demostraron que inferencias estadísticas hechas en base a sistemas cointegrados son sensibles a la validez de algunos de los supuestos del MCE, como es el caso de residuos sesgados y también, que estas inferencias son moderadamente robustas en

casos de exceso de curtosis y residuos heterocedásticos. En virtud de lo anterior, en esta subsección estudiamos la sensibilidad en la estimación máximo verosímil de los parámetros del MCE de un sistema cointegrado introduciendo un esquema de perturbación que nos permite trabajar el supuesto de heterocedasticidad de una manera analítica utilizando la técnica de Influencia Local propuesta por Cook (1986).

Como se esbozó anteriormente, la influencia local es un método de análisis de sensibilidad que permite estudiar la *influencia* de pequeñas perturbaciones introducidas en un modelo estadístico el cual tiene un vector de parámetros $\boldsymbol{\theta}$ de dimensión $(m \times 1)$ y función de log-verosimilitud $\ell(\boldsymbol{\theta})$, donde $\hat{\boldsymbol{\theta}}$ es el estimador máximo verosímil (EML) de $\boldsymbol{\theta}$.

Consideremos un vector de perturbaciones $\boldsymbol{\omega}$ de dimensión $(q \times 1)$ perteneciente a algún conjunto $\Omega \in \mathbb{R}^q$ y sea $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ la función de log-verosimilitud del modelo perturbado. De esta manera, existe un vector $\boldsymbol{\omega}_0 \in \Omega$ tal que $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}|\boldsymbol{\omega}_0)$ para todos los valores de $\boldsymbol{\omega}$ en el espacio paramétrico.

El estimador máximo verosímil de $\boldsymbol{\theta}$ del modelo perturbado lo denotamos por $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$. Si la diferencia entre $\hat{\boldsymbol{\theta}}$ y $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ es pequeña cuando $\boldsymbol{\omega}$ varía en Ω , entonces decimos que la perturbación tiene poco efecto sobre los estimadores del modelo inicial. En caso contrario, esto indica que el proceso de estimación es altamente sensible al esquema de perturbación.

Cook (1986) propone estudiar el desplazamiento de verosimilitudes dado por

$$LD(\boldsymbol{\omega}) = 2 \left[\ell(\hat{\boldsymbol{\theta}}) - \ell(\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}) \right], \quad (3.1)$$

para medir la diferencia entre $\hat{\boldsymbol{\theta}}$ y $\hat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$. En particular tenemos que $\hat{\boldsymbol{\theta}} = (\text{vec } \hat{\boldsymbol{\alpha}}, \text{vec } \hat{\boldsymbol{\beta}}, \text{vech } \hat{\boldsymbol{\Sigma}})$, el esquema de perturbación a desarrollar está dado por $\mathbf{e}_t \sim N_K(\boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t}, \boldsymbol{\Sigma}/\omega_t)$ y las funciones de log-verosimilitud para el modelo postulado y el modelo perturbado son dadas por

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{t=1}^T \ln(\det(\boldsymbol{\Sigma})) + (\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t})' \boldsymbol{\Sigma}^{-1} (\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t}),$$

$$\begin{aligned} \ell(\boldsymbol{\theta}|\boldsymbol{\omega}) &= -\frac{1}{2} \sum_{t=1}^T \left(\ln(\det(\boldsymbol{\Sigma})) - \frac{1}{T} \ln(\boldsymbol{\omega}_t) \right) + \\ &\quad \sum_{t=1}^T \boldsymbol{\omega}_t \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right), \end{aligned}$$

Para desarrollar la influencia local en el MCE denotaremos como $\ddot{\mathbf{L}}$ a la matriz de información dada por

$$\ddot{\mathbf{L}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \quad (3.2)$$

y por $\boldsymbol{\Delta}$ a la matriz

$$\boldsymbol{\Delta} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\omega}'} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \quad (3.3)$$

$C_{\mathbf{h}}(\boldsymbol{\theta})$ es la denominada influencia local sobre la estimación de $\boldsymbol{\theta}$, del esquema de perturbación en la dirección \mathbf{h} . Cook (1986) mostró que la curvatura normal en la dirección \mathbf{h} está dada por

$$C_{\mathbf{h}}(\boldsymbol{\theta}) = 2\mathbf{h}' \boldsymbol{\Delta}' \{-\ddot{\mathbf{L}}\}^{-1} \boldsymbol{\Delta} \mathbf{h} \quad , \quad \|\mathbf{h}\| = 1 \quad (3.4)$$

\mathbf{h}_{max} determina la dirección de máxima curvatura normal, por lo que C_{max} es igual dos veces el mayor valor propio de $\mathbb{B} = \boldsymbol{\Delta}' \{-\ddot{\mathbf{L}}\}^{-1} \boldsymbol{\Delta}$ asociado al vector propio \mathbf{h}_{max} . Esto nos sugiere cómo perturbar el modelo o los datos para obtener los mayores cambios locales en el desplazamiento de verosimilitudes. Una importante consideración que se debe tener en cuenta es que al trabajar la influencia local sobre procesos estocásticos únicamente nos remitiremos a identificar aquellas observaciones que ejercen una mayor influencia sobre la estimación máximo verosímil del MCE debido a que la eliminación de alguna observación en procesos de esta característica crea otro problema de mayor dificultad llamado estimación de datos faltantes.

Otra dirección importante es la dada por $\mathbf{h} = \mathbf{h}_{tT}$ el cual corresponde a la dirección asociada a la t -ésima observación, es decir \mathbf{h}_{tT} denota el vector de \mathbb{R}^q con un 1 en la t -ésima posición y todas las restantes iguales a 0. En tal caso, la máxima curvatura normal de la t -ésima observación está dada por

$$C_{\mathbf{h}_{tT}}(\boldsymbol{\theta}) = 2\boldsymbol{\Delta}'_t \{-\ddot{\mathbf{L}}\}^{-1} \boldsymbol{\Delta}_t, \quad (3.5)$$

donde $\boldsymbol{\Delta}_t$ es la t -ésima columna de $\boldsymbol{\Delta}$ con $t = 1, \dots, T$.

En nuestro caso la matriz de información escrita en forma particionada está

dada por

$$\ddot{\mathbf{L}} = \begin{pmatrix} L_{\alpha\alpha} & L_{\beta\alpha} & L_{\Sigma\alpha} \\ & L_{\beta\beta} & L_{\Sigma\beta} \\ & & L_{\Sigma\Sigma} \end{pmatrix},$$

donde

$$\begin{aligned} L_{\alpha\alpha} &= -T \left(\hat{\boldsymbol{\beta}}' \mathbf{S}_{11} \hat{\boldsymbol{\beta}} \otimes \hat{\boldsymbol{\Sigma}}^{-1} \right) \\ L_{\beta\alpha} &= -T \left(\hat{\boldsymbol{\beta}}' \mathbf{S}_{11} \otimes \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} \right) \\ L_{\Sigma\alpha} &= -T \left(\left(\hat{\boldsymbol{\beta}} \mathbf{S}_{10} - \hat{\boldsymbol{\beta}}' \mathbf{S}_{11} \hat{\boldsymbol{\beta}} \hat{\boldsymbol{\alpha}}' \right) \hat{\boldsymbol{\Sigma}}^{-1} \otimes \hat{\boldsymbol{\Sigma}}^{-1} \right) \mathbf{D}_K \\ L_{\beta\beta} &= -T \left(\hat{\boldsymbol{\alpha}}' \hat{\boldsymbol{\Sigma}}^{-1} \hat{\boldsymbol{\alpha}} \otimes \mathbf{S}_{11} \right) \\ L_{\Sigma\beta} &= -T \left(\hat{\boldsymbol{\alpha}}' \hat{\boldsymbol{\Sigma}}^{-1} \otimes \left(\mathbf{S}_{10} - \mathbf{S}_{11} \hat{\boldsymbol{\beta}} \hat{\boldsymbol{\alpha}}' \right) \hat{\boldsymbol{\Sigma}}^{-1} \right) \mathbf{D}_K \\ L_{\Sigma\Sigma} &= -\sum_{t=1}^T \mathbf{D}'_K \left(\mathbf{I}_K \otimes \hat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right)' \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right) \right) \mathbf{D}_K, \end{aligned}$$

donde \mathbf{D}_K representa una matriz de duplicación de orden $K^2 \times \frac{K(K+1)}{2}$. El desarrollo para obtener la matriz $\ddot{\mathbf{L}}$ y $\boldsymbol{\Delta}$ se muestra en el Apéndice A.

