

CUADERNOS DE MATEMÁTICA  
DE LA ESCUELA POLITÉCNICA NACIONAL

HOLGER CAPA SANTOS

---

---

SERIES TEMPORALES

*La ciencia y el arte de los pronósticos*

---

---



## **Cuaderno de Matemática No. 10**

**SERIES TEMPORALES: LA CIENCIA Y EL ARTE DE LOS PRONÓSTICOS**

**HOLGER CAPA SANTOS**

**Responsable de la Edición: ?**

**Revisión técnica: ?**

**Asistentes: Andrés Merino**

**Portada: Andrés Merino**

Registro de derecho autoral No.

ISBN:

Publicado por la Unidad de Publicaciones de la Facultad de Ciencias de la Escuela Politécnica Nacional, Ladrón de Guevara E11-253, Quito, Ecuador.

Primera edición: 2015

Primera impresión: 2015

© Escuela Politécnica Nacional 2015

# Índice general

<b>1</b>	<b>Modelos de Heteroscedasticidad Condicional</b>	<b>3</b>
1.1	Modelos Arch–Garch Simétricos . . . . .	3
1.2	Modelos GARCH asimétricos . . . . .	5
1.3	Metodología de la Modelación . . . . .	7
1.4	Ejemplos con Heteroscedasticidad Condicional . . . . .	8
<b>2</b>	<b>Modelos Multivariantes de Series Temporales</b>	<b>11</b>
2.1	Procesos Estacionarios . . . . .	11
2.1.1	Proceso estrictamente estacionario . . . . .	11
2.1.2	Proceso débilmente estacionario . . . . .	11
2.2	Matrices de Correlación Cruzada (Cross-Correlation) . . . . .	12
2.2.1	Dependencia Lineal . . . . .	13
2.3	Modelos de Vectores Autoregresivos (VAR) . . . . .	17
2.3.1	El caso bivariante . . . . .	17
2.3.2	Representación general de un VAR . . . . .	19
2.4	Representación VARMA de una Serie Multivariante . . . . .	21
2.4.1	La representación VMA . . . . .	21
2.4.2	Representación lineal de un VAR(p) . . . . .	21
2.4.3	La representación VARMA . . . . .	23
2.5	Formulación de un modelo VAR . . . . .	24
2.5.1	Estimación . . . . .	24
	<b>Índice Alfabético</b>	<b>27</b>



# Índice de figuras

1.1	FAC y FACP estimadas residuales del Modelo 3 para la SVM . . . . .	8
1.2	FAC y FACP estimadas de los residuos cuadráticos del Modelo 3 para la SVM	8
1.3	FAC y FACP estimadas residuales del Modelo 3-ARCH(1) para la SVM . . . .	9
1.4	FAC y FACP estimadas de los residuos cuadráticos del Modelo 3-ARCH(1) para la SVM . . . . .	9
2.1	Correlaciones cruzadas entre $X_{1t}$ y $X_{2t}$ . . . . .	16
2.2	Correlaciones cruzadas entre $X_{1t}$ y $X_{3t}$ . . . . .	16
2.3	Correlaciones cruzadas entre $X_{2t}$ y $X_{3t}$ . . . . .	17



# Índice de tablas

1.1	Información sobre los coeficientes del Modelo 3 para la SVM . . . . .	8
1.2	Información estadística para el Modelo 3-ARCH(1) para la SVM . . . . .	9
2.1	Resumen de estadísticas y matrices de correlación cruzada para $X_{1t}$ , $X_{2t}$ y $X_{3t}$	15





