



UNIVERSIDAD DEL BÍO-BÍO
FACULTAD DE CIENCIAS

Transmisión Asimétrica de Precios en el
sector de la palta en Chile:
Evidencia desde un modelo TVECM

Autor:

Héctor Garrido Henríquez

Profesor(es) Guía(s):

Dr. Sergio Contreras Espinoza

Dra. Monia Ben Kaabia

*Tesis para optar al grado de Magíster en Matemática con mención
en Estadística*

Departamento de Estadística

Departamento de Matemática

18 de abril de 2018

Índice general

1. Introducción	13
2. Transmisión de precios	15
3. Modelos y contrastes estadísticos utilizados en esta tesis	17
3.1. Pronósticos de una serie temporal según la metodología Box-Jenkins .	17
3.2. Contrastes de raíz unitaria/estacionariedad	17
3.2.1. Contraste de Dickey-Fuller Aumentado	17
3.2.2. Contraste de Phillips Perron (1992)	18
3.2.3. Contraste de Elliot, Rothenberg & Stock (1996)	19
3.2.4. Contraste Kiatkowsky, Pesaran, Schmidt & Shin (1992)	21
3.3. Análisis de cointegración	22
3.3.1. Cointegración y análisis de las relaciones de largo plazo	22
3.4. Estimación basada en la verosimilitud para el modelo VAR irrestricto	25
3.4.1. Contraste de razón de verosimilitud	27
3.4.1.1. El contraste de Ljung-Box	28
3.4.1.2. Contrastes para Heteroscedasticidad Residual	28
3.4.1.3. Contrastes de normalidad	29
3.4.2. Modelo Vectorial de Corrección del Error (VECM)	29
3.4.3. Estimación de un modelo VECM	31
3.4.4. El contraste de cointegración de Johansen	34
3.4.5. Estabilidad del modelo VECM	35
3.5. Modelo de vector de corrección del error por umbrales (TVECM) . .	36
3.5.1. Estimación del modelo TVECM	36
3.5.2. El contraste de Hansen & Seo (2002)	38
3.5.3. Función de impulso respuesta no lineal	40

3.6.	Imputación de datos faltantes a través del filtro de Kalman	40
3.6.1.	El filtro de Kalman	40
3.6.2.	Formulación en espacio estado de un modelo ARIMA(p,d,q) .	40
3.6.3.	imputación de valores faltantes	40
4.	Metodología y Datos	41
4.1.	Fuentes de información	41
5.	Resultados	43
5.1.	Imputación de valores perdidos	43
5.2.	Análisis del orden de integración de las series	45
5.2.1.	Análisis gráfico	45
5.2.1.1.	Análisis de las series en nivel	45
5.2.1.2.	Análisis de las series en diferencias	46
5.2.2.	Contrastes de raíz unitaria	48
5.2.2.1.	Análisis de las series originales	48
5.2.2.2.	Análisis de las series en diferencias	51
5.3.	Determinación del rango de cointegración y estimación del modelo del modelo lineal de corrección del error	53
5.4.	Diagnósticos del modelo lineal	58
5.5.	Determinación del rango de cointegración y estimación del modelo del modelo lineal de corrección del error	60
5.6.	Diagnósticos del modelo no lineal	62
6.	Conclusiones	65
7.	Bibliografía	67
8.	Anexo	69
8.1.	Algunos resultados útiles de cálculo matricial	69
8.1.1.	Derivada de una forma cuadrática	69
8.1.2.	Derivada de un determinante	70

Índice de figuras

5.1. Evolución de precios del palta Hass de primera calidad,2008-2016 . . .	44
5.2. Imputación de valores perdidos a través del filtro de Kalman	44
5.3. Evolución del logaritmo del precio mayorista de la palta, 2008-2016 .	45
5.4. Evolución del logaritmo del precio en supermercado de la palta, 2008- 2016	46
5.5. Evolución del logaritmo del precio mayorista de la palta, 2008-2016 .	47
5.6. Evolución del logaritmo del precio en supermercado de la palta, 2008- 2016	47
5.7. Número de Rezagos para el contraste de Independencia	55
5.8. Función ortogonal de impulso respuesta	56
5.9. Función ortogonal de impulso respuesta	57
5.10. Función acumulada de impulso respuesta	58
5.11. Función acumulada de impulso respuesta	58
5.12. Residuos del modelo	59
5.13. Histograma bivariado de los residuos del modelo	59
5.14. Mapa de calor de los residuos del modelo	60
5.15. Modelo de corrección del error por umbrales	62
5.16. Histograma bivariado de los residuos del modelo	63
5.17. Mapa de calor de los residuos del modelo	63

Índice de cuadros

3.1. Hipótesis del contraste de Dickey-Fuller	18
5.1. Contraste de Dickey-Fuller aumentado (con drift)	48
5.2. Contraste de Dickey-Fuller aumentado (con tendencia)	49
5.3. Contraste KPSS (con tendencia determinista)	49
5.4. Contraste KPSS (sin tendencia determinista)	50
5.5. Contraste Phillips & Perron ^a (con tendencia determinista)	50
5.6. Contraste Phillips & Perron ^a (con tendencia determinista)	50
5.7. Contraste de Elliot, Rothenberg & Stock (con tendencia determinista)	51
5.8. Contraste de Elliot, Rothenberg & Stock (con constante)	51
5.9. Contraste de Dickey-Fuller aumentado (sin drift y tendencia)	52
5.10. Contraste KPSS (sin tendencia determinista)	52
5.11. Contraste Phillips & Perron ^a (con tendencia determinista)	53
5.12. Contraste Phillips & Perron ^a (con tendencia determinista)	53
5.13. Contraste de Elliot, Rothenberg & Stock (con tendencia determinista)	54
5.14. Contraste de Elliot, Rothenberg & Stock (con constante)	54
5.15. Contraste de la <i>la traza</i> de cointegración de Johansen	55
5.16. Contraste del <i>máximo autovalor</i> de cointegración de Johansen	56
5.17. Contraste de Elliot, Rothenberg & Stock (con constante)	57
5.18. Diagnósticos del modelo	60

Agradecimientos

Abstract

Capítulo 1

Introducción

Capítulo 2

Transmisión de precios

Meyer, J., & Cramon-Taubadel, S. (2004). Asymmetric price transmission: a survey. *Journal of agricultural economics*, 55(3), 581-611.

Capítulo 3

Modelos y contrastes estadísticos utilizados en esta tesis

En este capítulo se exponen los principales modelos y contrastes estadísticos utilizados en esta tesis. Se comienza por la descripción de la estimación de modelos de series de tiempo univariados según la metodología de Box-Jenkins y los contrastes de raíz unitaria univariantes que permiten realizar una exploración preliminar de los datos, para luego abordar la formulación del modelo de regresiones vectoriales autorregresivas (**VAR**), el modelo vectorial de corrección del error (**VECM**), el contraste de cointegración de Johansen, el modelo vectorial de corrección del error por umbrales (**TVECM**) y el contraste de Hansen y Seo (2002). Además, se aborda el problema de imputación de datos faltantes a través del filtro de Kalman.

3.1. Pronósticos de una serie temporal según la metodología Box-Jenkins

EN ESTE APARTADO DEBO CITAR EL LIBRO DE BOX JENKINS Y REINSEL

3.2. Contrastes de raíz unitaria/estacionariedad

3.2.1. Contraste de Dickey-Fuller Aumentado

El contraste más utilizado en la investigación aplicada, dada su simplicidad, es el contraste propuesto por [fuller1976] y [dickey1981]. Para aplicar este contraste existen dos posibles modelos

Si y_t satisface la siguiente ecuación

$$y_t = \alpha + \rho y_{t-1} + \epsilon_t \quad (t = 1, \dots, n) \quad (3.2.1)$$

Donde $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

Si y_t satisface la siguiente ecuación

Como puede observarse en el cuadro 3.1, existen 3 estadísticos, Φ_1 , Φ_2 y Φ_3 , y sus respectivas hipótesis que pueden ser utilizados. Mientras Φ_1

$$y_t = \alpha + \beta \left(t - 1 - \frac{1}{2}n \right) + \rho y_{t-1} + \epsilon_t \quad (t = 1, \dots, n) \quad (3.2.2)$$

Donde $\epsilon_t \sim \mathcal{N}(0, \sigma^2)$.

Cuadro 3.1: Hipótesis del contraste de Dickey-Fuller

Estadístico	\mathcal{H}_0	\mathcal{H}_a
τ	$\rho = 1$	$\rho = 0$
Φ_1	$(\alpha, \rho) = (0, 1)$	$(\alpha, \rho) \neq (0, 1)$
Φ_2	$(\alpha, \beta, \rho) = (0, 0, 1)$	$(\alpha, \beta, \rho) \neq (0, 0, 1)$
Φ_3	$(\alpha, \beta, \rho) = (\alpha, 0, 1)$	$(\alpha, \beta, \rho) \neq (\alpha, 0, 1)$

Fuente: Elaboración propia basado en Dickey y Fuller (1981)

3.2.2. Contraste de Phillips Perron (1992)

De manera similar al contraste anterior [phillips1988] proponen un contraste no paramétrico para la hipótesis nula de raíz unitaria. A diferencia del contraste de Dickey Fuller este contraste resiste dependencia débil y heteroscedasticidad del término de error. El contraste está construido sobre la base de las siguientes formas funcionales:

$$y_t = \mu + \alpha y_{t-1} + \varepsilon_t, \quad (3.2.3)$$

$$y_t = \mu + \beta \left(t - \frac{1}{2}T \right) + \alpha y_{t-1} + \varepsilon_t \quad (3.2.4)$$

Luego de esto definen los siguientes estadísticos de prueba:

$$Z(\hat{\alpha}) = T(\hat{\alpha} - 1) - \hat{\lambda}/\bar{m}_{yy}, \quad (3.2.5)$$

$$Z(\tau_{\hat{\alpha}}) = (\hat{s}/\hat{\sigma}_{Tl})t_{\hat{\alpha}} - \hat{\lambda}'\hat{\sigma}_{Tl}/\bar{m}_{yy}^{1/2}, \quad (3.2.6)$$

$$Z(\tau_{\hat{\mu}}) = (\hat{s}/\hat{\sigma}_{Tl})t_{\hat{\mu}} + \hat{\lambda}'\hat{\sigma}_{Tl}m_y/\bar{m}_{yy}^{1/2}m_{yy}^{1/2} \quad (3.2.7)$$

Donde $\bar{m}_{yy} = T^{-2} \sum (y_t - \bar{y})^2$, $m_{yy} = T^{-2} \sum y_t^2$, $m_y = T^{-3/2} \sum y_t$ y $\hat{\lambda} = \frac{1}{2}(\hat{\sigma}_{Tl}^2 - \hat{s}^2)$, donde $\hat{\sigma}^2$ es la varianza muestral de los residuos, $\hat{\lambda}' = \hat{\lambda}/\hat{\sigma}_{Tl}^2$. Luego, la varianza de largo plazo es estimada de la siguiente forma:

$$\hat{\sigma}_{Tl}^2 = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 + 2T^{-1} \sum_{s=1}^l w_{sl} \sum_{t=s+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-s} \quad (3.2.8)$$

Donde $w_{sl} = 1 - s/(l+1)$

De manera similar, el contraste permite la inclusión de una tendencia determinista, modificando los estadísticos de prueba de la siguiente manera:

$$Z(\tilde{\alpha}) = T(\hat{\alpha} - 1) - \hat{\lambda}/M, \quad (3.2.9)$$

$$Z(t_{\tilde{\alpha}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\alpha}} - \tilde{\lambda}'\tilde{\sigma}_{Tl}/M^{1/2}, \quad (3.2.10)$$

$$Z(t_{\tilde{\mu}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\mu}} - \tilde{\lambda}'\tilde{\sigma}_{Tl}m_y/M^{1/2}(M + m_y^2)^{1/2}, \quad (3.2.11)$$

$$Z(t_{\tilde{\beta}}) = (\tilde{s}/\tilde{\sigma}_{Tl})t_{\tilde{\beta}} - \tilde{\lambda}'\tilde{\sigma}_{Tl} \left(\frac{1}{2}m_y - m_{ty} \right) / (M/12)^{1/2}\bar{m}_{yy}^{1/2} \quad (3.2.12)$$

Donde m_y , \bar{m}_{yy} , $\tilde{\lambda}$, $\tilde{\lambda}'$ y $\tilde{\sigma}_{Tl}$ son definidos al igual que en las ecuaciones 3.2.5 a 3.2.7 y $m_{ty} = T^{5/2} \sum t_{yt}$, $t_{\tilde{\mu}}$, $t_{\tilde{\beta}}$ y $t_{\tilde{\alpha}}$ son los estadísticos t de $\tilde{\mu}$, $\tilde{\alpha}$ y $\tilde{\beta}$, respectivamente. Por último la constante $M = (1 - T^{-2})m_{yy} - 12m_{ty}^2 + 12(1 + T^{-1})m_{ty}m_y - (4 + 6T^{-1} + 2T^{-2})m_y^2$

3.2.3. Contraste de Elliot, Rothenberg & Stock (1996)

Un defecto de los contrastes de raíz unitaria recién expuestos es su baja potencia si el verdadero proceso generador de datos es AR(1) cuyo coeficiente sea cercano a uno. Para mejorar la potencia de estos contrastes, Elliott, Rothenberg & Stock (1996) propusieron quitar los términos deterministas de la serie de tiempo. Los autores desarrollaron unos contrastes de punto-óptimo factible, denotados por P_T^μ y P_T^τ , los cuales toman en cuenta posibles problemas de autocorrelación en el término de error. El segundo contraste es denotado como el $DF - GLS$, el cual consiste en una modificación del contraste de Dickey-Fuller Aumentado. Se asume la siguiente forma para el proceso generador de los datos:

$$y_t = d_t + u_t \quad (3.2.13)$$

$$u_t = au_{t-1} + v_t \quad (3.2.14)$$

Donde $d_t = \beta' z_t$ representa a los componentes determinísticos, v_t es un proceso de error estacionario de media cero. En el caso en que $a = 1$, las ecuaciones 3.2.13 y 3.2.14 implican que el proceso es $I(1)$, mientras que si $|a| < 1$ significa que la serie es estacionaria.

Entonces el estadístico de punto óptimo factible será

$$P_T = \frac{S(a = \bar{a}) - \bar{a}S(a = 1)}{\hat{\omega}^2} \quad (3.2.15)$$

Donde $S(a = \bar{a})$ y $S(a = 1)$ son las sumas de cuadrados residuales de una regresión de mínimos cuadrados de y_a sobre Z_a con

$$y_a = (y_1, y_2 - ay_1, \dots, y_T - ay_{T-1}), \quad (3.2.16)$$

$$Z_a = (z_1, z_2 - az_1, \dots, z_T - az_{T-1}) \quad (3.2.17)$$

Por lo tanto, y_a es un vector columna $T - dimensional$ y Z_a defin una matriz de dimensiones $T \times q$. El estimador para la varianza del proceso de error v_t puede ser estimado como

$$\hat{\omega} = \frac{\hat{\sigma}_v^2}{(1 - \sum_{i=1}^p \hat{a}^i)^2} \quad (3.2.18)$$

Donde $\hat{\sigma}_v^2$ y \hat{a}_i para $i = 1, \dots, p$ son tomados desde la regresión de mínimos cuadrados auxiliar.

$$\Delta y_t = a_0 + a_1 \Delta y_{t-1} + \dots + \Delta y_{t-p} + a_{p+1} + v_t \quad (3.2.19)$$

Finalmente, la cantidad escalar \bar{a} es fijada como $\bar{a} = 1 + \frac{\bar{c}}{T}$, donde \bar{c} denota una constante. Dependiendo de los términos deterministas incluidos originalmente, el valor de \bar{c} será -7 para el caso de una constante, mientras que será de -13.5 en el caso de que se presente una tendencia lineal.