La matriz $\boldsymbol{\Delta}$ está dada por $(\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \dots, \boldsymbol{\Delta}_T)$ con

$$\boldsymbol{\Delta}_t = \begin{pmatrix} \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \text{vec } \boldsymbol{\alpha} \partial \omega_t} \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \text{vec } \boldsymbol{\beta} \partial \omega_t} \\ \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \text{vech } \boldsymbol{\Sigma} \partial \omega_t} \end{pmatrix} = \begin{pmatrix} \Delta_{\alpha t} \\ \Delta_{\beta t} \\ \Delta_{\Sigma t} \end{pmatrix} \quad t=1,2,\dots,T,$$

donde

$$\begin{aligned} \Delta_{\alpha t} &= \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right)' \hat{\boldsymbol{\Sigma}}^{-1} \otimes \mathbf{R}'_{1t} \hat{\boldsymbol{\beta}} \\ \Delta_{\beta t} &= \hat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right) \otimes \mathbf{R}_{1t} \\ \Delta_{\Sigma t} &= \mathbf{D}'_K \left(\hat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right) \otimes \hat{\boldsymbol{\Sigma}}^{-1} \left(\mathbf{R}_{0t} - \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{R}_{1t} \right) \right). \end{aligned}$$

La función de curvatura normal $C_{\mathbf{h}}(\boldsymbol{\theta})$ usada para caracterizar el desplazamiento de la verosimilitud en la dirección de máxima curvatura \mathbf{h} alrededor de $a = 0$ denotada por $LD(\boldsymbol{\omega}_0 + a\mathbf{h}_{max})$, puede ser representada en un gráfico

bidimensional dado por $LD(\boldsymbol{\omega}_0 + a\mathbf{h}_{max})$ versus $a \in \mathbb{R}$.

En el próximo capítulo, se presenta una aplicación de las técnicas aquí descritas a datos reales de acceso público ligados a la exportación al mundo de productos frutícolas y que tienen un importante rol en la economía regional de Valparaíso.

APLICACIÓN

En el presente capítulo ilustramos las técnicas descritas en los capítulos anteriores en base a la información de exportaciones de paltas y que son desarrolladas utilizando el software estadístico R.

4.1 Estudio de caso

En Chile la palta es un producto con gran demanda, internamente se consume aproximadamente el 45% de la producción mientras que el restante 55% se exporta principalmente a los E.E.U.U, Europa, Japón y Argentina. Se encuentra presente en el mercado durante todo el año y su importancia económica radica principalmente en la superficie dedicada a su cultivo ya que ocupa el cuarto lugar entre las frutas y hortalizas que se cultivan en nuestro país la cual se extiende a lo largo de todo el territorio nacional continental concentrándose mayoritariamente en las regiones de Valparaíso, Metropolitana y Coquimbo. Si a lo anterior adicionamos la alta demanda de mano de obra que su cultivo requiere y su posición en las exportaciones de productos agrícolas sin duda hacen a este producto un elemento interesante de análisis. Una problemática importante según el estudio publicado por la ODEPA Mercado de las Paltas 2009-2010 es que existe una gran volatili-

dad en los precios de exportación medidos en dólares al pasar de un año de producción a otro, esto se explica principalmente por el aumento constante del volumen de producción esperado, la agresiva competencia externa y el aumento sostenido de la producción californiana el cual a su vez, afecta las exportaciones desde Chile a Estados Unidos. Así los agricultores optimistas como consecuencia de una temporada de precios altos aumentan la superficie de cultivo generando, muchas veces, la oportunidad que se pueda incurrir en pérdidas importantes para los agricultores. Por ello, para este proyecto de tesis es importante implementar un modelo en base a relaciones de variables macroeconómicas que permitan estudiar este fenómeno y así realizar un aporte a la discusión del tema.

Durante los años 2007 a 2010 como parte de los instrumentos de planificación, el Consejo Nacional de Innovación para la Competitividad desarrolló un programa de mejoramiento de la productividad con el objeto claro de identificar ejes o áreas de prioridad investigativa y desarrollo regional en Chile. La implementación de este programa dio origen a un estudio relacionado con la competitividad en conglomerados de la economía chilena, el cuál definió sectores estratégicos de alto potencial de conglomerización dados por la fruticultura y horticultura primaria y el turismo para la región de Valparaíso. Conforme a esto, nos interesamos en estudiar procesos estocásticos no estacionarios del área de la fruticultura y horticultura primaria, el cual podemos entender como el proceso agrícola que abarca la producción de fruta y hortaliza, su selección, empaque y comercialización en el mercado doméstico como externo. De esta manera, con la intención de desarrollar una aplicación de interés regional, y teniendo en cuenta que generalmente el tipo de cambio afecta con fuerza el precio de la exportación de frutas y hortalizas como también el valor del petróleo, principalmente por el hecho que este afecta directamente el precio de los fertilizantes y transporte, en el presente proyecto realizamos el diagnóstico del análisis de cointegración para una serie multivariada cuyas componentes son la exportación al mundo en dólares FOB (free on board) de paltas en estado fresco, el tipo de cambio real peso/dólar y el valor en dólares por barril de petróleo crudo. La información contenida en las series bajo estudio es mensual y corresponde al periodo entre Enero de 1998 y Agosto de 2010 la cual fue obtenida desde los sitios web oficiales de la oficina de estudios y políticas agrarias del gobierno de Chile (<http://www.odepa.gob.cl>) y del Banco Central de Chile (<http://www.bcentral.cl>).

La Figura 4.1 muestra el comportamiento del proceso en estudio. Cabe mencionar, que el valor de la serie del precio en dólares por barril de petróleo

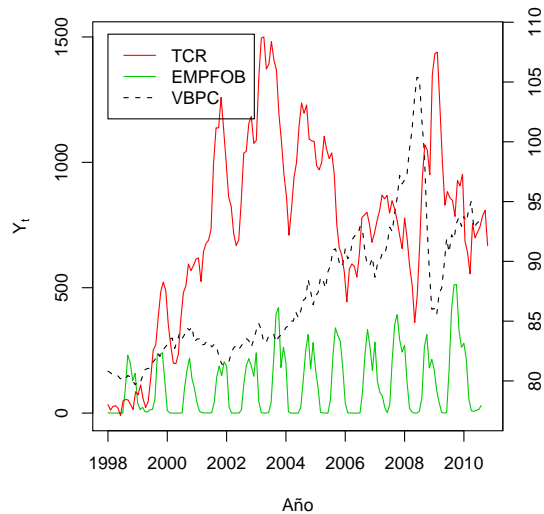


Figura 4.1: serie del EMPFOB, el TCR y el VBPC entre enero de 1998 y agosto de 2010.

crudo el valor de la serie de exportación al mundo de paltas en dólares FOB ha sido multiplicada convenientemente por 10 y por 1/100 respectivamente para obtener una mejor representación gráfica. Además los valores en el eje vertical derecho representan los valores de la serie del tipo de cambio real peso/dólar. Se puede observar que la exportación al mundo en dólares FOB de paltas en estado fresco (EMPFOB) tiene un comportamiento altamente estacional con picks muy bien marcados en los años 2003 y 2009. Además se aprecia que el tipo de cambio real peso/dólar (TCR) y el valor en dólares por barril de petróleo crudo (VBPC) tienen un comportamiento altamente volátil.

Mediante un análisis gráfico de las funciones de autocorrelación y autocorrelación parcial de las componentes del proceso en estudio dadas en la Figura 4.1, podemos decir por simple inspección que el TCR y el VBPC tienen un comportamiento no estacionario y además reafirmamos la hipótesis del comportamiento estacional con ciclos de 12 meses en el nivel de exportación al mundo en dólares FOB de paltas en estado fresco.