# Notaciones

- $E(.)$  : Esperanza matemática de una variable aleatoria.
- $V(.)$  : Varianza de una variable aleatoria.
- $\text{Cov}(.,.)$  : Covarianza de dos variables aleatorias.
- $N(\mu, \sigma^2)$  : Distribución normal con media  $\mu$  y varianza  $\sigma^2$ .
- $\hat{\cdot}$  : Denota un estimador del parámetro estadístico que se encuentra bajo el símbolo.
- $c$  : De manera general, representa a una constante numérica.
- i.i.d : Representa a la frase: independientes e idénticamente distribuidas.
- i.i.d (0,1) : Variables aleatorias independientes e idénticamente distribuidas con esperanza matemática 0 y varianza 1.
- $\gamma(l)$  : Función de autocovarianza de orden  $l$ .
- $\rho(l)$  : Función de autocorrelación de orden  $l$ .
- $r(l)$  : Función de autocorrelación parcial de orden  $l$ .
- $1_A$  : Función indicatriz sobre el conjunto  $A$ .
- $L_c^2(\Omega)$  : Espacio de variables aleatorias de  $\Omega$  en  $C$ , de cuadrado integrable.
- $L_c^2(E)$  : Espacio de funciones medibles de  $E$  en  $C$ , de cuadrado integrable.
- $\Sigma$  : Matriz de varianzas - covarianzas
- $\Sigma_u$  : Matriz de varianzas – covarianzas residuales
- $\therefore$  : Significa: por tanto



# Capítulo 1

## Modelos de Heteroscedasticidad Condicional

### 1.1 Modelos Arch–Garch Simétricos

En esta sección, se presentará un breve resumen de la teoría desarrollada alrededor de los modelos ARCH – GARCH, los cuales determinan un patrón de comportamiento estadístico para la varianza condicional, denominados modelos Auto-Regresivos con Heteroscedasticidad Condicional.

El análisis de series temporales económicas, tradicionalmente se ha centrado en el estudio de modelos para la media condicional en los que se asume que la varianza condicional es constante; en este caso se dice que existe homoscedasticidad. Si este no fuera el caso, se estaría enfrentado un problema de no estacionariedad de la serie.

En 1982, Robert Engle revolucionó los modelos de volatilidad introduciendo el estudio de las estructuras cuadráticas, ampliando así la visión de la metodología Box–Jenkins, en la cual los modelos lineales de tipo ARIMA admiten que las innovaciones son un ruido blanco, con media cero y varianza constante.

Los modelos de volatilidad condicional son importantes por el papel que juega el riesgo y el concepto de incertidumbre en el desarrollo de las teorías modernas de modelos financieros, que relacionan de forma directa el riesgo con la volatilidad. Estos modelos permiten relacionar el valor de la varianza condicional (no constante) en función del conjunto de información disponible en periodos anteriores; están específicamente diseñados para modelar y pronosticar varianzas condicionales.

**Definición 1.1.** Un Modelo ARCH( $r$ ) se define por:

$$\begin{aligned}Z_t &= \sqrt{h_t}u_t \\h_t &= \alpha_0 + \alpha_1 Z_{t-1}^2 + \dots + \alpha_r Z_{t-r}^2\end{aligned}$$

donde los  $(u_t)$  son independientes e idénticamente distribuidos con media cero y varianza uno (i.i.d.  $(0,1)$ ),  $\alpha_0 > 0$  y  $\alpha_i \geq 0$  para  $i > 0$ .

En la práctica se supone que  $u_t \sim N(0,1)$  o que siguen una distribución t-student. Los coeficientes  $\alpha_i$  deben satisfacer ciertas condiciones (en general no negatividad:  $\alpha_0 > 0$  y

$\alpha_i \geq 0$  para  $i > 0$  y  $\sum_{i=1}^r \alpha_i < 1$ ), dependiendo del tipo de restricciones que se coloquen sobre el proceso  $Z_t$ .

Las restricciones de signo de los coeficientes de la ecuación de varianza garantizan que la varianza condicional sea positiva en todos los períodos. Lo que se necesita es que, una vez que se haya estimado el modelo, genere una serie de varianzas positiva, lo que puede suceder aún si algunos de los coeficientes  $\alpha_i$  fueran negativos. Esto se puede considerar como un contraste de validez del modelo.

Por la propia definición, a valores grandes de  $Z_t$  les siguen otros valores grandes de la serie.