A continuación Elliot et al (1996) han propuesto otro contraste basado en el contraste de Dickey Fuller, el cuál es un estadístico para probar $\alpha_0 = 0$ basados en el siguiente

modelo.

$$\Delta y_t^d = \alpha_0 y_{t-1}^d + \alpha_1 \Delta_{t-1}^d + \dots + \alpha_p \Delta_{t-p}^d + \varepsilon_t \quad (3.2.20)$$

Donde y_t^d son los residuos en la regresión auxiliar $y_t^d \equiv y_t - \hat{\beta} z_t$.

3.2.4. Contraste Kiatkowsky, Pesaran, Schmidt & Shin (1992)

Este contraste está pensando para detectar ya sea estacionariedad en tendencia o en nivel. A diferencia de los contrastes anteriores donde la hipótesis nula implicaba afirmar la presencia de una raíz unitaria, en este caso, la hipótesis nula dice relación con que el proceso sea estacionario. Para construir el contraste se considera el siguiente modelo de base:

$$y_t = \zeta t + r_t + \varepsilon_t \quad (3.2.21)$$

$$r_t = r_{t-1} + u_t \quad (3.2.22)$$

Donde r_t es una caminata aleatoria y el término de error se asume i.i.d $(0, \sigma_u^2)$. La hipótesis nula de este contraste consiste en afirmar que $\mathcal{H}_0 : \sigma_u^2 = 0$, en cuyo caso, la tendencia r_t sería igualmente una tendencia determinista en lugar de estocástica.

Para construir el estadístico de prueba se sigue el siguiente procedimiento: i) se realiza una regresión de y_t sobre una constante o bien sobre una tendencia lineal y una constante, dependiendo de si se desea contrastar la hipótesis de estacionariedad en nivel o en pendiente. Luego se calcula la suma parcial de los residuos $\hat{\varepsilon}_t$ de la regresión como

$$S_t = \sum_{i=1}^t \hat{\varepsilon}_i, \quad t = 1, 2, \dots, T \quad (3.2.23)$$

El estadístico de prueba está definido como

$$LM = \frac{\sum_{t=1}^T S_t^2}{\hat{\sigma}_\varepsilon^2} \quad (3.2.24)$$

Donde $\hat{\sigma}_\varepsilon^2$ es una estimación de la varianza del error a un paso. Los autores sugieren utiliza una ventana de Bartlett $w(s, l) = 1 - s/(l + 1)$ como una función que da un peso optimo para calcular la varianza de largo plazo, esto es

$$\hat{\sigma}_\varepsilon^2 = s^2(l) = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t^2 + 2T - 1 \sum_{s=1}^l 1 - \frac{s}{l+1} \sum_{t=s+1}^T \hat{\varepsilon}_t \hat{\varepsilon}_{t-1} \quad (3.2.25)$$

3.3. Análisis de cointegración

3.3.1. Cointegración y análisis de las relaciones de largo plazo

Apuntes de Farías (2017a)

Dutoit et al. (2010, p. 15) define la transmisión en este contexto como “(...) la relación entre los precios de dos mercados relacionados; por ejemplo, entre el precio internacional de un producto y su precio doméstico

En econometría, la estimación de las relaciones entre variables que presentan tendencia reviste complejidad, porque su estructura puede provocar que se consideren significativas relaciones completamente espurias (Granger y Newbold, 1974)

Estas técnicas (cointegración) corresponden a los modelos de cointegración, que en la actualidad son utilizados con frecuencia en los estudios de transmisión de precios (análisis horizontal) y de transferencia de costos (análisis vertical)

En este caso, Π puede ser factorizado en una matriz $\alpha\beta'$, donde α es una matriz de dimensión $n \times r$ que representa la velocidad de ajuste al equilibrio, mientras que β es una matriz de dimensión $n \times r$ que representa los coeficientes de largo plazo

Apuntes de Juselius

- The time series describing cumulated trend-adjusted shocks is usually called a stochastic trend. It is a cumulation of random shocks with zero mean and constant variance. If
- with a linear deterministic trend component. Thus, the difference between a stochastic and deterministic trend is that the increments of a stochastic trend change randomly, whereas those of a deterministic trend are constant over time.

-
- It is easy to see that if inflation rate is $I(1)$ with a non-zero mean, then prices will contain a integrated twice cumulated of order stochastic two, or in trend, t $s=1$ notation $i=1$ s i . pt We say $I(2)$. that trend-adjusted prices are
 - We shall argue below that, unless a unit root is given a structural interpretation, the choice of one representation or the other is as such not very important, as long as there is consistency between the economic analysis and the choice. However, from an econometric point of view the choice between the two representations is usually crucial for the whole empirical analysis and should therefore be carefully considered.
 - variable. Because a cointegrating relation does not necessarily correspond to an interpretable economic relation, we make a further distinction between the statistical concept of a ‘cointegration relation’ and the economic concept of a ‘long-run equilibrium relation’.
 - say second that stochastic the distinction trend, between u_{2i} , as a long-run a long-run and structural medium-run trend stochastic or not. trend Thus,in one this might case is between an $I(1)$ stochastic trend with no linear trend and a near $I(1)$ stochastic trend with a linear trend.

Definición 3.3.1. Sea $\{\mathbf{x}_t\}$ un proceso estocástico para $t = \dots, -1, 0, 1, 2, \dots$ Si

$$\mathbb{E}[\mathbf{x}_t] = -\infty < \mu < \infty \quad (3.3.1)$$

$$\mathbb{E}[\mathbf{x}_t - \mu]^2 = \Sigma_0 < \infty \quad \forall t \quad (3.3.2)$$

$$\mathbb{E}[(\mathbf{x}_t - \mu)(\mathbf{x}_{t+h} - \mu)] = \Sigma_h \quad \forall t \text{ y } h \quad (3.3.3)$$

Entonces \mathbf{x}_t es debilmente estacionario. La estacionariedad estricta requiere que la distribución de (x_{t1}, \dots, x_{tk}) es la misma que $(x_{t1+h}, \dots, x_{tk+h})$ para $h = \dots, -1, 0, 1, 2, \dots$

for time t based on the available information at time t_1 . For example, a VAR model with autocorrelated and or heteroscedastic residuals would describe agents that do not use all information in the data as efficiently as possible. This is because by including the

For example, a VAR model with autocorrelated and or heteroscedastic residuals would describe agents that do not use all information in the data as efficiently as possible.

Simulation studies have shown that valid statistical inference is sensitive to violation of some of the assumptions, such as parameter non-constancy, autocorrelated residuals (the higher, the worse) and skewed residuals, while quite robust to others, such as excess kurtosis and residual heteroscedasticity. This will be discussed in more detail in

- the use of intervention dummies to account for significant political or institutional events during the sample;
- conditioning on weakly or strongly exogenous variables;
- checking the measurements of the chosen variables;
- changing the sample period to avoid fundamental regime shift or splitting the sample into more homogenous periods.

and the model has been extended to contain \mathbf{D}_t , a vector of deterministic components, such as a constant, seasonal dummies and intervention dummies. The autoregressive formulation is useful for expressing hypotheses on economic behaviour, whereas the moving average representation is useful when examining the properties of the process.

Si asumimos un modelo $VAR(2)$ bi-dimensional

$$(\mathbf{I} - \mathbf{\Pi}_1 L - \mathbf{\Pi}_2 L^2) \mathbf{x}_t = \mathbf{\Phi} \mathbf{D}_t + \varepsilon_t \quad (3.3.4)$$

La función características es entonces

$$\mathbf{\Pi}(z) = \mathbf{I} - \begin{bmatrix} \pi_{1,11} & \pi_{1,12} \\ \pi_{1,21} & \pi_{1,22} \end{bmatrix} z - \begin{bmatrix} \pi_{2,11} & \pi_{2,12} \\ \pi_{2,21} & \pi_{2,22} \end{bmatrix} z^2 \quad (3.3.5)$$

$$= \mathbf{I} - \begin{bmatrix} \pi_{1,11} z & \pi_{1,12} z \\ \pi_{1,21} z & \pi_{1,22} z \end{bmatrix} - \begin{bmatrix} \pi_{2,11} z^2 & \pi_{2,12} z^2 \\ \pi_{2,21} z^2 & \pi_{2,22} z^2 \end{bmatrix} \quad (3.3.6)$$

$$= \begin{bmatrix} (1 - \pi_{1,11} z - \pi_{2,11} z^2) & (-\pi_{1,12} z - \pi_{2,12} z^2) \\ (-\pi_{1,21} z - \pi_{2,21} z^2) & (1 - \pi_{1,22} z - \pi_{2,22} z^2) \end{bmatrix} \quad (3.3.7)$$

y

$$|\mathbf{\Pi}(z)| = (1 - \pi_{1,11}z - \pi_{2,11}z^2)(1 - \pi_{1,22}z - \pi_{2,22}z^2) - (\pi_{1,12}z + \pi_{2,12}z^2)(\pi_{1,21}z + \pi_{2,21}z^2) \quad (3.3.8)$$

$$= 1 - a_1z - a_2z^2 - a_3z^3 - a_4z^4 \quad (3.3.9)$$

$$= (1 - \rho_1z)(1 - \rho_2z)(1 - \rho_3z)(1 - \rho_4z) \quad (3.3.10)$$

El determinante entrega información valiosa sobre el comportamiento dinámico del proceso.

Luego

$$\mathbf{x}_t = \frac{\mathbf{\Pi}^a(L)(\mathbf{\Phi}\mathbf{D}_t + \varepsilon_t)}{(1 - \rho_1z)(1 - \rho_2z)(1 - \rho_3z)(1 - \rho_4z)} + \tilde{\mathbf{X}}^0, \quad t = 1, \dots, T \quad (3.3.11)$$

$$= \left(\frac{\mathbf{\Pi}_1^a L + \mathbf{\Pi}_2^a L^2}{(1 - \rho_2z)(1 - \rho_3z)(1 - \rho_4z)} \right) \left(\frac{\varepsilon_t + \mathbf{\Phi}D_t}{(1 - \rho_1L)} \right) + \tilde{\mathbf{X}}^0, \quad t = 1, \dots, T \quad (3.3.12)$$

3.4. Estimación basada en la verosimilitud para el modelo VAR irrestricto

Cuando el modelo no tiene restricciones sobre sus parámetros (como las que pueden surgir debido a la presencia de raíces unitarias) el modelo puede estimarse por MCO, caso que coincide con el estimador de *Full information maximum likelihood*

Si escribimos el modelo en su versión apilada

$$\mathbf{x}_t = \mathbf{B}'\mathbf{Z}_t + \varepsilon_t, \quad t = 1, \dots, T \quad (3.4.1)$$

$$\varepsilon_t \sim IN_p(\mathbf{0}, \mathbf{\Omega}) \quad (3.4.2)$$

Donde:

- $\mathbf{B}' = [\boldsymbol{\mu}_0, \mathbf{\Pi}_1, \mathbf{\Pi}_2, \dots, \mathbf{\Pi}_k]$

- $\mathbf{Z}'_t = [1, \mathbf{x}'_{t-1}, \mathbf{x}'_{t-2}, \dots, \mathbf{x}'_{t-k}]$

- $\mathbf{X}^0 = [\mathbf{x}'_0, \mathbf{x}'_{-1}, \dots, \mathbf{x}'_{-k+1}]$

La función de verosimilitud será la siguiente:

$$\log L(\mathbf{B}, \mathbf{\Omega}, \mathbf{X}) = -T \frac{p}{2} \log(2\pi) - T \frac{1}{2} \log |\mathbf{\Omega}| - \frac{1}{2} \sum_{t=1}^T (\mathbf{x}_t - \mathbf{B}' \mathbf{Z}_t)' \mathbf{\Omega}^{-1} (\mathbf{x}_t - \mathbf{B}' \mathbf{Z}_t) \quad (3.4.3)$$

Si calculamos $\frac{\partial \log L}{\partial \mathbf{B}}$, tendremos

$$\sum_{t=1}^T \mathbf{x}_t \mathbf{Z}_t' = \tilde{\mathbf{B}}' \sum_{t=1}^T \mathbf{Z}_t \mathbf{Z}_t'$$

Entonces, el estimador de máxima verosimilitud es

$$\tilde{\mathbf{B}}' = \sum_{t=1}^T (\mathbf{x}_t \mathbf{Z}_t') \left(\sum_{t=1}^T \mathbf{Z}_t \mathbf{Z}_t' \right)^{-1} = \mathbf{M}_{xZ} \mathbf{M}_{ZZ}^{-1} \quad (3.4.4)$$

Luego calculando $\frac{\partial \log L}{\partial \mathbf{\Omega}} = \mathbf{0}$

$$\hat{\mathbf{\Omega}} = T^{-1} \sum_{t=1}^T (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t)(\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t)' = T^{-1} \sum_{t=1}^T \hat{\mathbf{\epsilon}}_t \hat{\mathbf{\epsilon}}_t' \quad (3.4.5)$$

El valor máximo de la función de Verosimilitud, será el siguiente:

$$\log L_{\max} = -\frac{P}{2} T \log(2\pi) - \frac{1}{2} T \log |\hat{\mathbf{\Omega}}| - \frac{1}{2} \sum_{t=1}^T (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t)' \hat{\mathbf{\Omega}}^{-1} (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t) \quad (3.4.6)$$

Mostraremos que $\log L_{\max} = -\frac{1}{2} T \log |\hat{\mathbf{\Omega}}| + K$, $K \in \mathbb{R}$

$$\begin{aligned} (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t)' \hat{\mathbf{\Omega}}^{-1} (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{Z}_t) &= \hat{\mathbf{\epsilon}}_t' \hat{\mathbf{\Omega}}^{-1} \hat{\mathbf{\epsilon}}_t \\ &= \sum_{ij} \hat{\epsilon}_{t,i} (\hat{\mathbf{\Omega}}^{-1})_{ij} \hat{\epsilon}_{t,j} \end{aligned} \quad (3.4.7)$$

$$\begin{aligned} &= \sum_{ij} (\hat{\mathbf{\Omega}}^{-1})_{ij} \hat{\epsilon}_{t,i} \hat{\epsilon}_{t,j} \\ &= \text{traza} \{ \hat{\mathbf{\Omega}}^{-1} \hat{\mathbf{\epsilon}}_t \hat{\mathbf{\epsilon}}_t' \} \end{aligned} \quad (3.4.8)$$

Luego, se tiene que

$$\sum_{t=1}^T (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{z}_t) \hat{\mathbf{\Omega}}^{-1} (\mathbf{x}_t - \hat{\mathbf{B}}' \mathbf{z}_t)' = \sum_{t=1}^T \text{traza}\{\hat{\mathbf{\Omega}}^{-1} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t'\} \quad (3.4.9)$$

$$= T \sum_{t=1}^T \text{traza}\{\hat{\mathbf{\Omega}}^{-1} \hat{\boldsymbol{\varepsilon}}_t \hat{\boldsymbol{\varepsilon}}_t' / T\} \quad (3.4.10)$$

$$= T \text{traza}\{\hat{\mathbf{\Omega}}^{-1} \hat{\mathbf{\Omega}}\} \quad (3.4.11)$$

$$= T \text{traza}\{\mathbf{I}_p\} = Tp \quad (3.4.12)$$

De donde se desprende que

$$\log L_{\text{máx}} = -T \frac{1}{2} \log |\hat{\mathbf{\Omega}}| - \underbrace{T \frac{p}{2} - T \frac{p}{2} \log(2\pi)}_{+K} \quad (3.4.13)$$

NOTA PARA RECORDAR: si las variables del modelo están formuladas en logaritmo la desviación estándar de cada una de estas puede ser interpretada como un porcentaje de error

3.4.1. Contraste de razón de verosimilitud

$$-2 \log Q(\mathcal{H}_k / \mathcal{H}_{k+1}) = T(\log |\hat{\mathbf{\Omega}}_k| - \log |\hat{\mathbf{\Omega}}_{k+1}|) \sim \chi_{p^2}^2 \quad (3.4.14)$$

Criterios de selección

$$\text{AIC} = \log |\hat{\mathbf{\Omega}}| + (p^2 k) \frac{2}{T} \quad (3.4.15)$$

$$\text{SC} = \log |\hat{\mathbf{\Omega}}| + (p^2 k) \frac{\log T}{T} \quad (3.4.16)$$

$$\text{Hannah-Quinn} = \log |\hat{\mathbf{\Omega}}| + (p^2 k) \frac{2 \log \log T}{T} \quad (3.4.17)$$

Todos los criterios en común están basados en el máximo valor que alcanza la función de verosimilitud del modelo, más un factor que penaliza por el número de parámetros estimados.