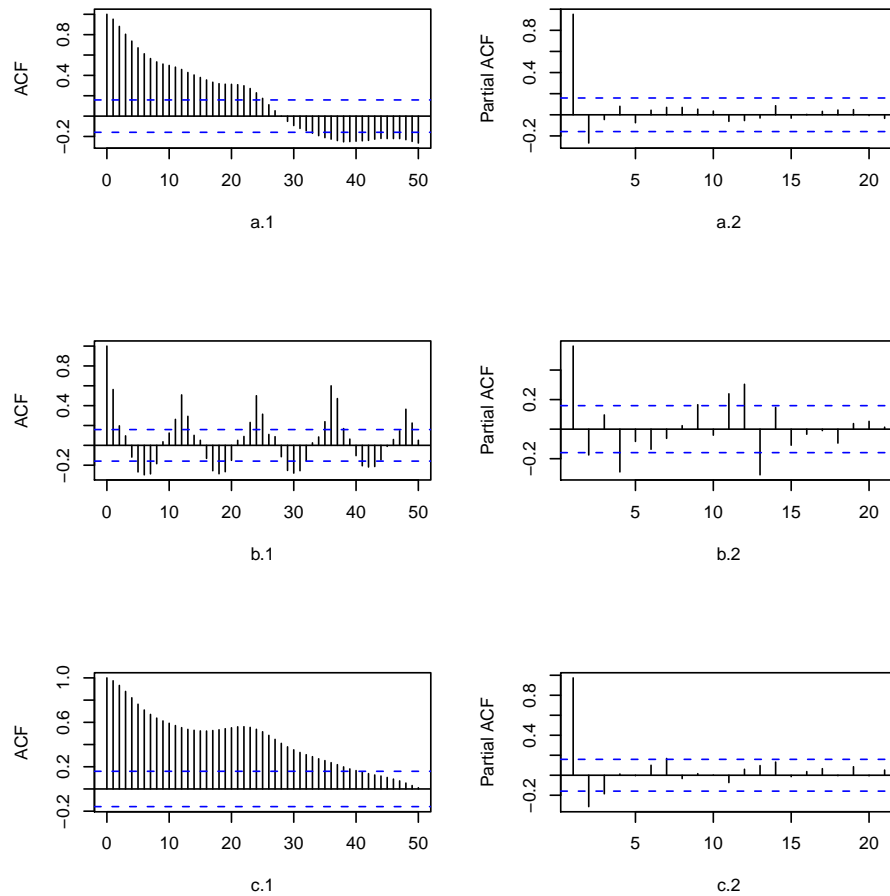


Figura 4.2: gráfica ACF y ACF parcial del TCR (a.1 y a.2), EMPFOB (b.1 y b.2) y el VBPC (c.1 y c.2) entre enero de 1998 y agosto de 2010.

El modelo VAR(2) estimado para este proceso temporal multivariado está dado por la ecuación

$$\mathbf{y}_t = \boldsymbol{\nu} + \hat{\mathbf{A}}_1 \mathbf{y}_{t-1} + \hat{\mathbf{A}}_2 \mathbf{y}_{t-2} + \boldsymbol{\epsilon}_t, \quad (4.1)$$

con matrices de coeficientes estimadas

$$\hat{\mathbf{A}}_1 = \begin{pmatrix} 1,23 & 0,00 & 0,00 \\ 0,00 & 0,62 & 0,00 \\ 0,00 & 0,00 & 1,31 \end{pmatrix}, \quad \hat{\mathbf{A}}_2 = \begin{pmatrix} -0,27 & 0,00 & 0,00 \\ 0,00 & -0,2 & 0,00 \\ 0,00 & 0,04 & -0,43 \end{pmatrix}.$$

A continuación se muestra el análisis de cointegración aplicado al proceso en estudio. Según la información extraída de las Tablas 4.1 y 4.2 no hay evidencia suficiente para rechazar la hipótesis de la existencia de una relación de cointegración estacionaria para todos los niveles de significancia dados. Esta relación de cointegración corregida por los efectos de corto plazo se observa claramente en $\hat{\beta}_1 \mathbf{R}_{1t}$ dada en la Figura 4.1.

Tabla 4.1: test de Johansen, estadístico del máximo valor propio para rango de cointegración.

	Estadístico	10%	5%	1%
$r = 2$	3,06	6,50	5,18	11,65
$r = 1$	7,89	12,91	14,90	19,19
$r = 0$	48,68	18,90	21,07	25,75

Tabla 4.2: test de Johansen, estadístico de la traza para rango de cointegración.

	Estadístico	10%	5%	1%
$r \leq 2$	3,06	6,50	8,18	11,65
$r \leq 1$	10,95	15,66	17,95	23,52
$r = 0$	59,63	28,71	31,52	37,22

Cabe mencionar que la gráfica de la relación de cointegración varía según la normalización que se haga del vector de valores propios. Las Tablas 4.3 y 4.4 muestran los valores propios estimados y los vectores propios normalizados y la matriz de cargas para los datos del nivel de exportación al mundo en dólares FOB de paltas en estado fresco.

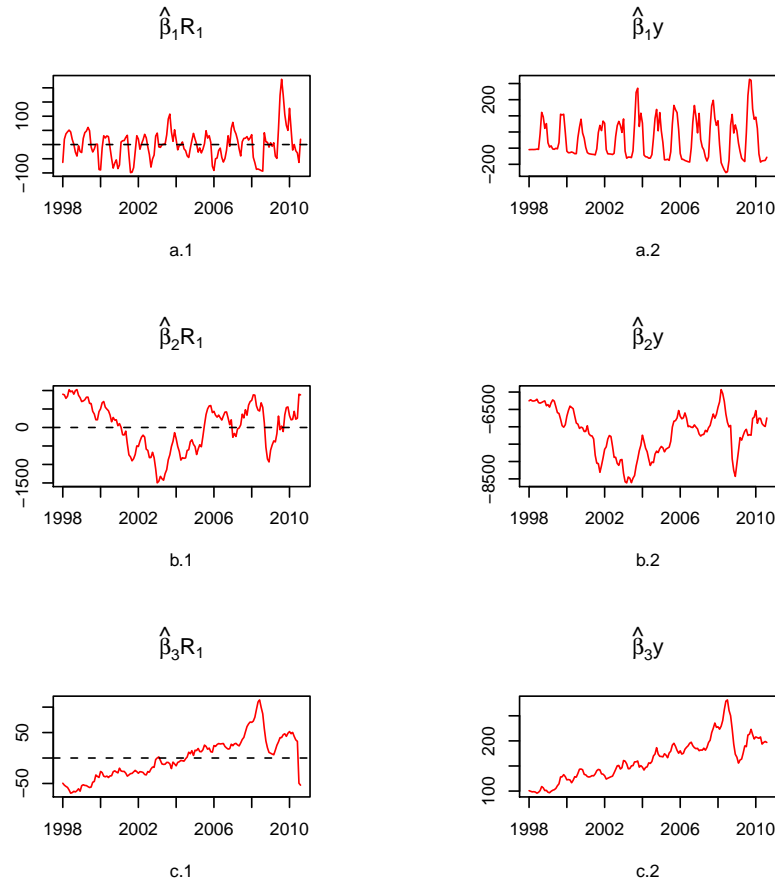


Figura 4.3: gráficas de relaciones de cointegración.

En base a la información de los vectores de cointegración y la matriz de cargas a los vectores propios, el modelo reducido estimado está dado por la ecuación

$$\mathbf{R}_{0t} = \hat{\mathbf{\Pi}}\mathbf{R}_{1t} + \mathbf{e}_t,$$

con

$$\hat{\mathbf{\Pi}} = \begin{pmatrix} -0,05 & 0,69 & -0,38 \\ -0,01 & -0,38 & 0,20 \\ -0,01 & 0,02 & -0,03 \end{pmatrix},$$

Tabla 4.3: valores propios estimados y vectores de cointegración.

$\hat{\lambda}_i$	$\hat{\beta}_i$	TCR	EMPFOB	VBPC
0,28	$\hat{\beta}_1$	-1,20	1,00	-0,1
0,05	$\hat{\beta}_2$	-82,27	0,00	1,00
0,02	$\hat{\beta}_3$	1,00	0,04	0,14

Tabla 4.4: matriz de cargas a los vectores propios.

Variable	$\hat{\alpha}_1$	$\hat{\alpha}_2$	$\hat{\alpha}_3$
TCR	-0,01	0,37	0,21
EMPFOB	6,92	-3,98	1,04
VBPC	-0,01	0,02	-0,04

y cuya relación de cointegración estacionaria es

$$\text{EMPFOB}_t = 1,2 \times \text{TCR}_t + 0,1 \times \text{VBPC}_t.$$

En la Tabla 4.5 se muestra un análisis de rendimiento con predicciones hechas por el MCE para el periodo Septiembre 2010, Marzo 2011. Se deduce que la relación de cointegración logra capturar la estructura estacional del nivel de exportación al mundo en dólares FOB de paltas en estado fresco y de acuerdo a esto, podemos decir con una confianza del 90% que para el periodo 2011 no se esperan grandes volúmenes de exportación como los registrados en el periodo 2007 y 2010 sino que un volumen mas bien similar al registrado en el año 2008 y en el que el punto más alto se alcanza durante el mes de Octubre de 2010 con un valor de predicción esperado cercano a 30.000 dólares FOB. El índice MAPE dado por $\frac{1}{n} \sum_{t=1}^n \left| \frac{rv_t - f_t}{rv_t} \right|$ donde rv_t denota el valor real en el tiempo t y f_t el valor predicho, es igual a 14%. Esta medida, entrega indicios que la relación de cointegración que explica la exportación de paltas tiene una exactitud igual a un 86%.