$h_t$  es la varianza condicional de  $Z_t$  dado  $Z_s$ ,  $s < t$ .

**Observación 1.1.** Se puede demostrar que un proceso  $ARCH(r)$  implica que se puede representar  $X_t^2$  como un proceso  $AR(r)$ , con residuos que no son gaussianos. Además, el coeficiente de apuntamiento (curtosis) es mayor que 3, por lo cual las colas de la distribución serán más pesadas que en la distribución normal.

**Definición 1.2.** Un modelo  $GARCH(r, s)$ , está definido por:

$$Z_t = \sqrt{h_t} u_t$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i Z_{t-i}^2 + \sum_{j=1}^s \beta_j h_t$$

donde los  $u_t$  son i.i.d.  $(0, 1)$  con  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  para  $i > 0$  y

$$\sum_{i=1}^r \alpha_i + \sum_{j=1}^s \beta_j < 1$$

Como en el caso de los modelos  $ARCH$ , usualmente se supone que los  $u_t$  son normales o siguen una distribución  $t$ -student.

Las restricciones de no negatividad impuestas sobre los coeficientes en la definición son para asegurarse que la varianza no llegue a tomar un valor negativo. Además, en algunos casos se pueden encontrar condiciones menos restrictivas para estos coeficientes, que aseguren que la varianza sea positiva.

**Observación 1.2.** Se puede demostrar que un proceso  $GARCH(rs)$  implica que se puede representar  $Z_t^2$  como un proceso  $ARMA(rs)$ , con residuos que no son gaussianos (más precisamente con residuos que representan una diferencia martingala). También, en este caso, el coeficiente de apuntamiento (curtosis) es mayor que 3, por lo cual las colas de la distribución serán más pesadas que en la distribución normal.

## Modelos IGARCH

Este modelo fue descrito originalmente por Engle y Bollerslev (1986). Si el modelo polinomial AR del modelo  $GARCH$  tiene una raíz unitaria, se tiene un modelo  $IGARCH$ . Los modelos  $IGARCH$  son modelos  $GARCH$  con una raíz unitaria. Estos modelos tienen la característica que los impactos de los choques al cuadrado sobre  $Z_t^2$  son persistentes; esto se

puede expresarse por:

$$\eta_{t-i} = Z_{t-i}^2 - h_{t-i}, \quad \text{para } i > 0$$

**Definición 1.3.** Un modelo IGARCH( $r, s$ ), se define por:

$$\begin{aligned} Z_t &= \sqrt{h_t} u_t \\ h_t &= \alpha_0 + \sum_{i=1}^r \alpha_i Z_{t-i}^2 + \sum_{j=1}^s (1 - \beta_j) h_{t-j} \end{aligned}$$

tal que,

$$\sum_{j=1}^s \beta_j + \sum_{i=1}^r \alpha_i = 1$$

Como en el caso de los modelos ARCH, usualmente se supone que los  $u_t$  son normales o siguen una distribución  $t$ -student.

Los modelos IGARCH son un caso específico dentro de la familia de los “modelos con varianza persistente” en los que la información actual (en el instante “ $t$ ”) es importante para realizar predicciones óptimas con cualquier horizonte temporal.

## 1.2 Modelos GARCH asimétricos

Una característica de los modelos GARCH, dado que la varianza depende fundamentalmente de valores cuadráticos pasados, es que la volatilidad que generan frente a cambios positivos o negativos inesperados de la variable (piénsese en retornos de inversiones), dan una respuesta simétrica a estos. Sin embargo, se ha probado empíricamente que la reacción que tiene la volatilidad o varianza condicional de muchas variables financieras a este tipo de cambios es asimétrica; es decir, existe diferencia en la respuesta de la volatilidad de la variable, dependiendo de si el cambio es positivo o negativo.

Con el fin de modelar esta respuesta asimétrica se han desarrollado una variedad de modelos asimétricos; los más representativos son los modelos EGARCH, TARCH, PARCH.