Al momento de la determinación del número de rezagos, volver a revisar tabla 4.5 de la página 92

$$\text{Trace correlation} = 1 - \text{traza}(\hat{\Omega}[\text{Cov}(\Delta \mathbf{x}_t)]^{-1})/p \quad (3.4.18)$$

3.4.1.1. El contraste de Ljung-Box

$$\text{Ljung-Box} = T(T+2) \sum_{h=1}^{T/4} (T-h)^{-1} \text{traza}(\hat{\Omega}'_h \hat{\Omega}^{-1} \hat{\Omega}'_h \hat{\Omega}^{-1}) \quad (3.4.19)$$

Donde $\hat{\Omega}_h = T^{-1} \sum_{t=1}^T \hat{\varepsilon}_t \hat{\varepsilon}'_{t-h}$. El estadístico se considera distribuido aproximadamente según una χ^2 con $p^2(T/4 - k + 1) - p^2$ grados de libertad.

También puede utilizarse un contraste propuesto por Godfrey(1988), el cual consiste en regresar los residuos del modelo VAR estimado, $\hat{\varepsilon}_t$, sobre k variables rezagadas, $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{t-k}$ y el j -ésimo residuo rezagado

$$\hat{\varepsilon}_t = \mathbf{A}_1 \mathbf{x}_{t-1} + \mathbf{A}_2 \mathbf{x}_{t-2} + \dots + \mathbf{A}_k \mathbf{x}_{t-k} + \mathbf{A}_\varepsilon \hat{\varepsilon} \quad (3.4.20)$$

Donde los primeros j valores están perdidos $\hat{\varepsilon}_{-j}, \dots, \hat{\varepsilon}_{-1}$, los que son fijados a cero. El Estadístico de prueba, de tipo multiplicador de Lagrange es calculado de la siguiente forma

$$LM(j) = -(T - p(k+1) - \frac{1}{2}) \log \left(\frac{|\tilde{\Omega}(j)|}{\|\tilde{\Omega}\|} \right) \quad (3.4.21)$$

El estadístico se distribuye aproximadamente como una χ^2 con p^2 grados de libertad. Porque

3.4.1.2. Contrastes para Heteroscedasticidad Residual

El contraste ARCH m -ésimo es calculado como $(T+k-m) \times R^2$. Aquí R^2 se obtiene de la siguiente regresión auxiliar

$$\hat{\varepsilon}_{i,t}^2 = \gamma_0 + \sum_{j=1}^m \gamma_j \hat{\varepsilon}_{i,t-j}^2 + error \quad (3.4.22)$$

El estadístico se distribuye aproximadamente como una χ^2 con m grados de libertad.

3.4.1.3. Contrastes de normalidad

Para construir un contraste adecuado para verificar la hipótesis de normalidad (multivariada) de los residuos del modelo, se utiliza lo siguiente

$$\text{skewness}_i = \sqrt{\hat{b}_{1i}} = T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_i / \hat{\sigma}_i)_t^3 \quad (3.4.23)$$

$$\text{kurtosis}_i = \hat{b}_{2i} = T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_i / \hat{\sigma}_i)_t^4 \quad (3.4.24)$$

Bajo el supuesto de que los residuos se distribuyen normal, el skewness y la kurtosis de los residuos $\hat{\varepsilon}_i$ son asintóticamente normales con las siguientes medias y varianzas

$$\sqrt{T}(\text{skewness}_i - 0) \overset{a}{\sim} \mathcal{N}(0, 6) \quad (3.4.25)$$

y

$$\sqrt{T}(\text{kurtosis}_i - 3) \overset{a}{\sim} \mathcal{N}(0, 24) \quad (3.4.26)$$

Entonces, la varianza de skewness es más pequeña que la varianza de la kurtosis, lo cual significa que los contrastes de normalidad son más sensibles a desviaciones sobre el supuesto de skewness (a menudo como resultado de los outliers) que por el exceso de kurtosis (las colas pesadas o demasiados residuos cercanos a la media). Basado en lo anterior, se puede construir un contraste para normalidad univariada de la siguiente forma

3.4.2. Modelo Vectorial de Corrección del Error (VECM)

Suponga que cada componente de una serie de tiempo K -dimensional y_t es $I(1)$. Entonces, la ecuación (VAR) no será una formulación adecuada de este modelo debido a que los términos $y_t, y_{t-1}, \dots, y_{t-p}$ son todos no estacionarios. De todas formas, sustituyendo

$$\mathbf{A}_1 = \mathbf{I}_k + \mathbf{\Gamma}_1 \quad (3.4.27)$$

$$\mathbf{A}_i = \mathbf{\Gamma}_i - \mathbf{\Gamma}_{i-1} \quad i = 1, \dots, p-1 \quad (3.4.28)$$

$$\mathbf{A}_p = -\mathbf{\Gamma}_{p-1} \quad (3.4.29)$$

En la ecuación (2.8), reagrupando términos y utilizando que $\Delta \mathbf{y}_i = \mathbf{y}_i - \mathbf{y}_{i-1} \quad \forall i$, podemos reescribir esta ecuación como

$$\Delta \mathbf{y}_t = \mu + \mathbf{\Gamma}_1 \Delta \mathbf{y}_{t-1} + \mathbf{\Gamma}_2 \Delta \mathbf{y}_{t-2} + \dots + \mathbf{\Gamma}_{p-1} \Delta \mathbf{y}_{t-p+1} + \mathbf{u}_t \quad (3.4.30)$$

Naturalmente, ambas ecuaciones describen el mismo modelo, pero preferimos usar la ecuación (3.4.30) cuando \mathbf{y}_t es $I(1)$, debido a que cada término es estacionario en este caso. Entonces, cuando \mathbf{y}_t es $I(1)$, podemos encontrar un modelo apropiado para y_t diferenciando cada componente de \mathbf{y}_t una vez, y llevando a cabo la regresión basada en la ecuación (3.4.30). De todas formas, entonces no podremos tomar en cuenta que podría haber dependencias entre algunos de los componentes de \mathbf{y}_t . Por ejemplo, dos de los componentes podrían tener una tendencia en común, o podría existir una combinación lineal de los componente de y_t la cual fuera estacionaria. Este problema suele resolverse utilizando incluyendo un **término de corrección del error** $\mathbf{\Pi} \mathbf{y}_{t-1}$ en la ecuación (3.4.30), donde $\mathbf{\Pi}$ es una matriz $K \times K$ de cuyo rango $\text{rank}(\mathbf{\Pi}) < K$, debido a que si $\mathbf{\Pi}$ tuviera rango completo, entonces $\mathbf{\Pi}$ es invertible, de manera que la variable no estacionaria \mathbf{y}_{t-1} puede ser escrita como la suma de términos estacionarios, lo que es una contradicción. Entonces, $\text{rank}(\mathbf{\Pi}) = r < K$ lo cual implica que existen $(K \times r)$ – matrices $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ de rango r tales que $\mathbf{\Pi} = \boldsymbol{\alpha} \boldsymbol{\beta}'$. Entonces, cada una de las r filas de $\boldsymbol{\beta}' \mathbf{y}_{t-1}$ es una combinación lineal estacionaria de los componentes de y_t y es llamada una **relación de cointegración**. El número r , el cual es igual al número de relación de cointegración es llamado el **rango de cointegración**. Como la matriz $\boldsymbol{\beta}$ contiene todos los coeficiente de las relaciones de cointegración, es llamada **la matriz de cointegración**. La matriz $\boldsymbol{\alpha}$, la cual es la matriz de coeficientes de los términos estacionarios $\boldsymbol{\beta}' \mathbf{y}_{t-1}$ en la ecuación (3.4.31), es llamada la matriz de carga.

Definición 3.4.1. Un modelo **VECM** de orden p se define como

$$\Delta \mathbf{x}_t = \mathbf{\Gamma}_1 \Delta \mathbf{x}_{t-1} + \dots + \mathbf{\Gamma}_{p-1} \Delta \mathbf{x}_{t-p+1} + \boldsymbol{\alpha} \boldsymbol{\beta}' \mathbf{x}_{t-1} + \boldsymbol{\Phi} \mathbf{D}_t + \boldsymbol{\varepsilon}_t \quad t = 1, \dots, T \quad (3.4.31)$$

Donde $\mathbf{x}_t = [x_{1t}, \dots, x_{Kt}]'$ es un vector aleatorio de $k \times 1$, \mathbf{D}_t representa regresores determinísticos, $\boldsymbol{\alpha}$ y $\boldsymbol{\beta}$ son matrices $(K \times r)$ tales que $\text{rank}(\boldsymbol{\alpha}) = \text{rank}(\boldsymbol{\beta}) < K$,

3.4.3. Estimación de un modelo VECM

En esta sección se reproduce parcialmente lo expuesto por (Juselius, 2006). Considere un modelo $VAR(k)$ en su forma de corrección del error con $\Pi = \alpha\beta'$:

$$\Delta \mathbf{x}_t = \Gamma_1 \Delta \mathbf{x}_{t-1} + \dots + \Gamma_{k-1} \Delta \mathbf{x}_{t-k+1} + \alpha\beta' \mathbf{x}_{t-1} + \Phi \mathbf{D}_t + \varepsilon_t \quad (3.4.32)$$

Tal y como se definió previamente

Para escribir el modelo de manera más compacta se utiliza la siguiente notación

$$\begin{aligned} \mathbf{Z}_{0t} &= \Delta \mathbf{x}_t \\ \mathbf{Z}_{1t} &= \mathbf{x}_{t-1} \\ \mathbf{Z}_{2t} &= [\Delta \mathbf{x}'_{t-1}, \Delta \mathbf{x}'_{t-1}, \Delta \mathbf{x}'_{t-2}, \dots, \Delta \mathbf{x}'_{t-k+1}, \mathbf{D}'_t] \end{aligned}$$

De manera que el modelo VECM anteriormente descrito se puede escribir de manera más compacta

$$\mathbf{Z}_{0t} = \alpha\beta' \mathbf{Z}_{1t} + \Psi \mathbf{Z}_{2t} + \varepsilon_t \quad (3.4.33)$$

Si se utiliza el teorema de Frisch-Waugh podemos expresar la ecuación anterior de forma aún más simplificada utilizando lo siguiente:

Primero, se utilizan las siguientes regresiones auxiliares

$$\mathbf{Z}_{0t} = \hat{\mathbf{B}}'_1 \mathbf{Z}_{2t} + \mathbf{R}_{0t} \quad (3.4.34)$$

$$\mathbf{Z}_{1t} = \hat{\mathbf{B}}'_2 \mathbf{Z}_{2t} + \mathbf{R}_{1t} \quad (3.4.35)$$

Utilizando los residuos de estas regresiones auxiliares puede escribirse el modelo concentrado de la siguiente forma:

$$\mathbf{R}_{0t} = \alpha\beta' \mathbf{R}_{1t} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}_p(\mathbf{0}, \Omega) \quad (3.4.36)$$

Para estimar el modelo se utiliza un procedimiento de dos pasos: Primero, se asume que β es conocido y se obtiene un estimador de α bajo el supuesto que $\beta' \mathbf{R}_{1t}$ es una variable conocida. Entonces se inserta $\alpha = \hat{\alpha}(\beta)$ en la expresión de la función de

verosimilitud, de forma que ésta esté en términos de β , pero no de α . Hecho esto, pueden encontrarse los valores de $\hat{\beta}$ que maximizan la función de verosimilitud. Luego, se puede proceder a estimar $\hat{\alpha} = \alpha(\hat{\beta})$

El primer paso consiste en utilizar la expresión (3.4.36) sin el término de error y post-multiplicar ésta por $\mathbf{R}'_{1t}\beta$

$$\mathbf{R}_{0t}\mathbf{R}'_{1t}\beta = \alpha\beta'\mathbf{R}_{1t}\mathbf{R}'_{1t}\beta \quad (3.4.37)$$

Sumando sobre t y dividiendo por T , se tiene:

$$\underbrace{T^{-1} \sum_t \mathbf{R}_{0t}\mathbf{R}'_{1t}}_{\mathbf{S}_{01}} \beta = \alpha \underbrace{\beta' T^{-1} \sum_t \mathbf{R}_{1t}\mathbf{R}'_{1t}}_{\mathbf{S}_{11}} \beta \quad (3.4.38)$$

Un estimador de mínimos cuadrados ordinarios de esta expresión puede obtenerse *despejando* la expresión anterior:

$$\hat{\alpha}(\beta) = \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1} \quad (3.4.39)$$

El segundo paso consiste en utilizar el método de máxima verosimilitud asumiendo que el término de error sigue una distribución normal multivariada. Donde para maximizar la función es necesario operar sobre el determinante de la matriz de varianza-covarianza. Es decir:

$$\mathcal{L}^{-2/T}(\beta, \alpha) = |\hat{\Omega}(\beta, \alpha)| + \text{términos constantes} \quad (3.4.40)$$