Tabla 4.5: predicciones con un 90% de confianza para la exportación en dólares FOB de paltas en estado fresco para el periodo Sep-2010 a Mar-2011.

Fecha	Predicción	Valor Real
Sep 2010	22.699,8	15.145,8
Oct 2010	28.822,9	27.114,6
Nov 2010	21.024,7	25.053,8
Dic 2010	20.013,4	21.155,1
Ene 2011	17.002,9	19.202,1
Feb 2011	7.407,8	7.174,9
Mar 2011	3.431,4	3.214,8

La Figura 4.4 muestra los gráficos índices h_{max} y h_2 para la estimación ML de θ , $\hat{\theta}$. Se observa una discrepancia en las observaciones influyentes

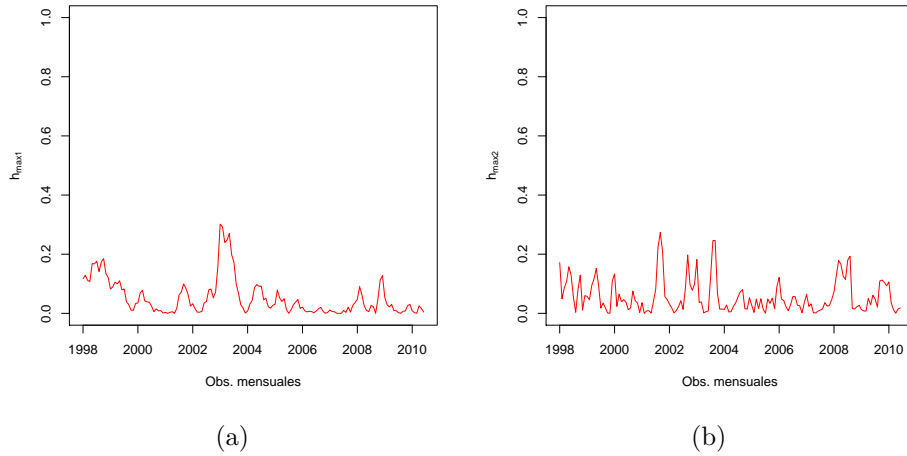


Figura 4.4: (a) gráfica índice del máximo vector propio, h_{max} para $\hat{\theta}$ y (b) gráfica índice del segundo vector propio, h_2 para $\hat{\theta}$.

para ambos vectores propios. Sin embargo, basándonos exclusivamente en el gráfico índice de las componentes de h_{max} podemos decir que una observación potencialmente influyente para la estimación ML, $\hat{\theta}$ en el MCE es la dada en Enero, Febrero y Marzo del año 2003.

Estudiamos el cambio relativo definido como $CR = \frac{\|(\hat{\beta}_{est} - \hat{\beta}_I)\|}{\|\hat{\beta}_{est}\|}$ donde $\hat{\beta}_{est}$ denota la estimación ML de los parámetros dados en de la relación

lineal estacionaria del MCE y $\hat{\beta}_I$ denota la estimación ML de $\hat{\beta}_{est}$ después de corregir mediante un promedio móvil las observaciones potencialmente influyentes de 2003 indexadas en I . La tabla 4.6 muestra el cambio relativo de $\hat{\beta}_{est}$. Observamos que para la relación de equilibrio estacionaria dada por el MCE, la variación máxima en la estimación de sus parámetros es de un 13% y se produce cuando modificamos las observaciones registradas en Enero, Febrero y Marzo de 2003.

Tabla 4.6: cambio relativo de $\hat{\beta}_{est}$.

I:Obs. Influyente 2003	$\hat{\beta}_{TCR}$	$\hat{\beta}_{VBPC}$	CR
Ene, Feb, Mar	1,05	0,10	0,13
Ene, Feb	1,06	0,10	0,12
Ene, Mar	1,09	0,10	0,09
Feb, Mar	1,16	0,11	0,03
Ene	1,08	1,13	0,10
Feb	1,17	0,12	0,03
Mar	1,19	0,10	0,01

Con el fin de comprobar el óptimo funcionamiento de la metodología propuesta, se modificó las observaciones registradas el mes de septiembre de 2009 de cada una de las componentes de la serie multivariada, debido a que en esta fecha se registró el mayor valor en dólares FOB de exportación de paltas al mundo en estado fresco. La modificación se realizó adicionando la norma de cada vector a su valor correspondiente a la fecha indicada (esto transforma un dato extremo en un dato aún más extremo) y luego se replicó el método de estimación y el análisis de influencia local. El resultado entrega las gráficas que se muestran en la Figura 4.5 De la gráfica se observa que dada la modificación realizada, el método de influencia local propuesto detecta eficientemente la potencial observación influyente.

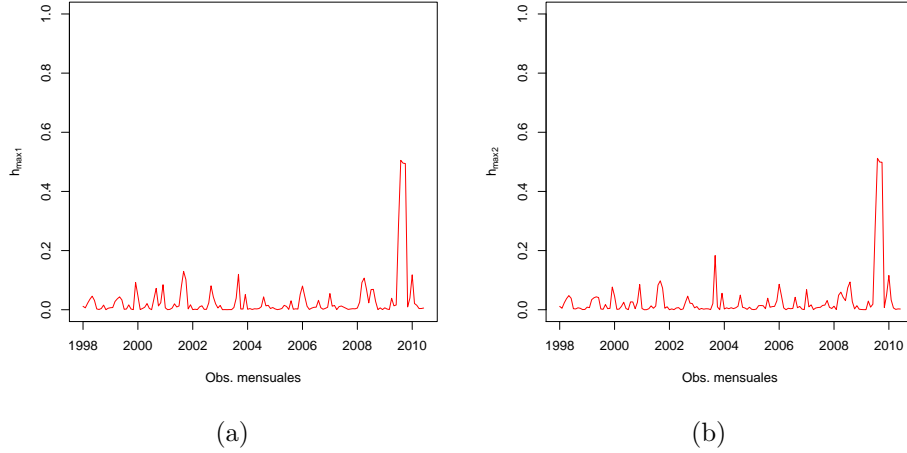


Figura 4.5: (a) gráfica índice del máximo vector propio, h_{max} para $\hat{\theta}$ con valores modificados en sep-2009 y (b) gráfica índice del segundo vector propio, h_2 para $\hat{\theta}$ con valores modificados en sep-2009

4.2 Conclusiones estudio de caso

Observamos que el comportamiento conjunto del proceso trivariado cuyas componentes están dadas por las series de exportación al mundo en dólares FOB de paltas en estado fresco, el tipo de cambio real peso/dólar y el valor por barril de petróleo crudo puede ser descrito efectivamente por un modelo VAR(2). De esta manera, consideramos un modelo VAR(2) en el modelo de corrección de errores dado por

$$\Delta \mathbf{y}_t = \boldsymbol{\nu} + \boldsymbol{\psi} \mathbf{d}_t + \hat{\mathbf{A}}_1 \Delta \mathbf{y}_{t-1} + \hat{\boldsymbol{\alpha}} \hat{\boldsymbol{\beta}}' \mathbf{y}_{t-2} + \boldsymbol{\epsilon}_t, \quad (4.2)$$

cuya relación de equilibrio estacionaria resulta estar determinada por la ecuación

$$\text{EMPF}_{t-1} = 1,2 \times \text{TCR}_t + 0,1 \times \text{VBPC}_t.$$

En base a lo anterior, las predicciones hechas por el modelo (4.2) para el valor de exportación al mundo en dólares FOB de paltas en estado fresco auguran valores normales para la temporada 2010-2011 sin mediar comportamientos climáticos y ambientales anormales en las regiones o áreas de cultivo.

Desde la figura (4.3-a.1) se desprende que una posible observación influyente

es la registrada en Marzo de 2009. Sin embargo, el análisis de influencia local muestra que las observaciones que mayor influencia ejercen bajo el supuesto de heterocedasticidad en la estimación ML de los parámetros del MCE son las dadas en Enero, Febrero y Marzo de 2003. Esto concuerda con antecedentes entregados por la ODEPA en su estudio Mercado de las Paltas 2009-2010. Si bien el modelo tiene un porcentaje de exactitud del 86%, es recomendable tener precaución en las decisiones o acciones a desarrollar en base a estos resultados debido a que se muestra que las observaciones con mayor influencia sobre la estimación de los parámetros de la relación de equilibrio fueron registradas hace casi 7 años atrás a la fecha de emisión de este trabajo. Por último, cabe señalar que los resultados presentados en esta aplicación no pueden ser generalizados ni extendidos a series de datos de otros productos frutícolas u hortícolas debido a que éstos son representativos solamente de las componentes en estudio anteriormente indicadas.