### Modelos EGARCH

Las restricciones de no negatividad para asegurar la positividad de la varianza en los modelos GARCH en muchas ocasiones son difíciles de lograr. Nelson (1991) propuso los modelos EGARCH (*Exponential GARCH*), como solución a este problema; éstos, además, incorporan efectos asimétricos.

**Definición 1.4.** Un modelo EGARCH( $r, s$ ), se define por:

$$\begin{aligned} Z_t &= \sqrt{h_t} u_t \\ \ln(h_t) &= \alpha_0 + \sum_{j=1}^s \beta_j \ln(h_{t-j}) + \sum_{i=1}^r (\alpha_i |u_t| + \gamma_i u_t) \end{aligned}$$

Nótese que el lado izquierdo de la ecuación es el logaritmo de la varianza condicional; esto implica que su efecto es exponencial y garantiza que las predicciones de ésta serán no

negativas. Cuando  $\gamma_i \neq 0$ , el efecto asimétrico deber incorporarse al Modelo GARCH.

Obsérvese también que se puede escribir  $\frac{Z_t}{\sqrt{h_t}}$  en lugar de  $u_t$ , en las fórmulas anteriores.

Como en el caso de los modelos ARCH, usualmente se supone que los  $u_t$  son normales o siguen una distribución  $t$ -student.

## Modelos TARCH

Los modelos TARCH (*Threshold ARCH*) fueron introducidos independientemente por Zakoian (1990) y Glosten, Jaganathan y Runklen (1993), por lo que también se conocen como GJR-GARCH. Estos modelos incluyen una variable adicional  $d_t$ , que determina el carácter asimétrico del modelo.

**Definición 1.5.** Un modelo  $TARCH(r, s)$ , se define por:

$$Z_t = \sqrt{h_t} u_t$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i Z_{t-i}^2 + \gamma Z_{t-1}^2 d_{t-1} + \sum_{j=1}^s \beta_j h_{t-j}$$

donde los  $u_t$  son i.i.d.  $(0, 1)$  con  $\alpha_0 > 0$ ,  $\alpha_i \geq 0$ ,  $\beta_j \geq 0$  para  $i, j > 0$  y  $\gamma \neq 0$

$$d_t = \begin{cases} 1, & \text{si } Z_t < 0 \\ 0, & \text{si } Z_t \geq 0 \end{cases}$$

Si  $\gamma = 0$  se pierde el efecto asimétrico del modelo.

En este modelo, las malas noticias ( $Z_t < 0$ ) y las buenas noticias ( $Z_t \geq 0$ ) (piénsese otra vez en retornos), tienen efectos diferentes sobre la varianza condicional.

## Modelos PARCH

Los modelos PARCH (*Power ARCH*) desarrollados independientemente por Taylor (1986) y Schwert (1989), introducen la desviación estándar a los modelos ARCH; donde se modela la desviación estándar en lugar de la varianza. Este modelo fue generalizado por Ding y otros (1993).

En el modelo PARCH, el parámetro de potencia  $\delta$  de la desviación estándar puede ser estimado antes que impuesto y los parámetros opcionales  $\gamma$  se agregan para capturar la asimetría dentro de los datos.

**Definición 1.6.** Un modelo  $PARCH(r, s)$ , se define por:

$$Z_t = \sqrt{h_t} u_t$$

$$h_t^\delta = \alpha_0 + \sum_{i=1}^r \alpha_i (|u_{t-i}| - \gamma_i u_{t-i})^\delta + \sum_{j=1}^s \beta_j h_{t-j}^\delta$$

donde  $\delta > 0$ , es el parámetro del término de la potencia.  $\gamma_i$  se dicen los parámetros de apalancamiento.

En series de valores simétricos  $\gamma_i = 0$  para todo  $i$ . Nótese que si  $\delta = 1$  y  $\gamma_i = 0$  para todo  $i$ , el modelo PARCH es simplemente una especificación GARCH estándar. Si los  $\gamma_i = 0$  se pierde el efecto asimétrico del modelo.