Donde:

$$\hat{\Omega}(\beta, \alpha) = T^{-1} \sum (\mathbf{R}_{0t} - \alpha\beta'\mathbf{R}_{1t})(\mathbf{R}_{0t} - \alpha\beta'\mathbf{R}_{1t})' \quad (3.4.41)$$

$$= T^{-1} \left(\sum \mathbf{R}_{0t}\mathbf{R}_{0t}' - \sum \mathbf{R}_{0t}\mathbf{R}_{1t}'\alpha\beta' - \alpha\beta' \sum \mathbf{R}_{1t}\mathbf{R}_{0t}' + \alpha\beta' \sum \mathbf{R}_{1t}\mathbf{R}_{1t}'\alpha\beta' \right) \quad (3.4.42)$$

$$= \mathbf{S}_{00} - \mathbf{S}_{01}\alpha\beta' - \alpha\beta'\mathbf{S}_{10} + \alpha\beta'\mathbf{S}_{11}\alpha\beta' \quad (3.4.43)$$

Si se sustituye (3.4.39) en (3.4.41), puede expresarse la matriz de covarianza como

una función exclusiva de β

$$\hat{\Omega}(\beta) = \mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10} \quad (3.4.44)$$

Sustituyendo (3.4.40) en (3.4.44), puede expresarse la matriz de covarianza sólo en términos de β

$$\hat{\Omega}(\beta) = \mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10} \quad (3.4.45)$$

$$+ \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1} \underbrace{\beta'\mathbf{S}_{11}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}}_{\mathbf{I}} \mathbf{S}_{10} \quad (3.4.46)$$

De esta manera,

$$\hat{\Omega}(\beta) = \mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10} \quad (3.4.47)$$

Para encontrar el estimador de máxima verosimilitud de β que minimiza $|\hat{\Omega}(\hat{\beta})|$ puede utilizarse el siguiente resultado:

$$\left| \begin{array}{cc} \mathbf{A} & \mathbf{B} \\ \mathbf{B}' & \mathbf{C} \end{array} \right| = |\mathbf{A}| \cdot |\mathbf{A} - \mathbf{B}\mathbf{C}^{-1}\mathbf{B}'| \quad (3.4.48)$$

Donde \mathbf{A} y \mathbf{C} son matrices cuadradas no singulares. Ahora se realiza la siguiente sustitución:

$$\mathbf{S}_{00} = \mathbf{A} \quad (3.4.49)$$

$$\beta'\mathbf{S}_{11}\beta = \mathbf{C} \quad (3.4.50)$$

$$\mathbf{S}_{01}\beta = \mathbf{B} \quad (3.4.51)$$

De manera que la expresión anterior puede reescribirse como:

$$|\mathbf{S}_{00}| \cdot |\beta'\mathbf{S}_{11}\beta - \beta'\mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\beta| = |\beta'\mathbf{S}_{11}\beta| \cdot \underbrace{|\mathbf{S}_{00} - \mathbf{S}_{01}\beta(\beta'\mathbf{S}_{11}\beta)^{-1}\beta'\mathbf{S}_{10}|}_{|\hat{\Omega}(\beta)|} \quad (3.4.52)$$

De donde se deduce que:

$$|\hat{\Omega}(\beta)| = \frac{|\mathbf{S}_{00}| \cdot |\beta' \mathbf{S}_{11} \beta - \beta' \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01} \beta|}{|\beta' \mathbf{S}_{11} \beta|} \quad (3.4.53)$$

$$|\hat{\Omega}(\beta)| = |\mathbf{S}_{00}| \cdot \frac{|\beta' \mathbf{S}_{11} \beta - \beta' \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01} \beta|}{|\beta' \mathbf{S}_{11} \beta|} \quad (3.4.54)$$

El segundo término del lado derecho de la ecuación se conoce como *cociente de Rayleigh*, por lo que los valores de la matriz β , pueden encontrarse diagonalizando dicha forma cuadrática.

Los autovalores asociados pueden encontrarse como:

$$|\rho \mathbf{S}_{11} - \mathbf{S}_{11} + \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01}| = 0 \quad (3.4.55)$$

O equivalentemente,

$$|\underbrace{(1 - \rho)}_{\lambda} \mathbf{S}_{11} - \mathbf{S}_{10} \mathbf{S}_{00}^{-1} \mathbf{S}_{01}| = 0 \quad (3.4.56)$$

Luego el determinante de la matriz de covarianza puede expresarse como:

$$|\hat{\Omega}(\beta)| = |\mathbf{S}_{00}| \prod_{i=1}^p (1 - \lambda_i) \quad (3.4.57)$$

Luego, los autovectores asociados a cada autovalor $\mathbf{v}_1, \dots, \mathbf{v}_p$ se corresponderán con los vectores columna de la matriz $\hat{\beta}$, ordenados de manera que $\hat{\lambda}_1 > \dots > \hat{\lambda}_p$.

3.4.4. El contraste de cointegración de Johansen

Utilizando las expresiones (3.4.40) y (3.4.57), la función de logverosimilitud del modelo puede ser representada por:

$$-2 \log \mathcal{L}(\beta) = T \log |\mathbf{S}_{00}| + T \sum_{i=1}^p \log(1 - \lambda_i) \quad (3.4.58)$$

Cabe notar entonces que, como se señaló en la sección anterior, mientras mayor la magnitud de los autovalores, más *estacionaria* será la relación de cointegración que

representa. De esta manera, el contraste busca identificar qué autovalores - y por tanto vectores cointegrantes - llevan a relaciones estacionarias. De esta manera, el contraste de hipótesis puede realizarse de manera secuencial.

$$\mathcal{H}(p) : rk(\mathbf{\Pi}) = p \quad vs \quad \mathcal{H}(r) : rk(\mathbf{\Pi}) = r \quad (3.4.59)$$

Esto quiere decir que para el primer caso, se contrasta la hipótesis de que no existen raíces unitarias comunes entre las diferentes variables endógenas del modelo y por tanto, puede estimarse un modelo VAR en niveles. Mientras que en los contrastes posteriores se pone a prueba la hipótesis sobre el número de relaciones de cointegración que pueden considerarse estacionarias.

Para estos efectos se puede construir un contraste de razón de verosimilitud (LR rest) de la siguiente forma:

$$-2 \log |\mathcal{Q}(\mathcal{H}_r/\mathcal{H}_p)| = T \log \left\{ \frac{|\mathbf{S}_{00}|(1 - \hat{\lambda}_1) \dots (1 - \hat{\lambda}_r)}{|\mathbf{S}_{00}|(1 - \hat{\lambda}_1) \dots (1 - \hat{\lambda}_p)} \right\} \quad (3.4.60)$$

$$\tau_{p-r} = -T \log \left[(1 - \hat{\lambda}_{r+1}) \dots (1 - \hat{\lambda}_p) \right] \quad (3.4.61)$$

La magnitud de este estadístico proporcionará evidencia en torno a la validez o no de las hipótesis a contrastar de acuerdo a si sobrepasa o no algún valor crítico C_{p-r*} . Sin embargo, de manera similar a como ocurre con el contraste de raíz unitaria de Dickey-Fuller, la distribución de este estadístico no es estándar. Sin embargo Johansen et al. (1996) proveen los valores críticos del contraste bajo diferentes escenarios.

3.4.5. Estabilidad del modelo VECM

NOTA PARA RECORDAR: UNA PRIMERA ESTIMACIÓN PARA EL VECTOR COINTEGRANTE ES LA ESTIMACIÓN en dos etapas de DE ENGLE GRANGER

CITAR "FIVE ALTERNATIVE METHODS OF ESTIMATING LONG-RUN EQUILIBRIUM RELATIONSHIPS" PARA MOSTRAR QUE EL MODELO NO ES SENSIBLE A LAS DESVIACIONES DEL SUPUESTO DE NORMALIDAD. (GONZALO JEC1994)

3.5. Modelo de vector de corrección del error por umbrales (TVECM)

3.5.1. Estimación del modelo TVECM

En esta sección se reproduce lo expuesto en el artículo de Hansen & Seo (2002) y Lo & Zivot (2001). En esta sección puede reescribirse la formulación del modelo (3.4.31) apilando los diferentes elementos, de la siguiente forma:

$$\Delta \mathbf{x}_t = \mathbf{A}' \mathbf{X}_{t-1}(\boldsymbol{\beta}) + \mathbf{u}_t \quad (3.5.1)$$

Donde

$$\mathbf{x}_{t-1}(\boldsymbol{\beta}) = \begin{pmatrix} w_{t-1}(\boldsymbol{\beta}) \\ \Delta \mathbf{x}_{t-1} \\ \Delta \mathbf{x}_{t-1} \\ \vdots \\ \Delta \mathbf{x}_{t-1} \\ \mathbf{D}_t \end{pmatrix} \quad (3.5.2)$$

Donde $w_t(\boldsymbol{\beta}) = \boldsymbol{\beta}' \mathbf{x}_{t-1}$. La regresora $\mathbf{x}_{t-1}(\boldsymbol{\beta})$ tiene dimensión $k \times 1$ y \mathbf{A} tiene dimensión $k \times p$ donde $k = pl + 2$. El término de error $\boldsymbol{\varepsilon}_t$ se distribuye de la misma manera asumida anteriormente, es decir, tiene covarianza finita, $\boldsymbol{\Omega} = \mathbb{E}(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$.

La notación $w_{t-1}(\boldsymbol{\beta})$ y $\mathbf{x}_{t-1}(\boldsymbol{\beta})$ indica que las variables son evaluadas en valores genéricos de $\boldsymbol{\beta}$. Cuando sean evaluadas en los verdaderos valores del vector cointegrante, se denota como w_{t-1} y \mathbf{x}_{t-1} , respectivamente¹.

Como una extensión del modelo anterior, puede escribirse el modelo TVECM de la siguiente forma:

Definición 3.5.1.

$$\Delta \mathbf{x}_t = \begin{cases} \mathbf{A}'_1 \mathbf{X}_{t-1}(\boldsymbol{\beta}) + \boldsymbol{\varepsilon}_t & \text{si } w_{t-1} \boldsymbol{\beta} \leq \gamma \\ \mathbf{A}'_2 \mathbf{X}_{t-1}(\boldsymbol{\beta}) + \boldsymbol{\varepsilon}_t & \text{si } w_{t-1} \boldsymbol{\beta} > \gamma \end{cases} \quad (3.5.3)$$

¹Es importante notar que para este modelo, no se utiliza la versión concentrada, pues como se verá más adelante, los coeficientes que describen la dinámica de corto plazo del proceso también dependerán del régimen escogido

Donde γ es el parámetro que determina el umbral a partir del cual ocurre el cambio de estructura del modelo.

El modelo anterior puede representarse de manera más sucinta utilizando variables indicadoras:

$$\Delta \mathbf{x}_t = \mathbf{A}'_1 \mathbf{X}_{t-1}(\boldsymbol{\beta}) d_{1t}(\boldsymbol{\beta}, \gamma) + \mathbf{A}'_2 \mathbf{X}_{t-1}(\boldsymbol{\beta}) d_{2t}(\boldsymbol{\beta}, \gamma) + \varepsilon_t \quad (3.5.4)$$

Donde

$$d_{1t} = 1(w_{t-1}(\boldsymbol{\beta}) \leq \gamma) \quad (3.5.5)$$

$$d_{2t} = 1(w_{t-1}(\boldsymbol{\beta}) > \gamma) \quad (3.5.6)$$

El efecto del umbral sólo tiene sentido si $0 < \mathbb{P}(w_{t-1} \leq \gamma) < 1$. Lo que se puede representar a través de la siguiente restricción:

$$\pi_0 \leq \mathbb{P}(w_{t-1} \leq \gamma) \leq 1 - \pi_0 \quad (3.5.7)$$

Donde $\pi_0 > 0$ es un parámetro que restringe el espacio del parámetro γ , para las aplicaciones dicho parámetro suele fijarse en $\pi_0 = 0,05$. Luego el modelo puede estimarse por el método de máxima verosimilitud, siendo la función a maximizar la siguiente:

$$\mathcal{L}(\mathbf{A}_1, \mathbf{A}_2, \boldsymbol{\Omega}, \boldsymbol{\beta}, \gamma) = -\frac{T}{2} \log |\hat{\boldsymbol{\Omega}}| - \frac{1}{2} \sum_{t=1}^T \varepsilon_t(\mathbf{A}_1, \mathbf{A}_2, \boldsymbol{\Omega}, \boldsymbol{\beta}, \gamma)' \boldsymbol{\Omega}^{-1} \varepsilon_t(\mathbf{A}_1, \mathbf{A}_2, \boldsymbol{\Omega}, \boldsymbol{\beta}, \gamma) \quad (3.5.8)$$

Para efectos prácticos, resulta útil resolver la ecuación anterior para los estimadores de la dinámica de corto plazo como funciones de los parámetros $(\boldsymbol{\beta}, \gamma)$. De manera que las estimaciones OLS son las siguientes:

$$\hat{\mathbf{A}}_1(\boldsymbol{\beta}, 1) = \left(\sum_{t=1}^T \mathbf{X}_{t-1}(\boldsymbol{\beta}) \mathbf{X}_{t-1}(\boldsymbol{\beta})' d_{1t}(\boldsymbol{\beta}, \gamma) \right)^{-1} \left(\sum_{t=1}^T \mathbf{X}_{t-1}(\boldsymbol{\beta}) \Delta \mathbf{x}_{t-1}(\boldsymbol{\beta})' d_{1t}(\boldsymbol{\beta}, \gamma) \right) \quad (3.5.9)$$

$$\hat{\mathbf{A}}_2(\boldsymbol{\beta}, 1) = \left(\sum_{t=1}^T \mathbf{X}_{t-1}(\boldsymbol{\beta}) \mathbf{X}_{t-1}(\boldsymbol{\beta})' d_{2t}(\boldsymbol{\beta}, \gamma) \right)^{-1} \left(\sum_{t=1}^T \mathbf{X}_{t-1}(\boldsymbol{\beta}) \Delta \mathbf{x}_{t-1}(\boldsymbol{\beta})' d_{2t}(\boldsymbol{\beta}, \gamma) \right) \quad (3.5.10)$$

$$\hat{\boldsymbol{\varepsilon}}_t(\boldsymbol{\beta}, \gamma) = \boldsymbol{\varepsilon}_t(\hat{\mathbf{A}}_1(\boldsymbol{\beta}, \gamma), \hat{\mathbf{A}}_2(\boldsymbol{\beta}, \gamma), \boldsymbol{\beta}, \gamma) \quad (3.5.11)$$

$$\hat{\boldsymbol{\Omega}}(\boldsymbol{\beta}, \gamma) = \frac{1}{T} \sum_{t=1}^T \hat{\boldsymbol{\varepsilon}}_t(\boldsymbol{\beta}, \gamma) \hat{\boldsymbol{\varepsilon}}_t(\boldsymbol{\beta}, \gamma)' \quad (3.5.12)$$

Luego los estimadores de máxima verosimilitud son $\hat{\mathbf{A}}_1 = \hat{\mathbf{A}}_1(\hat{\boldsymbol{\beta}}, \hat{\gamma})$ y $\hat{\mathbf{A}}_2 = \hat{\mathbf{A}}_2(\hat{\boldsymbol{\beta}}, \hat{\gamma})$. Este procedimiento no da pie a una solución analítica para encontrar los parámetros $(\boldsymbol{\beta}, \gamma)$, por lo que se suele utilizar una *gridsearch*, que puede entenderse como una búsqueda exhaustiva de los parámetros a través de tanteo. Esto implica que la búsqueda sólo ha de realizarse sobre la base de un modelo con dos variables $k = 2$, y teniendo en cuenta la normalización para la identificación del vector cointegrante, se tiene que sólo se pueden buscar dos parámetros: $\boldsymbol{\beta} = (1, \beta)$ y γ . De utilizarse más variables endógenas el proceso de búsqueda se vuelve demasiado costoso computacionalmente.