CONCLUSIONES

El análisis de cointegración provee un método para analizar modelos que describen el co-movimiento dinámico de un proceso estocástico multivariado no estacionario. Utilizando la metodología de Johansen, establecimos mediante la estimación previa de un modelo VAR(p) que es posible encontrar un MCE que corrige las tendencias determinísticas o estocásticas presentes en la serie temporal y entrega relaciones de equilibrio estacionarias. En particular encontramos la relación dada por la ecuación

$$\text{EMPFOB}_t = 1,2 \times \text{TCR}_t + 0,1 \times \text{VBPC}_t.$$

En ella, logramos explicar que el comportamiento en el largo plazo de las exportaciones al mundo en dólares FOB de paltas en estado fresco está explicado en un 86% por el tipo de cambio real peso/dólar y el valor por barril de petróleo crudo multiplicados ambos por un factor estimado.

La inclusión de un MCE deriva del hecho que existan grandes desviaciones de observaciones que no respondan al hecho que una relación cointegrada permanezca en un estado de equilibrio. Estas grandes desviaciones pueden tener una fuerte influencia en las estimaciones máximo verosímiles del MCE pero no necesariamente pueden ser identificadas mediante un análisis gráfico de las relaciones de cointegración.

Mediante el diagnóstico de influencia mostramos evidencia empírica útil para

identificar observaciones influyentes en las estimaciones máximo verosímil de los parámetros del modelo de corrección de errores el cual entrega indicios que las observaciones registradas en Enero, Febrero y Marzo de 2003 son influyentes en la estimación de los parámetros de largo plazo bajo el esquema de perturbación propuesto. En particular perturbamos el parámetro de escala del modelo postulado el cual nos permite estudiar el supuesto de homocedastidad. La metodología aquí descrita, permite concluir mediante el diagnóstico de influencias en sistemas econométricos cointegrados que el supuesto de homocedasticidad en los modelos de corrección de errores es robusto a cambios de escala y que además, nos permite identificar de manera eficiente aquellas observaciones que ejercen mayor influencia en la estimación máximo verosímil del modelo. Sin embargo un buen ejercicio futuro es probar esta metodología para errores provenientes de otras distribuciones estadísticas como, por ejemplo, la distribución t-Student para la identificación de outliers.

Cabe mencionar además, que las técnicas presentadas en este trabajo de tesis pueden ser implementadas en cualquier software estadístico que permita la realización de las iteraciones o rutinas computacionales aquí provistas (Apéndice B) ya que son de fácil implementación computacional y están creadas para su libre uso y distribución.

BIBLIOGRAFÍA

- Alexander, C. (2001). *Market Models: A Guide to Financial Data Analysis*. Wiley, New York.
- Página web, Banco Central de Chile. URL:<http://www.bcentral.cl>.
- Banerjee, A. and Hendry D.F., (1992). Testing integration and cointegration: An overview *Oxford Bulletin of Economics and Statistics* **54**, 225-255.
- Bohn, H., (2008). Influential observations in cointegrated VAR models: Danish money demand 1973–2003 *Econometrics Journal* **11**, 39-57.
- Cademartori, D., Romo, C., Campos, R., Galea, M., (2003). Robust estimation of systematic risk using the t-distribution in the chilean stock markets *Applied Economics Letters* **10**, 447-453.
- Cadigan, N., Farrell, P., (2002). Generalized local influence with applications to fish stock cohort analysis *Journal of Royal Statistical Society, Series B* **51**, 469-483.
- Cook, D.,(1986). Assessment of local influence *Journal of Royal Statistical Society, Series B* **48**, 133-169.
- Engle, R.F. and Granger W.J., (1987). Co-integration and error correction: Representation, estimation and testing *Econometrica* **55**, 251-276.
- Franses, P.H. and Lucas A., (1998). Outlier detection in cointegration analysis *Journal of Business and Economics Statistics* **16**, 459-468.

- Granger, W.J.,(1981). Long memory relationships and aggregation of dynamic models *Journal of Econometrics* **14**, 227-238.
- Granger, W.J. and Newbold P.E.,(1974). Spurious regression in econometrics *Journal of Econometrics* **2**, 111-120.
- Hamilton, J.D., (1994). *Times Series Analysis*. Princeton University Press, Princeton.
- Hendry, D.F., (1995). *Dynamic Econometrics*. Oxford University Press, London.
- Johansen, S.,(1988). Statistical analysis of cointegration vectors. *Journal of Economics Dynamics and Control* **12**, 231-254.
- Johansen, S.,(1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autorregressive models. *Econometrica* **59**, 1551-1581.
- Juselius, K.,(2006). *The Cointegrated VAR Model: Methodology and Applications*. Oxford University Press, London.
- Lütkepohl, H., (2006). *New Introduction to Multiple Time Series Analysis*. [2da edición] Springer, New York.
- Magnus, J., Neudecker H., (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. [3ra edición] Willey, New York.
- Moretín, P., (2006). *Econometria Financeira. Um Curso em Series Temporais Financeiras*. Associacao brasileira de estatistica, Sao Paulo.
- Oficina de Estudios y Políticas Agrarias del Gobierno de Chile (2010). Mercado de la Palta 2009-2010. URL:<http://www.odepa.gob.cl>.
- Osorio, F., (2006). *Diagnóstico de Influência em Modelos Elípticos com Efeitos Mistos*. Tesis no publicada. Universidade de Sao Paulo.
- Pan, J.X, Fang, K.T, (2000). *Growth Curve Models and Statistical Diagnostics*. Springer, New York.
- Rachev, S.T., Mittnik, S., Fabozzi, F.J., Focardi, S.M., Jasic, T., (2007). *Financial Econometrics; From Basic to Advanced Modeling Techniques*. Wiley, New York.

R Development Core Team (2008). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

Zhang, X., Yang, B., Zhang, T., Zhang, S., (2001). Local influence on the error-correction variable in a cointegrated system. *Journal of Systems Engineering & Electronics* **12**, 1-8.

Zhang, X., King, L. K., (2005). Influence diagnostics in generalized autoregressive conditional heterocedasticity processes *Journal of Business & Economic Statistics* **23**, 118-129.

APENDICE A

En este apéndice se presentan los cálculos para determinar la curvatura normal en el MCE. En una primera parte se muestra el desarrollo para obtener la matriz de información y la matriz $\mathbf{\Delta}$, ambas dados en el capítulo 3.

Sean las matrices $\boldsymbol{\alpha}$ de orden $(K \times r)$, $\boldsymbol{\beta}$ de orden $(K \times r)$, $\boldsymbol{\Sigma}$ de orden $(K \times K)$, \mathbf{R}_{0t} de orden $(K \times 1)$ y \mathbf{R}_{1t} de orden $(K \times 1)$. La función de log-verosimilitud dada por

$$\ell(\boldsymbol{\theta}) = -\frac{1}{2} \left\{ \sum_{t=0}^T \log(\det(\boldsymbol{\Sigma})) + \sum_{t=1}^T \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right) \right\}$$

y la matriz de información

$$\ddot{\mathbf{L}} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}'}$$

entonces para el cálculo de la matriz de información tenemos lo siguiente:

$$\begin{aligned} d\boldsymbol{\alpha}\ell(\boldsymbol{\theta}) &= -\frac{1}{2} \left\{ \sum_{t=1}^T d \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right) \right\} \\ &= \sum_{t=1}^T \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\alpha})\boldsymbol{\beta}'\mathbf{R}_{1t} \\ &= \sum_{t=1}^T \text{tr} \left[\boldsymbol{\beta}'\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\alpha}) \right] \\ &= \sum_{t=1}^T \text{tr} \left[\left(\boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha}\boldsymbol{\beta}'\mathbf{R}_{1t} \right) \mathbf{R}_{1t}'\boldsymbol{\beta} \right)' d(\boldsymbol{\alpha}) \right] \end{aligned}$$