Los modelos GARCH asimétricos, se estiman por el método de máxima verosimilitud condicional, por lo cual se requiere de ciertos supuestos acerca del comportamiento de los errores. Por lo general, se suponen i.i.d con distribución normal o incluso con una distribución  $t$ -student.

## 1.3 Metodología de la Modelación

El objetivo es encontrar un modelo que represente adecuadamente a los datos históricos de una determinada variable, combinando especificaciones tanto para la media como para la varianza condicional. Los tipos de modelos que se considerarán serán los ARIMA – GARCH, de tal manera que la media condicional de la serie sea descrita por un modelo del tipo ARIMA y su varianza condicional por uno de la familia de modelos ARCH – GARCH o de sus extensiones asimétricas PARCH, TARCH y EGARCH. La modelación se realizará utilizando el paquete EViews.

El primer paso es, por tanto, modelar la serie de datos por un modelo del tipo ARIMA o incluso SARIMA, con lo que se obtiene un modelo para la media condicional de la serie.

Luego de haberse eliminado toda correlación lineal en la serie, se debe indagar si existe heteroscedasticidad condicional residual, para lo cual deben analizarse los residuos estandarizados estimados al cuadrado; el correleograma correspondiente, permite llevar a cabo un análisis gráfico de identificación, para ver si algún valor es estadísticamente diferente de cero, y por tanto, existe autocorrelación en su forma residual cuadrática.

Si se verifica la existencia de heteroscedasticidad condicional en los residuos, se rechaza el supuesto de la varianza constante; se intentará entonces obtener una especificación para la varianza condicional, a través de la modelación de los residuos estimados obtenidos por el modelo ARIMA, mediante un modelo del tipo ARCH – GARCH o sus extensiones asimétricas.

Inicialmente se mantiene la estructura para la media condicional, obtenida por el modelo ARIMA, pero esta puede modificarse con la nueva especificación. Los residuos estimados deben analizarse, tanto en su forma simple como en la cuadrática, para eliminar toda evidencia de autocorrelación lineal (deben aceptarse como un ruido blanco).

La estimación y verificación permiten encontrar uno o varios modelos que cumplan las condiciones que se impusieron en la modelación ARIMA; es decir, todos los coeficientes deben ser significativos; las raíces de los polinomios característicos, tanto de la parte autorregresiva como de la media móvil, deben estar fuera del círculo unidad, para así asegurar la estacionariedad e invertibilidad del proceso. Además, los coeficientes de la ecuación de la varianza condicional deben satisfacer las restricciones de no negatividad para la varianza (modelos ARCH – GARCH).

Para la verificación de la presencia de una estructura ARIMA en los residuos (simples o cuadráticos) pueden utilizarse la FAC y la FACP; además, también se debe realizar la prueba global (estadístico  $Q$ ) de Box – Pierce – Ljung.

Una vez que un modelo ha sido estimado y ha superado las diversas verificaciones, se convierte en un instrumento útil para las predicciones de valores futuros. Como en la modelación ARIMA, si varios modelos son plausibles, se elige entre estos al mejor, mediante

los criterios ya citados previamente.

## 1.4 Ejemplos con Heteroscedasticidad Condicional

Aunque los datos de las ventas que se vienen utilizando no corresponden al ámbito financiero, sirven muy bien para ilustrar la modelación para la varianza condicional. En esta ocasión se adoptará el Modelo 3 con el cual se modeló la media condicional (SARIMA). La Figura 4.2 no permite aceptar la hipótesis de que la serie tenga una varianza constante.

Una posibilidad para amortiguar los efectos de varianza no constante es utilizar la transformación logaritmo o, en general, la transformación de Box y Cox; sin embargo, en esta ocasión se tratará de modelar directamente la varianza a través de los Modelos ARCH-GARCH o sus extensiones asimétricas.