Aún así, para disminuir el tiempo de búsqueda se suelen exponer diferentes restricciones. Por ejemplo, se sugiere utilizar como punto de referencia para β la estimación del modelo lineal y para el caso de γ utilizar el soporte empírico de la relación de cointegración obtenida a través de la estimación del modelo **VECM**.

El proceso de estimación puede resumirse en los siguientes pasos:

1. Determinar los posibles valores de los parámetros estableciendo un intervalo de búsqueda dado por $[\gamma_L, \gamma_U]$ y $[\beta_L, \beta_U]$ basándose en la estimación lineal $\tilde{\beta}$ y el soporte de la relación de cointegración.
2. Para cada par (β, γ) sobre el espacio de búsqueda, evaluar $\hat{\mathbf{A}}_1(\beta, \gamma)$, $\hat{\mathbf{A}}_2(\beta, \gamma)$, y $\hat{\boldsymbol{\Omega}}(\beta, \gamma)$
3. Encontrar los estimadores $(\hat{\beta}, \hat{\gamma})$ como los valores de (β, γ) sobre el espacio de búsqueda que lleva al valor más pequeño de la expresión $\log |\hat{\boldsymbol{\Omega}}(\beta, \gamma)|$
4. Fijar $\hat{\boldsymbol{\Omega}} = \hat{\boldsymbol{\Omega}}(\hat{\beta}, \hat{\gamma})$

3.5.2. El contraste de Hansen & Seo (2002)

Sea \mathcal{H}_0 la clase de modelos lineales **VECM** y \mathcal{H}_1 la clase de modelos de umbrales de dos regímenes². Cabe notar que ambos modelos están anidados, por lo que un modelo

²Nótese que en esta sección se utiliza el modelo bivariado, razón por lo cual se deja de utilizar notación $\boldsymbol{\beta}$ que denota una matriz y se utiliza solamente β , que denota un solo parámetro, pues

lineal será un modelo de dos regimenes en que se cumpla que $\mathbf{A}_1 = \mathbf{A}_2$. El objeto de esta sección es entonces determinar como contrastar la evidencia estadística a favor de cointegración lineal o bien de cointegración por umbrales.

En primera instancia se asume que los parámetros (β, γ) son fijos y conocidos. El modelo bajo \mathcal{H}_0 es el siguiente:

$$\Delta \mathbf{x}_t = \mathbf{A}' \mathbf{X}_{t-1}(\beta) + \boldsymbol{\varepsilon}_t \quad (3.5.13)$$

y bajo \mathcal{H}_1

$$\Delta \mathbf{x}_t = \mathbf{A}'_1 \mathbf{X}_{t-1}(\beta) d_{1t}(\beta, \gamma) + \mathbf{A}'_2 \mathbf{X}_{t-1}(\beta) d_{2t}(\beta, \gamma) + \boldsymbol{\varepsilon}_t \quad (3.5.14)$$

Al asumir ambos parámetros de interés conocidos, el modelo se vuelve lineal, por lo que el estimador de máxima verosimilitud coincide con el de mínimos cuadrados. Siguiendo este mismo razonamiento, se puede utilizar un contraste de tipo multiplicador de lagrange (LM) que sea robusto a problemas de heteroscedasticidad. Para estos fines se cambia ligeramenta la notación. Sean $\mathbf{X}_1(\beta, \gamma)$ y $\mathbf{X}_2(\beta, \gamma)$ matrices formadas por los vectores apilados de $\mathbf{X}_{t-1}(\beta) d_{1t}(\beta, \gamma)$ y $\mathbf{X}_{t-1}(\beta) d_{2t}(\beta, \gamma)$, respectivamente. Sean además $\boldsymbol{\xi}_1(\beta, \gamma)$ y $\boldsymbol{\xi}_2(\beta, \gamma)$ matrices de las filas apiladas $\tilde{\boldsymbol{\varepsilon}}_t \otimes \mathbf{X}_{t-1}(\beta) d_{1t}(\beta, \gamma)$ y $\tilde{\boldsymbol{\varepsilon}}_t \otimes \mathbf{X}_{t-1}(\beta) d_{2t}(\beta, \gamma)$, respectivamente. Con $\tilde{\boldsymbol{\varepsilon}}_t$ representando los residuos de un modelo **VECM** lineal. Luego se define el producto externo:

$$\mathbf{M}_1(\beta, \gamma) = \mathbf{I}_p \otimes \mathbf{X}_1(\beta, \gamma)' \mathbf{X}_1(\beta, \gamma) \quad (3.5.15)$$

$$\mathbf{M}_2(\beta, \gamma) = \mathbf{I}_p \otimes \mathbf{X}_2(\beta, \gamma)' \mathbf{X}_2(\beta, \gamma) \quad (3.5.16)$$

y

$$\boldsymbol{\Xi}_1(\beta, \gamma) = \boldsymbol{\xi}_1(\beta, \gamma)' \boldsymbol{\xi}_1(\beta, \gamma) \quad (3.5.17)$$

$$\boldsymbol{\Xi}_2(\beta, \gamma) = \boldsymbol{\xi}_2(\beta, \gamma)' \boldsymbol{\xi}_2(\beta, \gamma) \quad (3.5.18)$$

$$(3.5.19)$$

Acto seguido se puede definir las matrices $\hat{\mathbf{V}}_1(\beta, \gamma)$ y $\hat{\mathbf{V}}_2(\beta, \gamma)$, como estimadores de la varianza al *textit* Eicker-White

se asume además que el vector cointegrante ha sido normalizado

$$\hat{\mathbf{V}}_1(\beta, \gamma) = \mathbf{M}_1(\beta, \gamma)^{-1} \Xi_1(\beta, \gamma) \mathbf{M}_1(\beta, \gamma)^{-1} \quad (3.5.20)$$

$$\hat{\mathbf{V}}_2(\beta, \gamma) = \mathbf{M}_1(\beta, \gamma)^{-1} \Xi_1(\beta, \gamma) \mathbf{M}_1(\beta, \gamma)^{-1} \quad (3.5.21)$$

Lo que lleva a un estimador LM resistente a problemas de heteroscedasticidad:

$$LM(\beta, \gamma) = \text{vec}(\hat{\mathbf{A}}_1(\beta, \gamma) - \hat{\mathbf{A}}_1(\beta, \gamma))' (\hat{\mathbf{V}}_1(\beta, \gamma) + \hat{\mathbf{V}}_2(\beta, \gamma))^{-1} \text{vec}(\hat{\mathbf{A}}_1(\beta, \gamma) - \hat{\mathbf{A}}_1(\beta, \gamma)) \quad (3.5.22)$$

Si β y γ fueran conocidos, éste sería el estadístico. Pero, dado que no lo son, el estadístico debe ser evaluado bajo \mathcal{H}_0 . El estimador de β ha usarse será el proveniente del modelo lineal $\tilde{\beta}$, pero no existe un estimador de γ bajo \mathcal{H}_0 , por lo que, como señalan Hansen & Seo (2012), se utiliza el criterio de unión intersección que propone utilizar como estadístico

$$SupLM = \sup_{\gamma_L \leq \gamma < \gamma_U} LM(\tilde{\beta}, \gamma) \quad (3.5.23)$$

3.5.3. Función de impulso respuesta no lineal

Mencionar que para obtener los intervalos de confianza en este contexto se deben utilizar métodos de bootstrapping y citar a Kreiss & Lahiri (2012) y también a Koop, Pesaran & Potter (1996)

3.6. Imputación de datos faltantes a través del filtro de Kalman

CITAR A Durbin y Koopman (2012)

3.6.1. El filtro de Kalman

3.6.2. Formulación en espacio estado de un modelo ARI-MA(p,d,q)

3.6.3. imputación de valores faltantes

Capítulo 4

Metodología y Datos

4.1. Fuentes de información

Capítulo 5

Resultados

En este capítulo se muestra una aplicación

Todos los cálculos aquí expuestos fueron realizados a través del lenguaje y entorno de programación orientado a la estadística, **R** (R core team, 2017).

5.1. Imputación de valores perdidos

Dada la frecuencia relativamente alta de los datos, se tiene un número relativamente importante de observaciones faltantes. Para solucionar este problema se procedió a estimar un modelo ARIMA a través comando `auto.arima` del paquete `forecast` de Hyndman & Khandakar (2007).

A continuación se presentan los resultados de la estimación para ambos modelos:

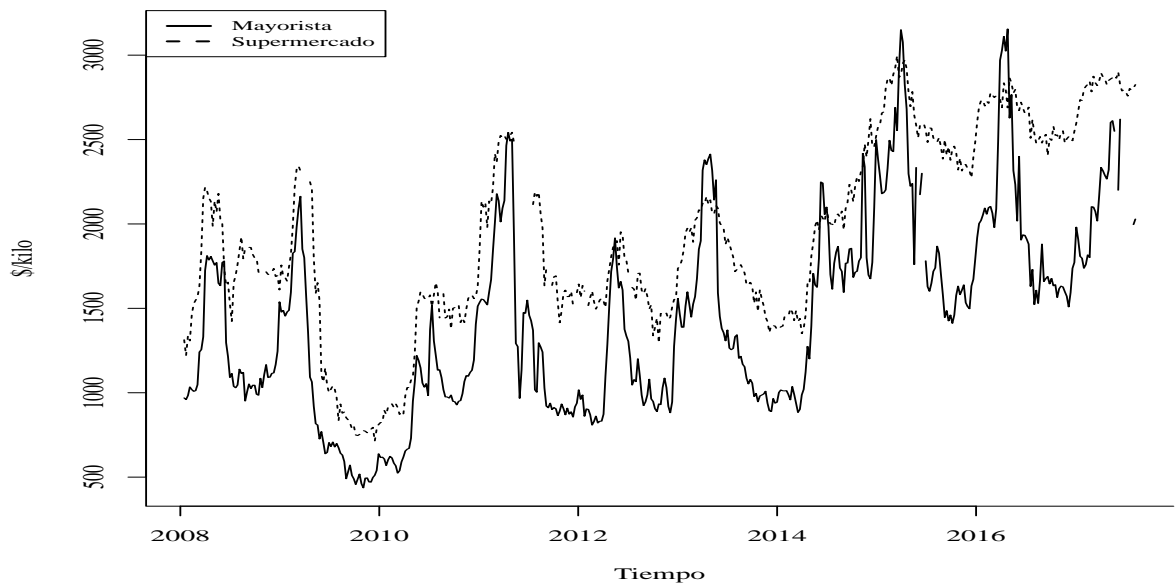


Figura 5.1: Evolución de precios del palta Hass de primera calidad, 2008-2016

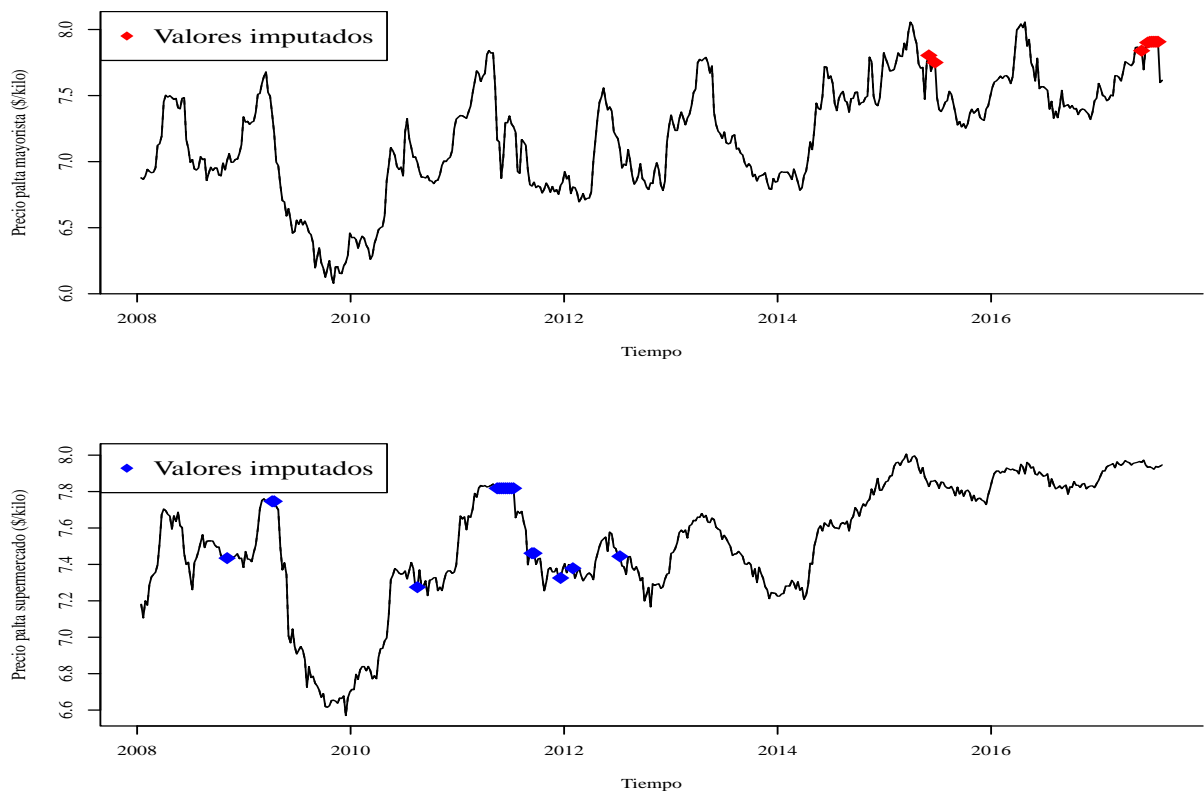


Figura 5.2: Imputación de valores perdidos a través del filtro de Kalman

5.2. Análisis del orden de integración de las series

5.2.1. Análisis gráfico

Para analizar el orden de integración de las series es de utilidad usar herramientas gráficas que entreguen cierta orientación sobre el comportamiento de las series antes de realizar un análisis formal a través de contrastes de hipótesis.

Cabe señalar que, como un intento de estabilizar la varianza de las series, estas serán trabajadas en logaritmo.

5.2.1.1. Análisis de las series en nivel

En el caso del logaritmo de los precios mayoristas de la palta, puede apreciarse una media no constante, aunque lo mismo no es directamente apreciable para la varianza. Por otro lado, su correlograma muestra autocorrelaciones significativas más allá de 52 períodos de rezago, lo que permite afirmar que el proceso tendría una memoria prolongada y que por tanto, no es estacionario (Véase figura 5.3).