$$\begin{aligned}
&= \sum_{t=1}^T \text{vec}' \left[\Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \mathbf{R}'_{1t} \boldsymbol{\beta} \right] d(\text{vec } \boldsymbol{\alpha}) \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \text{vec } \boldsymbol{\alpha}} &= \sum_{t=1}^T \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \mathbf{R}'_{1t} \boldsymbol{\beta} \\
d^2_{\boldsymbol{\alpha}} \ell(\boldsymbol{\theta}) &= \sum_{t=1}^T d \left[\left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\boldsymbol{\alpha}) \boldsymbol{\beta}' \mathbf{R}_{1t} \right] \\
&= - \sum_{t=1}^T \text{tr} \left(\mathbf{R}'_{1t} \boldsymbol{\beta} d(\boldsymbol{\alpha}') \Sigma^{-1} d(\boldsymbol{\alpha}) \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \\
&= - \sum_{t=1}^T \text{tr} \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \boldsymbol{\beta} d(\boldsymbol{\alpha}') \Sigma^{-1} d(\boldsymbol{\alpha}) \right) \\
&= - \sum_{t=1}^T d(\text{vec}' \boldsymbol{\alpha}) \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \boldsymbol{\beta} \otimes \Sigma^{-1} \right) d(\text{vec } \boldsymbol{\alpha}) \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vec } \boldsymbol{\alpha} \partial \text{vec } \boldsymbol{\alpha}'} &= - \sum_{t=1}^T \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \boldsymbol{\beta} \otimes \Sigma^{-1} \right) \\
&= -T \left(\boldsymbol{\beta}' \mathbf{S}_{11} \boldsymbol{\beta} \otimes \Sigma^{-1} \right) \\
d^2_{\boldsymbol{\beta} \boldsymbol{\alpha}} \ell(\boldsymbol{\theta}) &= - \sum_{t=1}^T \text{tr} \left(\mathbf{R}'_{1t} d(\boldsymbol{\beta}) \boldsymbol{\alpha}' \Sigma^{-1} d(\boldsymbol{\alpha}) \boldsymbol{\beta} \mathbf{R}_{1t} \right) \\
&= - \sum_{t=1}^T d(\text{vec}' \boldsymbol{\alpha}) \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \otimes \Sigma^{-1} \boldsymbol{\alpha} \right) d(\text{vec}' \boldsymbol{\beta}) \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vec } \boldsymbol{\beta} \partial \text{vec } \boldsymbol{\alpha}'} &= - \sum_{t=1}^T \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \mathbf{R}'_{1t} \otimes \Sigma^{-1} \boldsymbol{\alpha} \right) \\
&= -T \left(\boldsymbol{\beta}' \mathbf{S}_{11} \otimes \Sigma^{-1} \boldsymbol{\alpha} \right)
\end{aligned}$$

$$\begin{aligned}
d_{\Sigma}^2 \ell(\boldsymbol{\theta}) &= - \sum_{t=1}^T d \left[\left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\alpha}) \boldsymbol{\beta}' \mathbf{R}_{1t} \right] \\
&= - \sum_{t=1}^T \text{tr} \left[\boldsymbol{\beta}' \mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\alpha}) \right] \\
&= - \sum_{t=1}^T d(\text{vec}' \boldsymbol{\alpha}) \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \right) d(\text{vec } \boldsymbol{\Sigma}) \\
&= - \sum_{t=1}^T d(\text{vec}' \boldsymbol{\alpha}) \left(\boldsymbol{\beta}' \mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \right) \mathbf{D}_K d(\text{vech } \boldsymbol{\Sigma})
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vech } \boldsymbol{\Sigma} \partial \text{vec } \boldsymbol{\alpha}'} &= - \sum_{t=1}^T \mathbf{D}'_K \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \mathbf{R}'_{1t} \boldsymbol{\beta} \right) \\
&= -T \mathbf{D}'_K \left(\boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\Sigma}^{-1} \left(\mathbf{S}_{01} \boldsymbol{\beta} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{S}_{11} \boldsymbol{\beta} \right) \right)
\end{aligned}$$

$$\begin{aligned}
d_{\boldsymbol{\beta}} \ell(\boldsymbol{\theta}) &= \sum_{t=1}^T d \left[\left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d(\boldsymbol{\beta}') \mathbf{R}_{1t} \right] \\
&= \sum_{t=1}^T \text{tr} \left[\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d(\boldsymbol{\beta}') \right] \\
&= \sum_{t=1}^T \text{tr} \left[\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \right]' d(\boldsymbol{\beta}) \\
&= \sum_{t=1}^T \text{vec}' \left(\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \right) d(\text{vec } \boldsymbol{\beta}) \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \text{vec } \boldsymbol{\beta}} &= \sum_{t=1}^T \mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \\
d_{\boldsymbol{\beta}}^2 \ell(\boldsymbol{\theta}) &= - \sum_{t=1}^T d \left[\mathbf{R}'_{1t} d(\boldsymbol{\beta}) \boldsymbol{\alpha}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d(\boldsymbol{\beta}') \mathbf{R}_{1t} \right] \\
&= - \sum_{t=1}^T \text{tr} \left[\boldsymbol{\alpha}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d(\boldsymbol{\beta}') \mathbf{R}_{1t} \mathbf{R}'_{1t} d(\boldsymbol{\beta}) \right]
\end{aligned}$$

$$\begin{aligned}
&= -\sum_{t=1}^T d(\text{vec}' \boldsymbol{\beta}) \left(\boldsymbol{\alpha}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \otimes \mathbf{R}_{1t} \mathbf{R}'_{1t} \right) d(\text{vec } \boldsymbol{\beta}) \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vec } \boldsymbol{\beta} \partial \text{vec } \boldsymbol{\beta}'} &= -T \left(\boldsymbol{\alpha}' \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \otimes \mathbf{S}_{11} \right) \\
d_{\boldsymbol{\Sigma}} \ell(\boldsymbol{\theta}) &= -\sum_{t=1}^T \text{tr} \left[\left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d(\boldsymbol{\beta}') \mathbf{R}_{1t} \right] \\
&= -\sum_{t=1}^T \text{tr} \left[\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\Sigma}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} d\boldsymbol{\beta}' \right] \\
&= -\sum_{t=1}^T d(\text{vec } \boldsymbol{\beta}) \left(\mathbf{R}_{1t} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \otimes \boldsymbol{\alpha} \boldsymbol{\Sigma}^{-1} \right) d(\text{vec } \boldsymbol{\Sigma}) \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vech } \boldsymbol{\Sigma} \partial \text{vec } \boldsymbol{\beta}'} &= \sum_{t=1}^T \mathbf{D}'_K \left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \otimes \boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \mathbf{R}'_{1t} \right) \\
&= -T \mathbf{D}'_K \left(\boldsymbol{\Sigma}^{-1} \boldsymbol{\alpha} \otimes \boldsymbol{\Sigma}^{-1} \left(\mathbf{S}_{01} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{S}_{11} \right) \right) \\
d_{\boldsymbol{\Sigma}} \ell(\boldsymbol{\theta}) &= \frac{1}{2} \sum_{t=1}^T \text{tr} \left(\boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} d(\boldsymbol{\Sigma}) \right) \\
&= \frac{1}{2} \sum_{t=1}^T d\text{vec}' \left(\boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \right) d(\text{vec } \boldsymbol{\Sigma}) \\
\frac{\partial \ell(\boldsymbol{\theta})}{\partial \text{vech } \boldsymbol{\Sigma}} &= \frac{1}{2} \sum_{t=1}^T \text{vec}' \left(\boldsymbol{\Sigma}^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \boldsymbol{\Sigma}^{-1} \right) \mathbf{D}_K
\end{aligned}$$

$$\begin{aligned}
d_{\Sigma}^2 \ell(\boldsymbol{\theta}) &= \sum_{t=1}^T \text{tr} \left[d(\Sigma^{-1}) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\Sigma) \right] \\
&+ \text{tr} \left[\Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' d(\Sigma^{-1}) d(\Sigma) \right] \\
&= - \sum_{t=1}^T \text{tr} \left[d(\Sigma^{-1}) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\Sigma) \right] \\
&= - \sum_{t=1}^T d(\text{vec}' \Sigma) \left(\mathbf{I}_K \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \right) d(\text{vec} \Sigma^{-1}) \\
&= - \sum_{t=1}^T (d(\text{vech} \Sigma^{-1}) \mathbf{D}_K)' \\
&\quad \times \left(\mathbf{I}_K \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \right) (\mathbf{D}_K d(\text{vech} \Sigma^{-1}))
\end{aligned}$$

$$\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \text{vech} \Sigma \partial \text{vech} \Sigma} = - \sum_{t=1}^T \mathbf{D}'_K \left(\mathbf{I}_K \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \right) \mathbf{D}_K$$