En la Tabla 4.1 y en las figuras 4.1 y 4.2 se presentan la información estadística y residual para el Modelo 3 de la SVM:

**Tabla 1.1:** Información sobre los coeficientes del Modelo 3 para la SVM

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	156.1661	34.82541	4.484257	0.0000
AR(1)	0.325909	0.107494	3.031876	0.0033
AR(12)	-0.335945	0.114892	-2.924008	0.0045
MA(13)	0.480540	0.111556	4.307611	0.0000

**Figura 1.1:** FAC y FACP estimadas residuales del Modelo 3 para la SVM

**Figura 1.2:** FAC y FACP estimadas de los residuos cuadráticos del Modelo 3 para la SVM

Se observan fuertes correlaciones entre los residuos cuadráticos estandarizados estimados, por lo cual se hace necesaria la modelación de la varianza condicional del Modelo 3.

En general, es difícil establecer el orden para los modelos ARCH-GARCH. Lo usual es probar los modelos con parámetros (1,0), (1,1), (1,2) o (2,2). Para este caso se empezó probando con el modelo  $ARCH(1)$ ; los resultados aparecen en la Tabla 4.2 y en las Figuras 4.3 y 4.4.



**Tabla 1.2:** Información estadística para el Modelo 3-ARCH(1) para la SVM

Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	115.3958	36.57140	3.155355	0.0016
AR(1)	0.483776	0.103019	4.695993	0.0000
AR(12)	-0.449722	0.078830	-5.704935	0.0000
MA(13)	0.798959	0.041313	19.33915	0.0000
Variance Equation				
C	16129.50	3530.889	4.568113	0.0000
RESID(-1) <sup>2</sup>	0.610436	0.249248	2.449115	0.0143
R-squared	0.335639	Mean dependent var	157.1548	
Adjusted R-squared	0.310726	S.D. dependent var	258.2531	
S.E. of regression	214.4082	Akaike info criterion	13.23093	
Sum squared resid	3677671.	Schwarz criterion	13.40456	
Log likelihood	-549.6992	Hannan-Quinn criter.	13.30073	
Durbin-Watson stat	2.201103			
Inverted AR Roots	,96 - ,24i	,96 + ,24i	,71 + ,65i	,71 - ,65i
	,28 - ,89i	,28 + ,89i	-,21 + ,90i	-,21 - ,90i
	-,63 - ,66i	-,63 + ,66i	-,87 + ,24i	-,87 - ,24i
Inverted MA Roots	.95-.24i	.95+.24i	.74-.65i	.74+.65i
	,35 + ,92i	,35 - ,92i	-,12 - ,98i	-,12 + ,98i
	-,56 - ,81i	-,56 + ,81i	-,87 + ,46i	-,87 - ,46i
	,98			

**Figura 1.3:** FAC y FACP estimadas residuales del Modelo 3-ARCH(1) para la SVM**Figura 1.4:** FAC y FACP estimadas de los residuos cuadráticos del Modelo 3-ARCH(1) para la SVM

Las Figuras 4.3 y 4.4 evidencian que existen problemas ya no solo en los residuos cuadráticos, sino también en los residuos simples. En la figura 4.4, la FACP en el orden 13 es significativo (y cercano a la estacionalidad 12); por lo cual, se decidió incluir un término AR(13) en el Modelo 3; esto tampoco solucionó totalmente la falta de independencia de los residuos cuadráticos. Luego, de algunas pruebas se encontró como modelo final aquel que contiene términos c, SAR(12), MA(12) y AR(13) para la media (se lo llamará Modelo 4) y ARCH(1) para la varianza. Los resultados se muestran en la Tabla 4.3 y en las Figuras 4.5 y 4.6.



## Capítulo 2

# Modelos Multivariantes de Series Temporales

Una serie temporal multivariante es un proceso estocástico  $(X_t)_{t \in \mathbb{Z}}$ , con  $X_t$  un vector donde cada componente se define como una serie temporal univariante. En este documento se utilizará la notación de vectores como columnas. Así se denota:

$$X_t = (X_{1t}, \dots, X_{kt})' \text{ el vector de } k \text{ series univariantes en el instante } t$$

Lo importante de tratar series multivariantes es que, a más de considerar simultáneamente observaciones de dos o más series univariantes, también se puede analizar las correlaciones existentes entre ellas; esto evidentemente enriquece el análisis, aunque los procesos operativos serán más complejos que en el caso univariante.