Para el caso de la serie del logaritmo del precio de la palta en supermercados, se puede observar un comportamiento bastante similar. Es decir, una media que no es constante debido a una tendencia, cuya naturaleza se determina más adelante; y una memoria prolongada al observar el correlograma correspondiente (Véase figura 5.4).

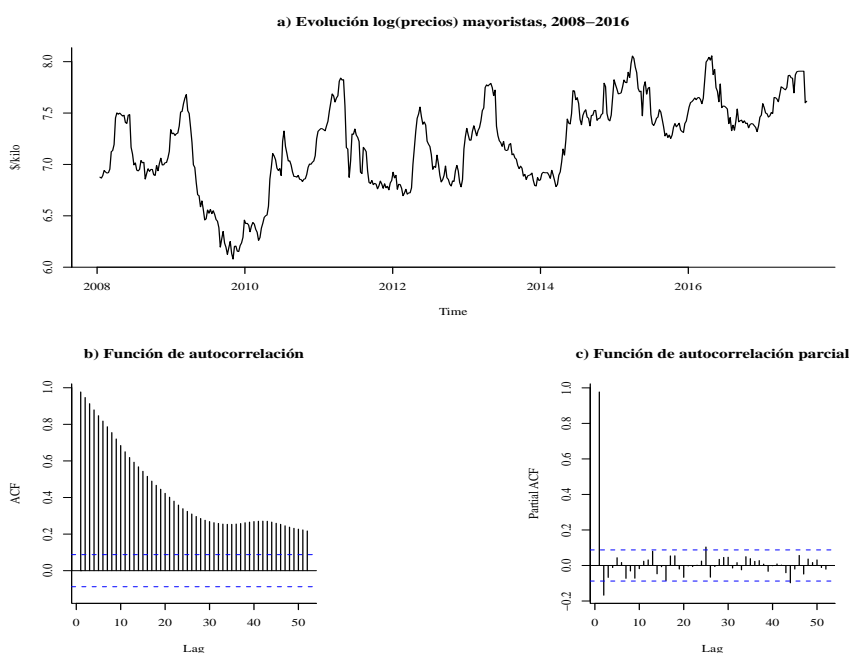


Figura 5.3: Evolución del logaritmo del precio mayorista de la palta, 2008-2016

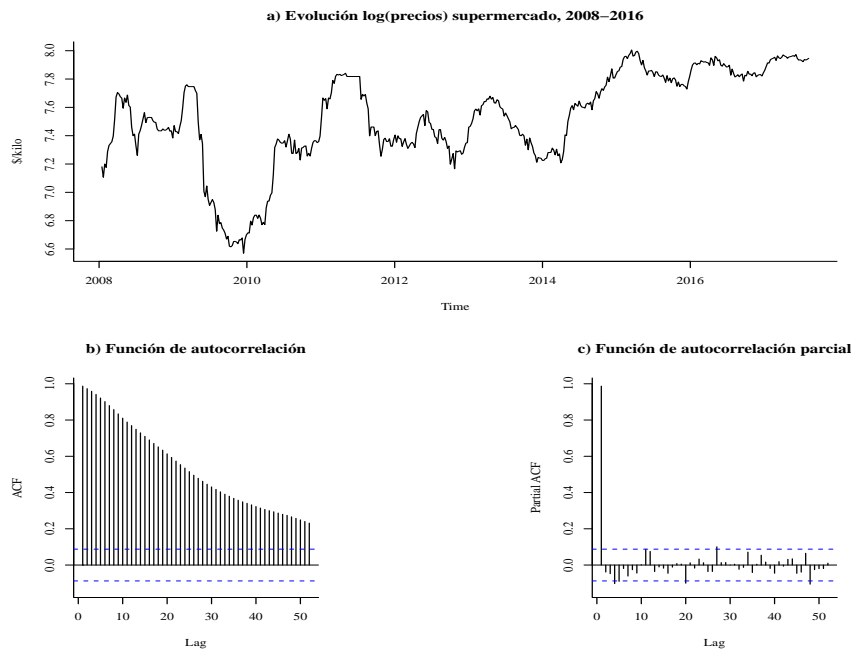


Figura 5.4: Evolución del logaritmo del precio en supermercado de la palta, 2008-2016

5.2.1.2. Análisis de las series en diferencias

A continuación se procedió a diferenciar las series, pues existe la sospecha de la presencia de raíz unitaria en ambas, cuestión que será determinada más adelante.

Al analizar las figuras 5.3 y 5.4 se observan algunos rezagos significativos en el correlograma, lo que puede servir de orientación para determinar el orden de integración de las series. A pesar de lo anterior, se observa de todas maneras que el comportamiento de la varianza no es estable a lo largo del tiempo en ambos casos.

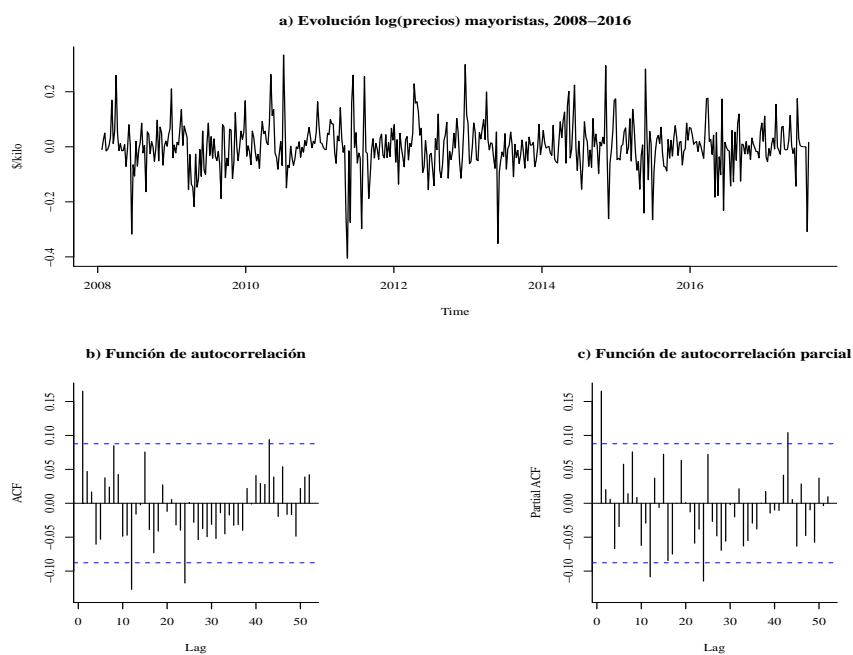


Figura 5.5: Evolución del logaritmo del precio mayorista de la palta, 2008-2016

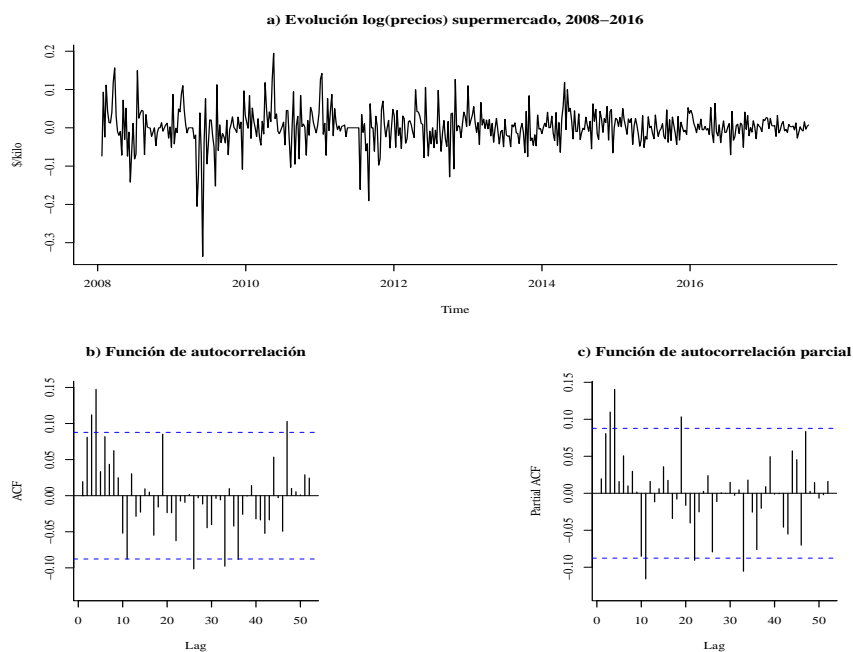


Figura 5.6: Evolución del logaritmo del precio en supermercado de la palta, 2008-2016

5.2.2. Contrastes de raíz unitaria

5.2.2.1. Análisis de las series originales

Para determinar de manera formal la presencia de raíces unitarias en las series analizadas se procede a continuación a aplicar una serie de contrastes comúnmente utilizados en la literatura.

A continuación se aplicará el contraste de Dickey-Fuller aumentado bajo el supuesto de que el proceso subyacente tiene drift y tendencia y bajo el supuesto de que sólo tiene drift.

Cuadro 5.1: Contraste de Dickey-Fuller aumentado (con drift)

		Estimaciones		Valores críticos		
		Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_2	-2.5725	-1.6393		-2.57	-2.87	-3.44
ϕ_1	3.4095	1.6354		3.79	4.61	6.47

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

El cuadro 5.9 muestra que la evidencia estadística provista por la realización de la serie es apenas suficiente para rechazar la hipótesis $\mathcal{H}_0 : \rho = 1$ a un nivel de significancia de 10 %. Por otro lado, la hipótesis $\mathcal{H}_0 : (\alpha, \rho) = (0, 1)$ no puede ser rechazada.

Con respecto a la serie de precios de supermercados la evidencia va en la misma dirección.

De acuerdo a lo anterior, la serie no contendría una raíz unitaria. De todas maneras, es necesario formular otras representaciones del proceso, como bien podría ser incluir una tendencia determinista (a continuación) o utilizar otros contrastes.

El cuadro 5.2 muestra que para el caso de la serie de precios mayoristas, la hipótesis $\mathcal{H}_0 : \rho = 1$ se rechaza a un nivel de significancia de 10 %. Mientras que la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (0, 0, 1)$ no muestra evidencia estadística suficiente para ser rechazada a los niveles de significancia propuestos. Por último, la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (\alpha, 0, 1)$ puede ser rechazada sólo a un 10 % de significancia.

Cuadro 5.2: Contraste de Dickey-Fuller aumentado (con tendencia)

	Estimaciones		Valores críticos		
	Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_3	-3.4073	-2.1143	-3.13	-3.42	-3.98
ϕ_2	3.9655	1.6848	4.05	4.71	6.15
ϕ_3	5.8467	2.2351	5.36	6.30	8.34

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

Para el caso de la serie de precios de supermercado, la hipótesis $\mathcal{H}_0 : \rho = 1$ no puede ser rechazada a ninguno de los niveles de significancia propuestos. Mientras que el contraste tampoco provee información suficiente para rechazar $\mathcal{H}_0 : (\alpha, \beta, \rho) = (0, 0, 1)$. Por último, la hipótesis $\mathcal{H}_0 : (\alpha, \beta, \rho) = (\alpha, 0, 1)$ puede ser rechazada sólo a un 10 % de significancia.

El resultado de este contraste es más bien contradictorio, pues no permite concluir que ambas series contengan una raíz unitaria, bajo los diferentes escenarios que sus hipótesis configuran. Queda la posibilidad entonces de que las series tengan un comportamiento estacionario en tendencia, situación que será abordada con el contraste KPSS.

Los resultados del cuadro 5.3 muestran que para ambas series se rechaza la hipótesis de estacionariedad de la serie a nivel de significancia del 1 %. Razón por la cual se descarta la hipótesis de estacionariedad en tendencia.

Cuadro 5.3: Contraste KPSS (con tendencia determinista)

	Estadístico		Valores críticos		
\mathcal{H}_0	Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_3	0.2405	0.3223	0.119	0.146	0.216

^a: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

^b: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

Por otro lado, los resultados del cuadro 5.4 indican que la hipótesis de estacionariedad en niveles se rechaza fuertemente en ambos casos.

Cuadro 5.4: Contraste KPSS (sin tendencia determinista)

Estadístico			Valores críticos		
\mathcal{H}_0	Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_3	3.5537	3.8913	0.347	0.463	0.739
^a : Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.					
^b : Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.					

El contraste de Phillips Perron indica que no se puede rechazar la hipótesis de raíz unitaria para ambas series.

Cuadro 5.5: Contraste Phillips & Perron^a (con tendencia determinista)

	Mayorista	Supermercado	90 %	95 %	99 %
$Z(t_{\hat{\alpha}})$	-22.1367	-11.958	-3.13	-3.42	-3.98
$Z(t_{\hat{\mu}})$	1.5224	1.0702	4.04	4.71	6.15
$Z(t_{\hat{\beta}})$	2.2482	1.6792	5.63	6.30	8.34
^a : Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.					

El cuadro 5.11 muestra que el contraste rechaza la hipótesis de raíz unitaria, aunque no permite identificar con claridad la estructura de los términos deterministas.

Cuadro 5.6: Contraste Phillips & Perron^a (con tendencia determinista)

	Mayorista	Supermercado	90 %	95 %	99 %
$Z(t_{\hat{\alpha}})$	-13.3185	-6.8276	-2.57	-2.87	-3.44
$Z(t_{\hat{\mu}})$	2.5034	1.8389	3.79	4.61	6.47
^a : Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.					

Al observar el cuadro 5.12 queda de manifiesto el rechazo de la hipótesis de raíz unitaria, aunque por cierto, no se puede afirmar con exactitud el comportamiento de los términos deterministas.

Hecho lo anterior, se procede a aplicar el contraste ERS, utilizando en primer lugar un modelo con constante y tendencia determinista. El cuadro 5.13 muestra que para el caso de los mayoristas no se puede rechazar la hipótesis nula de raíz unitaria para ninguno de los niveles de significancia prescritos, mientras que para el caso de la serie supermercado dicha hipótesis se rechaza al menos al 5 %.

Cuadro 5.7: Contraste de Elliot, Rothenberg & Stock
(con tendencia determinista)

Estadístico		Valores críticos		
Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
3.8715	10.0306	6.89	5.62	3.96

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

El cuadro 5.17

Cuadro 5.8: Contraste de Elliot, Rothenberg & Stock
(con constante)

Estadístico		Valores críticos		
Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
3.0049	12.0408	4.48	3.26	1.99

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

5.2.2.2. Análisis de las series en diferencias

Una vez establecida la presencia de raíces unitarias en las series, se procederá a aplicar los contrastes nuevamente para las series en diferencias

Por otro lado, los resultados del cuadro 5.4 indican que la hipótesis de estacionariedad en niveles se rechaza fuertemente en ambos casos.

El contraste KPSS muestra claramente como, ahora que la serie ha sido diferenciada, no se puede rechazar la hipótesis nula de estacionariedad.

El contraste de Phillips Perron indica que no se puede rechazar la hipótesis de raíz unitaria para ambas series.