La matriz $\boldsymbol{\Delta} = (\boldsymbol{\Delta}_1, \boldsymbol{\Delta}_2, \dots, \boldsymbol{\Delta}_T)$ la obtenemos por medio de las siguientes diferenciaciones

$$d_{\boldsymbol{\omega}_t} \ell(\boldsymbol{\theta} | \boldsymbol{\omega}) = \frac{1}{2\boldsymbol{\omega}_t} d(\boldsymbol{\omega}_t) - \frac{1}{2} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) d(\boldsymbol{\omega}_t)$$

$$\begin{aligned}
d_{\boldsymbol{\alpha} \boldsymbol{\omega}_t}^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega}) &= \frac{1}{2} \mathbf{R}'_{1t} \boldsymbol{\beta} d(\boldsymbol{\alpha}') \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) d(\boldsymbol{\omega}_t) \\
&\quad + \frac{1}{2} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\boldsymbol{\alpha}) \boldsymbol{\beta}' \mathbf{R}_{1t} d(\boldsymbol{\omega}_t) \\
&= \mathbf{R}'_{1t} \boldsymbol{\beta} d(\boldsymbol{\alpha}') \Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) d(\boldsymbol{\omega}_t) \\
&= \text{tr} \left[\Sigma^{-1} \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) d(\boldsymbol{\omega}_t) \mathbf{R}'_{1t} \boldsymbol{\beta} d(\boldsymbol{\alpha}') \right] \\
&= d(\text{vec}' \boldsymbol{\alpha}) \left(\left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right) \otimes \mathbf{R}'_{1t} \boldsymbol{\beta} \right) d(\boldsymbol{\omega}_t) \\
\frac{\partial^2 \ell(\boldsymbol{\theta} | \boldsymbol{\omega})}{\partial \text{vec} \boldsymbol{\alpha} \partial \boldsymbol{\omega}_t} &= \left(\mathbf{R}_{0t} - \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{R}_{1t} \right)' \Sigma^{-1} \otimes \mathbf{R}'_{1t} \boldsymbol{\beta}
\end{aligned}$$

$$\begin{aligned}
d_{\beta \omega_t}^2 \ell(\theta | \omega) &= \frac{1}{2} \mathbf{R}'_{1t} d(\beta) \alpha' \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) d(\omega_t) \\
&\quad + \frac{1}{2} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right)' \Sigma^{-1} \alpha d(\beta') \mathbf{R}_{1t} d(\omega_t) \\
&= \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right)' \Sigma^{-1} \alpha d(\beta') \mathbf{R}_{1t} d(\omega_t) \\
&= \text{tr} \left[\mathbf{R}_{1t} d(\omega_t) \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \Sigma^{-1} \alpha d(\beta') \right] \\
&= d(\text{vec}' \beta) \left(\Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \otimes \mathbf{R}_{1t} \right) d(\omega_t) \\
\frac{\partial^2 \ell(\theta | \omega)}{\partial \text{vec } \beta \partial \omega_t} &= \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \otimes \mathbf{R}_{1t}
\end{aligned}$$

$$\begin{aligned}
d_{\Sigma \omega_t}^2 \ell(\theta | \omega) &= \frac{1}{2} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\Sigma) \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) d(\omega_t) \\
&= \frac{1}{2} \text{tr} \left[\Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) d(\omega_t) \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right)' \Sigma^{-1} d(\Sigma^{-1}) \right] \\
&= d(\text{vec}' \Sigma) \left(\Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \right) d(\omega_t) \\
&= d(\text{vech}' \Sigma) \mathbf{D}'_K \left(\Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \right) d(\omega_t) \\
\frac{\partial^2 \ell(\theta | \omega)}{\partial \text{vech } \Sigma \partial \omega_t} &= \mathbf{D}'_K \left(\Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \otimes \Sigma^{-1} \left(\mathbf{R}_{0t} - \alpha \beta' \mathbf{R}_{1t} \right) \right)
\end{aligned}$$

Donde \mathbf{D}_K es una matriz de duplicación de orden $K^2 \times \frac{K(K+1)}{2}$.

APENDICE B

En la presente sección se anexa el script R para la aplicación del diagnóstico de influencia en sistemas econométricos cointegrados.

```
#-----  
#--ESTIMACION DEL MODELO DE CORRECIÓN DE ERRORES--  
#----EN DATOS DE EXPORTACIÓN DE PALTAS-----  
#-----  
  
library(vars)      # librería que desarrolla la estimación de modelos VAR.  
library(tseries)  # librería para el análisis de series de tiempo.  
library(urca)     # librería para el análisis de econometría y cointegración.  
dat01             #Datos Fuente  
attach(dat01)  
names(dat01)  
  
stc<-ts(TCreal,start=c(1998),frequency=12)  
spalta<-ts(Paltas,start=c(1998),frequency=12)  
spetro<-ts(Petroleo,start=c(1998),frequency=12)  
  
mser<-cbind(stc,spalta,spetro)  
cov2cor(cov(mser))  
  
#-----GRAFICAS TEMPORALES-----  
par(mar=rep(5,4))  
plot.ts(spalta,col=3,xlim=c(1998,2010.8),
```

```

ylim=c(-10,1500),
ylab=expression(Y[t]),xlab="Año")
lines(spetro,col=1,lty=2)
par(new=T)
plot.ts(stc,col=2,axes=FALSE,ylab=" ",xlab=" ")
axis(4,pretty(range(stc)))
legend(1998,109,c("TCR","EMPFOB","VBPC"),
col=c(2,3,1),lty=c(1,1,2),horiz=FALSE)

#ts.plot(stc,spalta,spetro,col=c(2,3,1),xlim=c(1998,2010.8)
,ylim=c(-10,1500),
#ylab=expression(Y[t]),xlab="Año")
#legend(1998,1500,c("Tipo Cambio Real","Paltas US$ FOB",
"US$ Barril de Petroleo"),
#col=c(2,3,1),lty=1,horiz=FALSE)

par(mfrow=c(3,2))
acf(ts(TCreal),lag.max=50,xlab="a.1",main="");
pacf(ts(TCreal),xlab="a.2",main="")
acf(ts(Paltas),lag.max=50,xlab="b.1",main="");
pacf(ts(Paltas),xlab="b.2",main="")
acf(ts(Petroleo),lag.max=50,xlab="c.1",main="");
pacf(ts(Petroleo),xlab="c.2",main="")
#acf(sipc,lag.max=50,main="IPC");pacf(sipc,main="IPC")
par(mfrow=c(1,1))

#-----TEST DE DICKEY-FULLER-----
#--para existencia de raices unitarias-----

DFtc<-adf.test(stc)
DFtc$p.value
DFtc$alternative
m1<-ur.df(stc,lags=1,type="drift")
plot(m1); summary(m1)

DFcebo<-adf.test(spalta)
DFcebo$p.value
m2<-ur.df(spalta,type="trend")

```

```
plot(m2); summary(m2)

DFpetro<-adf.test(spetro)
DFpetro$p.value
m3<-ur.df(spetro,type="drift")
plot(m3); summary(m3)

#-----ESTIMACIÓN DEL MODELO VAR-----
stri<-cbind(stc,spalta,spetro)
m.1<-VARselect(stri,lag.max=9,type="const")
modelo1<-VAR(stri,lag.max=2,ic="AIC"),type="both")
summary(modelo1)
Acoef(modelo1)

#-----PRUEBAS DE DIAGNÓSTICO MODELO VAR-----
ppmtu1<-serial.test(modelo1,lags.pt=2,type=c("ES"))
efearch1<-arch.test(modelo1,lags.multi=2,
multivariate.only=TRUE)
norm1<-normality.test(modelo1,multivariate.only = TRUE)

caus1<-causality(modelo1,cause="spalta")
caus2<-causality(modelo1,cause="stc")
caus3<-causality(modelo1,cause="spetro")

predi<-predict(modelo1,n.ahead=4,ci=0.90)
plot(predi,names="spalta")
fanchart(predi,names="spalta")

#-----ANÁLISIS DE COINTEGRACIÓN-----
vecm1<-ca.jo(stri,season=12,type = c("eigen"))
#season=12 es porque tenemos datos mensuales
vecm2<-ca.jo(stri,season=12,type = c("trace"))
summary(vecm1);summary(vecm2)
vecm1.r1<-cajorls(vecm1,r=1)

#-----VECM como VAR en Niveles-----
vecm1.nivel<-vec2var(vecm1,r=1)
serial.test(vecm1.nivel)
```

```

arch.test(vecm1.nivel)
normality.test(vecm1.nivel)
predivecm<-predict(vecm1.nivel,n.ahead=12,ci=0.90)
plot(predivecm,names="spalta",xlab="t",
col=c(4,3,1,1,2),ylab="Miles US$ FOB",main="")

#----ANALISIS GRÁFICO DE COINTERGACIÓN-----
beta01<-vecm1@V
alpa01<-vecm1@W
sigmaest<-vecm1@DELTA
R1<-vecm1@RK
R0<-vecm1@R0
R1t<-t(R1)
R0t<-t(R0)
S11<-(1/T)*(R1t%*%t(R1t))
S01<-(1/T)*(R0t%*%t(R1t))
S10<-t(S01)
beta01[,1]<-beta01[,1]/beta01[2,1]
beta01[,2]<-beta01[,2]/beta01[3,2]
beta01
com1<-ts(R1%*%beta01,start=c(1998),
end=c(2010,8),frequency=12)
com2<-ts(vecm1@x%*%beta01,start=c(1998),
end=c(2010,8),frequency=12)
plot(com1)
plot(com2)
co1.1.2<-ts(R1%*%beta01[,-c(1,2)],start=c(1998),
end=c(2010,8),frequency=12)
co2.1.2<-ts(vecm1@x%*%beta01[,-c(1,2)],start=c(1998),
end=c(2010,8),frequency=12)
co1.2.3<-ts(R1%*%beta01[,-c(2,3)],start=c(1998),
end=c(2010,8),frequency=12)
co2.2.3<-ts(vecm1@x%*%beta01[,-c(2,3)],start=c(1998),
end=c(2010,8),frequency=12)
co1.1.3<-ts(R1%*%beta01[,-c(1,3)],start=c(1998),
end=c(2010,8),frequency=12)
co2.1.3<-ts(vecm1@x%*%beta01[,-c(1,3)],start=c(1998),