### 2.1 Procesos Estacionarios

Para poder estimar las características de los procesos se necesita suponer que son estables a lo largo del tiempo; esto implica, que son estacionarios.

#### 2.1.1 Proceso estrictamente estacionario

Un proceso estocástico multivariado  $(X_t)_{t \in \mathbb{Z}}$ , con  $X_t = (X_{1t}, \dots, X_{kt})'$ , es estrictamente estacionario (o fuertemente estacionario) si las distribuciones conjuntas de cualquier conjunto finito de variables se mantienen por saltos.

Es decir, si:

$$F_{t_1+l, \dots, t_k+l}(x_{t_1+l}, \dots, x_{t_k+l}) = F_{t_1, \dots, t_k}(x_{t_1}, \dots, x_{t_k})$$

Para todo  $k \in \mathbb{N}$  y para todo  $t_1, \dots, t_k, l \in \mathbb{Z}$

Donde,  $F_{t_1, \dots, t_k}$  denota la distribución conjunta de  $X_{t_1}, \dots, X_{t_k}$ .

#### 2.1.2 Proceso débilmente estacionario

Un proceso estocástico multivariado  $(X_t)_{t \in \mathbb{Z}}$ , con  $X_t = (X_{1t}, \dots, X_{kt})'$ , se dice que es *débilmente estacionario* si sus momentos de primer y segundo orden son invariantes en el tiempo

(no dependen de  $t$ ); es decir:

1.  $E(X_t) = \mu \quad \forall t$  (el vector media es constante).
2.  $Cov(X_t, X_{t-l}) = E[(X_t - \mu)(X_{t-l} - \mu)'] = \Gamma_l \quad \forall t$ ; es decir, la matriz de *covarianzas cruzadas* entre  $X_t$  y  $X_{t-l}$  es independiente de  $t$  (solo depende del salto  $l$ ).

La media  $\mu$  es un vector  $k$ -dimensional compuesto por las esperanzas de las componentes de  $X_t$ . La matriz de covarianzas cruzadas es de orden  $k * k$ .

El  $i$ -ésimo elemento de la diagonal de  $\Gamma_0$  es la varianza de  $X_{it}$ ; mientras que, el elemento  $(i, j)$  de  $\Gamma_0$  es la covarianza entre  $X_{it}$  y  $X_{jt}$ . El elemento  $(i, j)$  de  $\Gamma_l$  es la covarianza entre  $X_{it}$  y  $X_{j,t-l}$ .

**Observación 2.1.** Se puede demostrar que si un proceso  $(X_t)_{t \in \mathbb{Z}}$  es débilmente estacionario entonces también lo será cada una de sus componentes.

## 2.2 Matrices de Correlación Cruzada (Cross-Correlation)

En lo que sigue se considera que  $(X_t)_{t \in \mathbb{Z}}$  es estacionaria.

Sea  $D$  una matriz diagonal de orden  $k * k$  compuesta por las desviaciones estándar de  $X_{it}$  para  $i = 1, \dots, k$ , que se denota por:  $D = \text{diag} \{ \sqrt{\Gamma_{11}(0)}, \dots, \sqrt{\Gamma_{kk}(0)} \}$ . La matriz de correlaciones cruzadas de  $X_t$  se define como:

$$\rho_0 \equiv [\rho_{ij}(0)] = D^{-1} \Gamma_0 D^{-1}$$

De manera particular, el elemento  $(i, j)$  de  $\rho_0$  es:

$$\rho_{ij}(0) = \frac{\Gamma_{ij}(0)}{\sqrt{\Gamma_{ii}(0)\Gamma_{jj}(0)}} = \frac{Cov(X_{it}, X_{jt})}{de(X_{it})de(X_{jt})},$$

donde,  $de(\cdot)$  es la desviación estándar.