Cuadro 5.9: Contraste de Dickey-Fuller aumentado (sin drift y tendencia)

	Estimaciones		Valores críticos		
	Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_3	-14.0093	-14.3253	-1.62	-1.95	-2.58

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

Cuadro 5.10: Contraste KPSS (sin tendencia determinista)

	Estadístico		Valores críticos		
\mathcal{H}_0	Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
τ_3	0.0236	0.0433	0.347	0.463	0.739

^a: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

^b: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

El cuadro 5.11 muestra que el contraste rechaza la hipótesis de raíz unitaria, y que como es de esperar, la serie no contiene término constante.

Al observar el cuadro 5.12 queda de manifiesto el rechazo de la hipótesis de raíz unitaria, aunque por cierto, no se puede afirmar con exactitud el comportamiento de los términos deterministas.

Hecho lo anterior, se procede a aplicar el contraste ERS, utilizando en primer lugar un modelo con constante y tendencia determinista. El cuadro 5.13 muestra que para el caso de los mayoristas no se puede rechazar la hipótesis nula de raíz unitaria para ninguno de los niveles de significancia prescritos, mientras que para el caso de la serie supermercado dicha hipótesis se rechaza al menos al 5 %.

El cuadro 5.17

A partir del análisis recién expuesto, se puede concluir que ambas series al ser diferenciadas tienen un comportamiento estacionario. De forma tal que ambas series son $I(1)$.

De esta forma, se procederá a determinar si la tendencia estocástica es compartida a través de un análisis de cointegración.

Cuadro 5.11: Contraste Phillips & Perron^a (con tendencia determinista)

	Mayorista	Supermercado	90 %	95 %	99 %
$Z(t_{\hat{\alpha}})$	-411.6423	-571.565	-3.13	-3.42	-3.98
$Z(t_{\hat{\mu}})$	0.3197	0.7824	4.04	4.71	6.15

^a: Con un rezago de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

Cuadro 5.12: Contraste Phillips & Perron^a (con tendencia determinista)

	Mayorista	Supermercado	90 %	95 %	99 %
$Z(t_{\hat{\alpha}})$	-13.3185	-6.8276	-2.57	-2.87	-3.44
$Z(t_{\hat{\mu}})$	2.5034	1.8389	3.79	4.61	6.47

^a: Con cinco rezagos de acuerdo a la siguiente regla $\sqrt[4]{4 \times (n/100)}$.

5.3. Determinación del rango de cointegración y estimación del modelo del modelo lineal de corrección del error

En este apartado se utiliza principalmente con el paquete `vars` (Pfaff, 2008).

Antes de la determinación del rango de cointegración del proceso, es necesario tener en cuenta que el uso del contraste de cointegración de Johansen que se utilizará para dichos fines es sensible a la especificación del modelo vectorial de corrección del equilibrio que subyace. Para la especificación de dicho modelo se recurrió a diferentes criterios.

De manera similar a como señalan Ben Kaabia & Gil (2008), se observa que al realizar los contrastes univariantes de raíz unitaria, no se rechaza la existencia de una raíz unitaria en torno a una constante no nula. Así también, al realizar los mismos contrastes para las series en primeras diferencias, las series resultaron estacionarias. Por último, en términos económicos, la presencia de una constante indica que existe una diferencia de escala entre los precios mayoristas y detallistas.

El número de rezagos a incluir en el modelo se determinó con el objetivo de prevenir problemas de autocorrelación. De esta manera, se estimó el modelo secuencialmente

Cuadro 5.13: Contraste de Elliot, Rothenberg & Stock
(con tendencia determinista)

Estadístico		Valores críticos		
Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
3.8715	10.0306	6.89	5.62	3.96

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

Cuadro 5.14: Contraste de Elliot, Rothenberg & Stock
(con constante)

Estadístico		Valores críticos		
Mayorista ^a	Supermercado ^b	90 %	95 %	99 %
3.0049	12.0408	4.48	3.26	1.99

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

utilizando entre 2 y 12 rezagos y se calculó el estadístico de Ljung-Box multivariado para los residuos rezagados entre el orden mínimo posible y aquellos de orden 52. Los p-valores correspondientes a cada uno de los contrastes están resumidos en la figura 5.17. De esta manera, se aprecia que la hipótesis nula de independencia no se puede rechazar a un nivel de significancia de 5 % en la mayoría de los casos para el modelo con ocho términos rezagados.

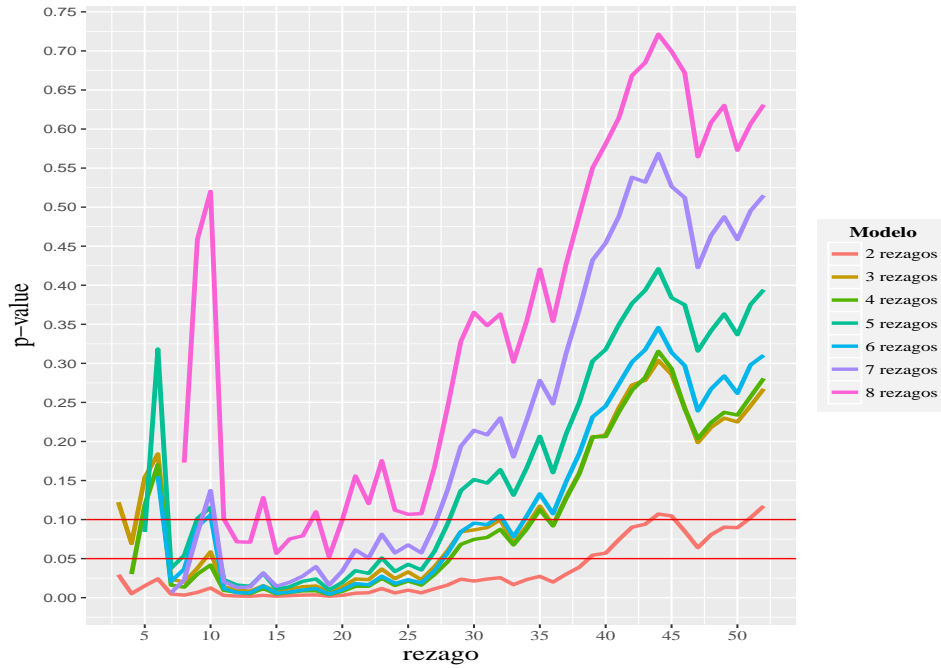


Figura 5.7: Número de Rezagos para el contraste de Independencia

De esta forma, el modelo definitivo a estimar resultó ser el siguiente:

$$\Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \sum_{i=1}^8 \Gamma_i \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t \quad (5.3.1)$$

Entonces, definido el número de rezagos a utilizar y la estructura del vector cointegrante, se procedió a aplicar el contraste de cointegración de Johansen, cuyos resultados pueden apreciarse en los cuadros 5.15 y 5.16

Cuadro 5.15: Contraste de la *la traza* de cointegración de Johansen

Estadístico		Valores críticos		
\mathcal{H}_0	$p = 8$	90 %	95 %	99 %
$r \leq 1$	5.82	7.52	9.24	12.97
$r = 0$	36.63	17.85	19.96	24.60

En el cuadro 5.15 se observa el rechazo de la hipótesis $\mathcal{H}_0 := 0$ lo que descarta que la especificación adecuada sea un modelo VAR en niveles. Por otro lado, la hipótesis nula $\mathcal{H}_0 \leq 1$ no puede rechazarse.

Cuadro 5.16: Contraste del *máximo autovalor* de cointegración de Johansen

	Estadístico	Valores críticos		
\mathcal{H}_0	$p = 8$	90 %	95 %	99 %
$r \leq 1$	5.82	7.52	9.24	12.97
$r = 0$	30.82	13.75	15.67	20.20

Al realizar la prueba de acuerdo al máximo autovalor, se observa el mismo resultado.

De esta manera, el modelo puede especificarse como

$$\Delta \mathbf{y}_t = \boldsymbol{\alpha} \boldsymbol{\beta}' \begin{bmatrix} 1 \\ \mathbf{y}_{t-1} \end{bmatrix} + \sum_{i=1}^8 \boldsymbol{\Gamma}_i \Delta \mathbf{y}_{t-i} + \boldsymbol{\varepsilon}_t \quad (5.3.2)$$

Donde $\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix}$ y $\boldsymbol{\beta}' = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 \end{bmatrix}$

[1] 0.5149079

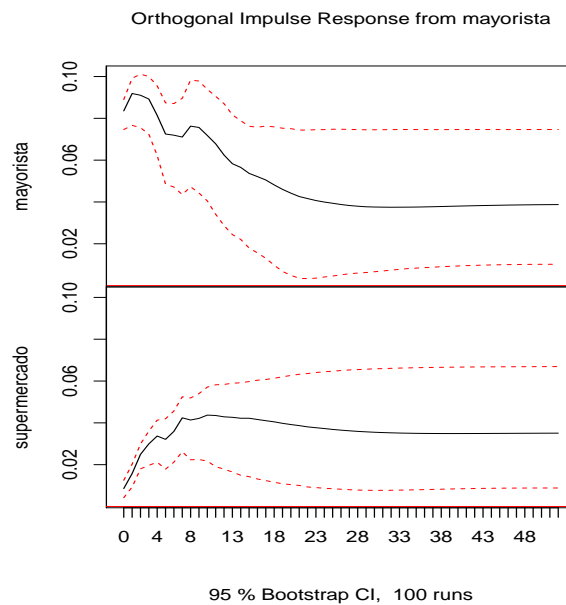


Figura 5.8: Función ortogonal de impulso respuesta

Cuadro 5.17: Contraste de Elliot, Rothenberg & Stock (con constante)

	$\Delta\text{mayorista}_t$	$\Delta\text{supermercado}_t$
corrección del error	-0.146592	0.027821
$\Delta\text{mayorista}_{t-1}$	0.065270	0.095569
$\Delta\text{supermercado}_{t-1}$	0.333719	-0.095055
$\Delta\text{mayorista}_{t-2}$	-0.075599	0.112612
$\Delta\text{supermercado}_{t-1}$	0.296286	-0.042207
$\Delta\text{mayorista}_{t-3}$	-0.095907	0.063245
$\Delta\text{supermercado}_{t-3}$	0.198395	0.006723
$\Delta\text{mayorista}_{t-4}$	-0.174854	0.046165
$\Delta\text{supermercado}_{t-4}$	0.212239	0.053372
$\Delta\text{mayorista}_{t-5}$	-0.175601	-0.004331
$\Delta\text{supermercado}_{t-5}$	0.167277	-0.057114
$\Delta\text{mayorista}_{t-6}$	-0.064872	0.068671
$\Delta\text{supermercado}_{t-1}$	0.148495	-0.011887
$\Delta\text{mayorista}_{t-7}$	-0.122949	0.101965
$\Delta\text{supermercado}_{t-7}$	0.332634	-0.042574
$\Delta\text{mayorista}_{t-8}$	-0.067370	0.009668
$\Delta\text{supermercado}_{t-1}$	0.107484	-0.033937

^a: Con un rezago, de acuerdo a criterio BIC.

^b: Con un rezago, de acuerdo a criterio BIC.