```

```

end=c(2010,8),frequency=12)
par(mfrow=c(3,2))
# las siguientes graficas comparan visualmente
# el MCE con plot(com1) y plot(com2)
plot(co1.2.3,col=2,xlab="a.1",ylab="",
main=expression(hat(beta)[1]*R[1]));abline(h=0,lty=2)
plot(co2.2.3,col=2,xlab="a.2",ylab="",
main=expression(hat(beta)[1]*y))
plot(co1.1.3,col=2,xlab="b.1",ylab="",
main=expression(hat(beta)[2]*R[1]));abline(h=0,lty=2)
plot(co2.1.3,col=2,xlab="b.2",ylab="",
main=expression(hat(beta)[2]*y))
plot(co1.1.2,col=2,xlab="c.1",ylab="",
main=expression(hat(beta)[3]*R[1]));abline(h=0,lty=2)
plot(co2.1.2,col=2,xlab="c.2",ylab="",
main=expression(hat(beta)[3]*y))

par(mfrow=c(1,1))
alpaest<-vecm1@PI%*%solve(t(beta01))
piest<-alpaest%*%t(beta01)

#-----
#SCRIPT PARA DIAGNOSTICO DE INFLUENCIA EN SISTEMAS COINTEGRADOS
#-----

# Elementos que se necesitan
# T = N
T<-150
alfa<-alpa01
beta<-beta01
sigma<-sigmaest
# Función vectorización de una matriz
vec <- function(X){
vector <- 0
n      <- dim(X)[1]
k      <- 1
for(i in 1:n){
for(j in 1:n){

```

```

vector[k] <- X[i,j]
k          <- k + 1
}
}
return(vector)
}
# Función para crear la matriz de duplicación
duplication <- function(m){
E <- array(0, dim = c(m, m))
T <- array(0, dim = c(m, m))
u <- array(0, c(m * (m + 1) / 2, 1))
D <- array(0, dim = c(m * m, m * (m + 1) / 2))
for(i in 1:m){
for(j in 1:i){
if(i != j){
E[i,j] <- 1
A      <- E
E      <- array(0, dim = c(m, m))
E[j,i] <- 1
B      <- E
T      <- A + B
E      <- array(0, dim = c(m, m))
}
else{
E[i,i] <- 1
T      <- E
E      <- array(0, dim = c(m, m))
}
if(i >= j){
u[(j - 1) * m + i - j * (j - 1) / 2] <- 1
D                                     <- D + vec(T) %*% t(u)
u                                     <- array(0, c(m * (m + 1) / 2, 1))
}
}
}
return(D)
}
Dk<-duplication(3)

```

```

#-----
# Para Matriz L
# elementos de dentro de cada Laa, Lba, etc...
a1<-(t(beta)%*%S11)%x%(solve(sigma))
a2<-(t(beta)%*%S11)%x%(solve(sigma)%*%alfa)
a3<-(((beta)%*%S10 - t(beta)%*%S11)%x%(t(alfa))
%*%solve(sigma))%x%(solve(sigma))%*%Dk
a4<-(t(alfa)%*%solve(sigma)%*%alfa)%x%S11
a5<-((t(alfa)%*%solve(sigma))%x%((S10 - S11)%*%beta)%*%
t(alfa))%*%solve(sigma))%*%Dk
Iden<-diag(3)
bb<-matrix(0,6,6)
bb
for(i in 1:T){
a<-t(Dk)%*(Iden%x%(solve(sigma)*c((t(R0t[,i]-
(alfa)%*%t(beta)%*%R1t[,i]))%*% (R0t[,i]-(alfa)%*%
t(beta)%*%R1t[,i]))))))%*%Dk
bb<- bb+a
}
bb
# Matriz L
#Laa
Laa<--T*a1 #9*9
#Lba
Lba<--T*a2 #9*9
#Lsa
Lsa<--T*a3 #9*6
#Lbb
Lbb<--T*a4 #9*9
#Lsb
Lsb<--T*a5 #9*6
#Lss
Lss<--bb #6*6
L<-matrix(0,24,24)
L[1:9,1:9]<-Laa
L[10:18,1:9]<-t(Lba)
L[19:24,1:9]<-t(Lsa)
L[1:9,10:18]<-Lba

```

```

L[10:18,10:18]<-Lbb
L[19:24,10:18]<-t(Lsb)
L[1:9,19:24]<-Lsa
L[10:18,19:24]<-Lsb
L[19:24,19:24]<-Lss
L
#-----
# Para Matriz delta
d1<-matrix(0,9,150)
d2<-matrix(0,9,150)
d3<-matrix(0,6,150)
for(i in 1:T){
b1<-t(R0t[,i]-(alfa**t(beta)**R1t[,i]))**
solve(sigma)%x%(t(R1t[,i])**beta)
b1<-t(b1)
d1[,i]<-b1
}
for(i in 1:T){
b2<-(solve(sigma)**(R0t[,i]-alfa**t(beta)**
R1t[,i]))%x%R1t[,i]
b2<-t(b2)
d2[,i]<-b2
}
for(i in 1:T){
b3<-t(Dk)**((solve(sigma)**(R0t[,i]-alfa**t(beta)**
R1t[,i]))%x%(solve(sigma)**(R0t[,i]-alfa**t(beta)**R1t[,i])))
b3<-t(b3)
d3[,i]<-b3
}
# Matriz Delta
delta<-matrix(0,24,150)
delta[1:9,1:150]<-d1
delta[10:18,1:150]<-d2
delta[19:24,1:150]<-d3
delta
#-----
# Matrix B
# L 24*24   delta 24*150

```

```
a<-svd(L)          # descomposición
d<-t(a$u) %*% L %*% a$v    # inversa L, transformación
B<-t(delta)%*%d%*%delta    # matriz B
r1<-eigen(B)$vectors[,1]   # primer vector propio
r2<-eigen(B)$vectors[,2]   # segundo vector propio
sum(r1^2)
sum(r2^2)
mv<-matrix(0,150,150)
for(i in 1:T){
mv[,i]<-eigen(B)$vectors[,i]
}
dv<-diag(mv)          #diagonal de matriz de vectores propios
sum(dv^2)             #plot(abs(B[,1]),type="l", main="Primera columna de B")
#plot(abs(diag(B)),type="l" ,main="Diagonal de B")
vmax1<-ts(abs(r1),start=c(1998),frequency=12)
vmax2<-ts(abs(r2),start=c(1998),frequency=12)
par(mfrow=c(2,1))
plot(vmax1,type="l",xlab="Obs. mensuales",ylab=expression(h [max1]),
main="",col=2)
plot(vmax2,type="l",xlab="Obs. mensuales",ylab=expression(h [max2]),
main="",col=2)
par(mfrow=c(1,1))
vmaxD<-ts(abs(dv),start=c(1998),frequency=12)
plot(vmaxD,type="l",xlab="Obs. mensuales",ylab="Curvatura local total",
main=" ",col=2)
```