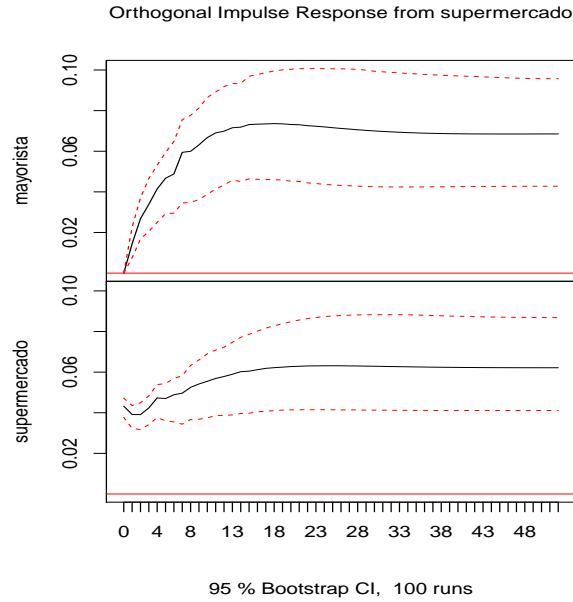


Figura 5.9: Función ortogonal de impulso respuesta

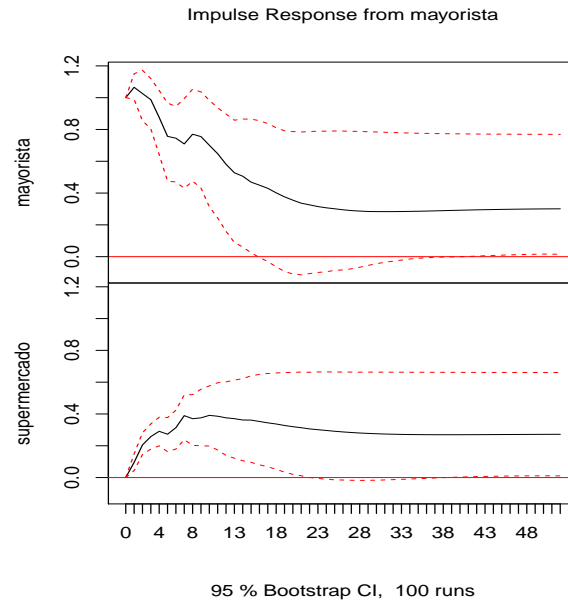


Figura 5.10: Función acumulada de impulso respuesta

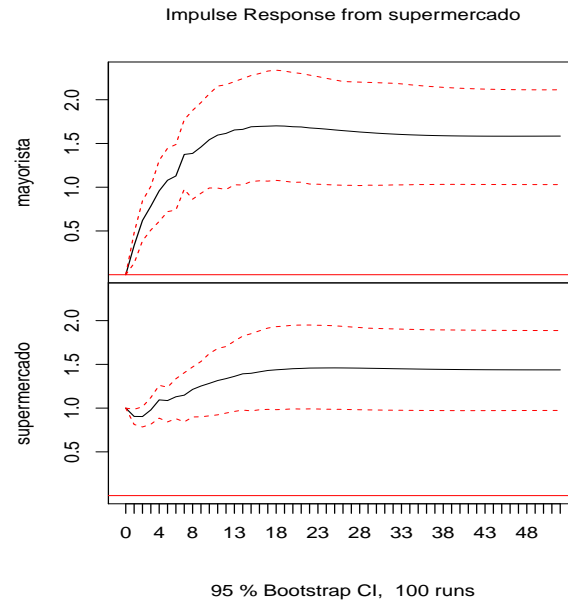


Figura 5.11: Función acumulada de impulso respuesta

5.4. Diagnósticos del modelo lineal

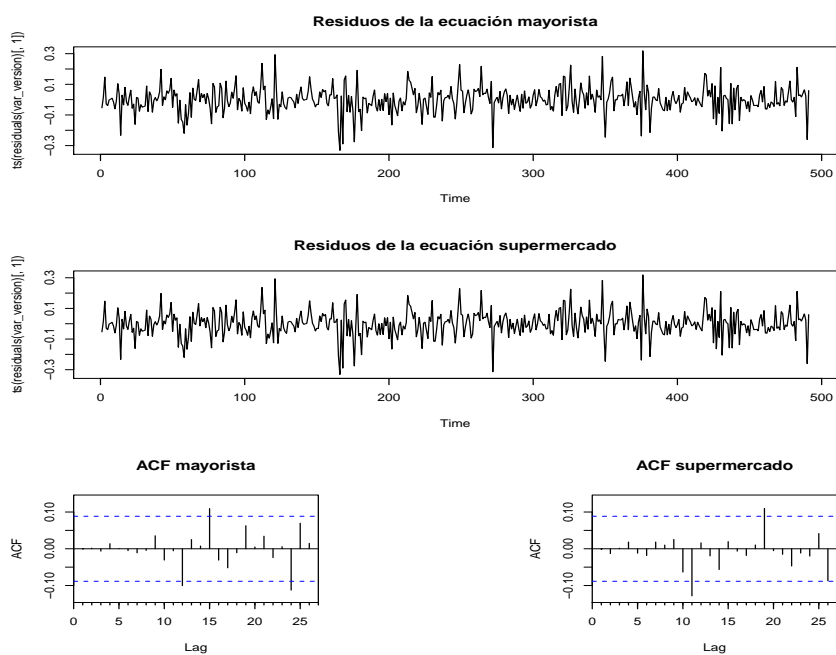


Figura 5.12: Residuos del modelo

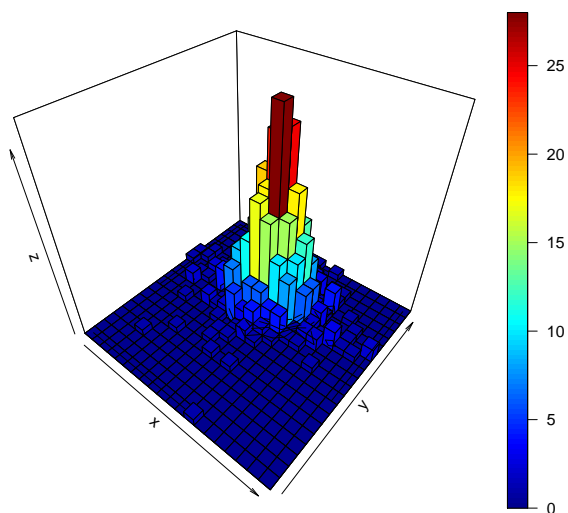


Figura 5.13: Histograma bivariado de los residuos del modelo

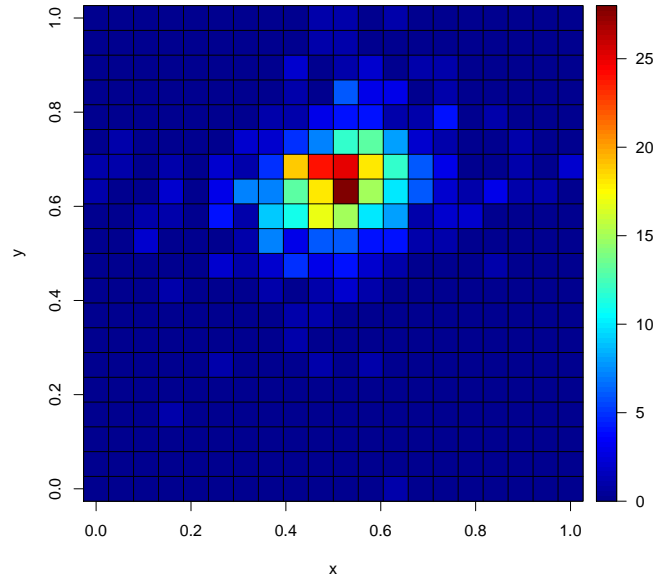


Figura 5.14: Mapa de calor de los residuos del modelo

Cuadro 5.18: Diagnósticos del modelo

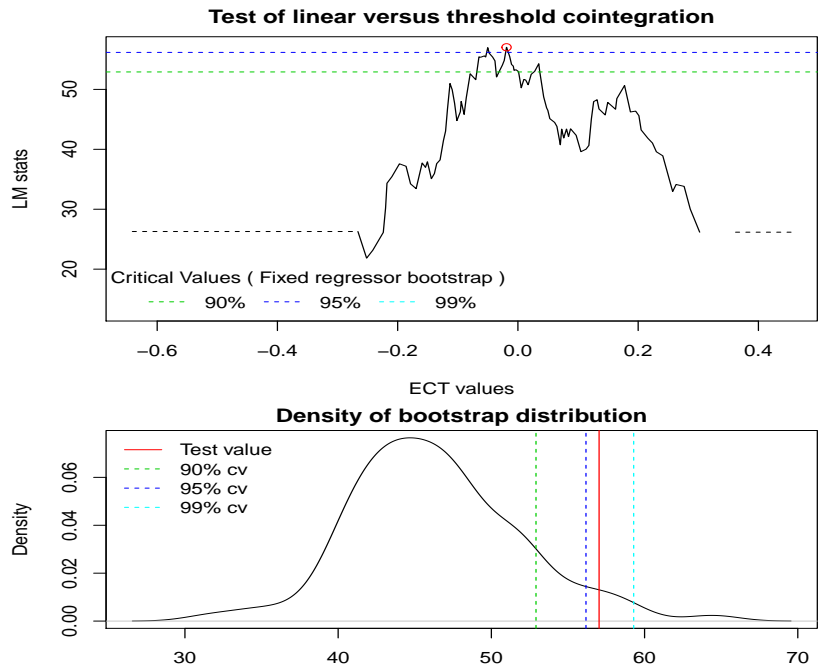
Contraste	E	df	$\mathbb{P}(\chi^2 > E)$
Doornik & Hansen	246.913	4	2.11×10^{-45}
Jarque-Bera (Multivariado)	30.141	4	$< 2.85 \times 10^{-7}$
ARCH (Multivariado)	144.4	81	1.89×10^{-5}

5.5. Determinación del rango de cointegración y estimación del modelo del modelo lineal de corrección del error

Los contrastes y estimaciones de esta sección fueron realizadas a través del paquete `tsDyn` de Di Narzo, Aznarte & Stigler(2009).

El paquete `tsDyn` y la función asociada a este, `TVECM.HStest` no permiten utilizar una constante en la relación de cointegración, por lo que para esto se debió modificar ligeramente el código original de la función recién mencionada, pues como ha sido descrito en el capítulo 3, la estimación del modelo VECM se construye sobre al

base de un modelo lineal, por lo que sólo bastó con incorporar la constante en la estimación del primer paso.



El modelo ha sido estimado con una constante en la relación de cointegración

38780 (97%) points of the grid lead to regimes with percentage of observations <

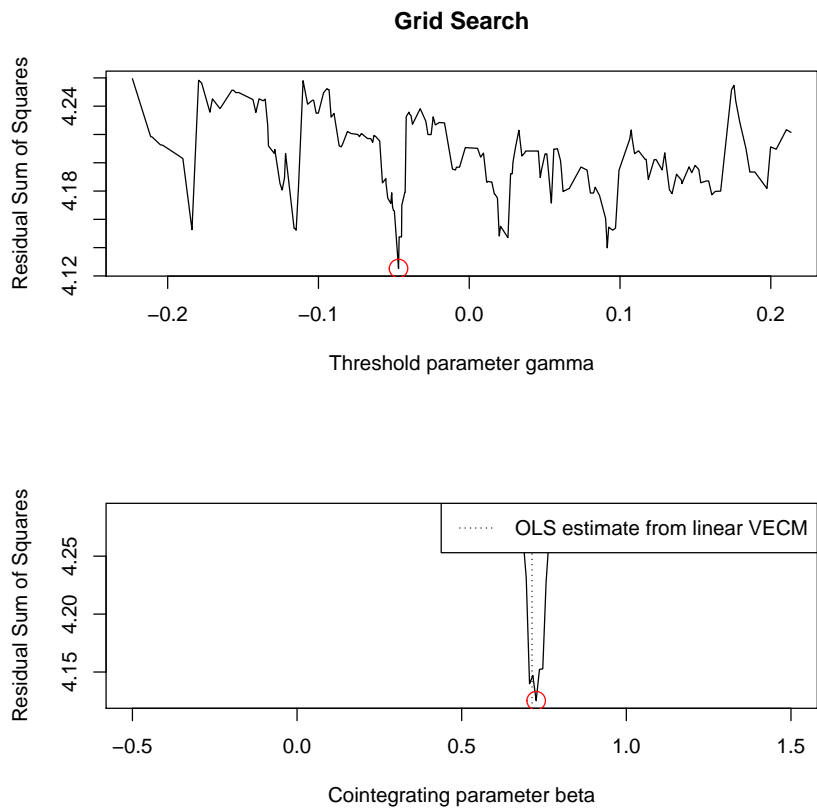


Figura 5.15: Modelo de corrección del error por umbrales

5.6. Diagnósticos del modelo no lineal

```
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
```

```
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
```

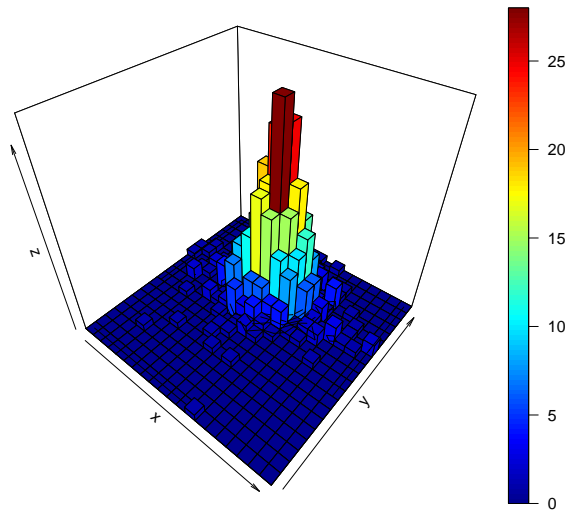


Figura 5.16: Histograma bivariado de los residuos del modelo

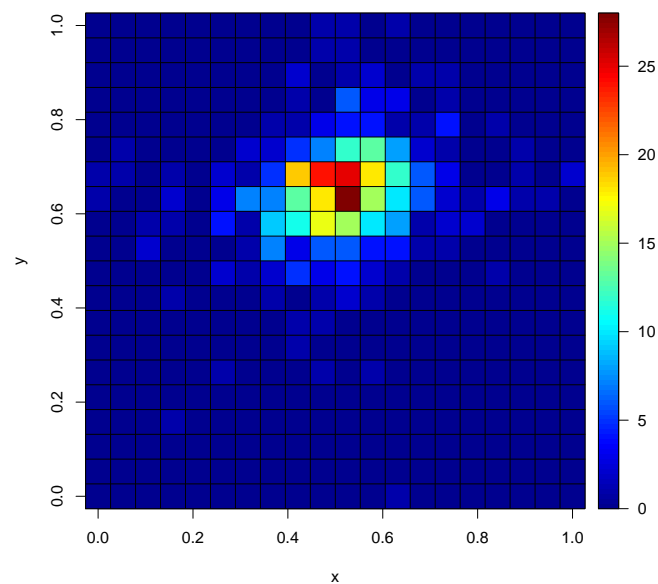


Figura 5.17: Mapa de calor de los residuos del modelo

```
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
```

```
## $multi
```

```
##           E df    P(Chi > E)
```

```
## 1 246.913  4 3.009846e-52
```

```
##
```

```
## $univ
```

```
##           E df    P(Chi > E)
```

```
## resids.of.mayorista    72.63378  2 1.689563e-16
```

```
## resids.of.supermercado 174.27920  2 1.431372e-38
```

```
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
```

```
## Error in residuals(mono_rest): objeto 'mono_rest' no encontrado
```


Capítulo 6

Conclusiones

Capítulo 7

Bibliografía

Ben-Kaabia, M., & Gil Roig, J. M. (2008). Asimetrías en la transmisión de precios en el sector del tomate en España. *Economía Agraria y Recursos Naturales*, 8(1), 57-82.

Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.

Di Narzo, A. F., Aznarte, J. L., & Stigler, M. (2009). tsDyn: Time series analysis based on dynamical systems theory. R package version 0.7.(URL: <http://stat.ethz.ch/CRAN/web/packages/tsDyn/vignettes/tsDyn.pdf>).

Durbin, J., & Koopman, S. J. (2012). *Time series analysis by state space methods* (Vol. 38). OUP Oxford.

Hansen, B. E., & Seo, B. (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of econometrics*, 110(2), 293-318.

Hyndman, R. J., & Khandakar, Y. (2007). Automatic time series for forecasting: the forecast package for R (No. 6/07). Monash University, Department of Econometrics and Business Statistics.

Koop, G., Pesaran, M. H., & Potter, S. M. (1996). Impulse response analysis in nonlinear multivariate models. *Journal of econometrics*, 74(1), 119-147.

Johansen, S., Danielsdottir, A. K., Kristinsson, K., Petersen, P. H., & Naevdal, G. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press

Juselius, K. (2006). *The cointegrated VAR model: methodology and applications*. Oxford university press.

Kreiss, J. P., & Lahiri, S. N. (2012). Bootstrap methods for time series. In Handbook of statistics (Vol. 30, pp. 3-26). Elsevier.

Lo, M. C., & Zivot, E. (2001). Threshold cointegration and nonlinear adjustment to the law of one price. *Macroeconomic Dynamics*, 5(4), 533-576.

Lütkepohl, H. (2005). New introduction to multiple time series analysis. Springer Science & Business Media.

Meyer, J., & Cramon-Taubadel, S. (2004). Asymmetric price transmission: a survey. *Journal of agricultural economics*, 55(3), 581-611.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, 27(4), 1-32.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Capítulo 8

Anexo

8.1. Algunos resultados útiles de cálculo matricial

8.1.1. Derivada de una forma cuadrática

La traza de una matriz cuadrada \mathbf{A} se define como la suma de los elementos de la diagonal a_{ii} de \mathbf{A} :

$$\text{tr}[\mathbf{A}] := \sum_i a_{ii} \quad (8.1.1)$$

Una propiedad útil de la traza es su invariabilidad ante permutaciones

Proposición 8.1.1. Sean $\mathbf{A}, \mathbf{B}, \mathbf{C}$ matrices cuadradas de $n \times n$, se tiene que

$$\text{tr}[\mathbf{ABC}] = \text{tr}[\mathbf{CAB}] = \text{tr}[\mathbf{BCA}] \quad (8.1.2)$$

Proposición 8.1.2. Sean $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ y $\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$. Entonces

$$\mathbf{x}^T \mathbf{A} \mathbf{x} = \sum_{j=1}^n \sum_{i=1}^n a_{ij} x_i x_j \quad (8.1.3)$$

Proposición 8.1.3. Sean $\mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ y $\mathbf{A} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix}$. Entonces

$$\frac{\partial}{\partial \mathbf{A}} \mathbf{x}^T \mathbf{A} \mathbf{x} = \frac{\partial}{\partial \mathbf{A}} \text{tr} [\mathbf{x} \mathbf{x}^T \mathbf{A}] = [\mathbf{x} \mathbf{x}^T]^T = \mathbf{x} \mathbf{x}^T \quad (8.1.4)$$

8.1.2. Derivada de un determinante

Proposición 8.1.4. *Formula de Jacobi*

Para cualquier matriz A