$\rho_{ij}(0)$  es el coeficiente de correlación lineal entre  $X_{it}$  y  $X_{jt}$ . En el análisis de series de tiempo, dicho coeficiente se conoce como de concurrencia (en el mismo instante). Es fácil ver que:

- i)  $\rho_{ij}(0) = \rho_{ji}(0)$
- ii)  $-1 \leq \rho_{ij}(0) \leq 1$
- iii)  $\rho_{ii}(0) = 1$

Así,  $\rho(0)$  es una matriz simétrica con 1 en la diagonal.

Hay que mencionar que las matrices  $\Gamma_l$  contienen las **relaciones en retardo** entre las componentes de las series. Por lo tanto, las matrices de correlación cruzada se utilizan para medir la fuerza de la dependencia lineal entre las series de tiempo.

La matriz de correlación cruzada de  $X_t$  con  $X_{t-l}$  se define como:

$$\rho_l \equiv [\rho_{ij}(l)] = D^{-1} \Gamma_l D^{-1}$$

donde,  $D$  es la matriz diagonal de las desviaciones estándar de las series individuales. De la definición se tiene:

$$\rho_{ij}(l) = \frac{\Gamma_{ij}(l)}{\sqrt{\Gamma_{ii}(0) \Gamma_{jj}(0)}} = \frac{\text{Cov}(X_{it}, X_{jt-l})}{\text{de}(X_{it}) \text{de}(X_{jt})} = \frac{\text{Cov}(X_{it}, X_{jt-l})}{\text{de}(X_{it}) \text{de}(X_{jt-l})}$$

que es el coeficiente de correlación lineal entre  $X_{it}$  y  $X_{jt-l}$ . Cuando  $l > 0$ , este coeficiente de correlación mide la dependencia lineal de  $X_{it}$  con respecto  $X_{jt-l}$  ( $X_{t-l}$  ocurre con anterioridad al instante  $t$ ). Consecuentemente, si  $\rho_{ij}(l) \neq 0$  y  $l > 0$ , se dice que la serie  $X_{jt}$  **conduce** a la serie  $X_{it}$  con retardo  $l$ .

Similarmente,  $\rho_{ji}(l)$  mide la dependencia lineal de  $X_{jt}$  con respecto a  $X_{it-l}$  y se puede decir que la serie  $X_{it}$  **conduce** a la serie  $X_{jt}$ , con retardo  $l$ , si  $\rho_{ij}(l) \neq 0$  y  $l > 0$ .

Se pueden mencionar las siguientes propiedades cuando  $l > 0$ :

1. En general,  $\rho_{ij}(l) \neq \rho_{ji}(l)$  para  $i \neq j$ , porque los dos coeficientes de correlación miden diferentes relaciones lineales entre las series. Por lo tanto,  $\Gamma_l$  y  $\rho_l$  son, generalmente, no simétricas.
2. Utilizando  $\text{Cov}(X, Y) = \text{Cov}(Y, X)$  y suponiendo que las series son estacionarias, se tiene:

$$\text{Cov}(X_{it}, X_{jt-l}) = \text{Cov}(X_{jt-l}, X_{it}) = \text{Cov}(X_{jt}, X_{it+l}) = \text{Cov}(X_{jt}, X_{it-(-l)})$$

Así que  $\Gamma_{ij}(l) = \Gamma_{ji}(-l)$ , donde  $\Gamma_{ji}(-l)$  es el elemento  $(j, i)$  de  $\Gamma_{-l}$ ; la igualdad se cumple para  $1 \leq i, j \leq k$ . Es decir,  $\Gamma_l = \Gamma'_{-l}$ .

### 2.2.1 Dependencia Lineal

Considérense las matrices de correlación cruzada  $\{\rho(l) | l = 0, 1, 2, \dots\}$  de una serie temporal vectorial estacionaria; éstas contienen la siguiente información:

1. Los elementos de la diagonal de la matriz de correlación cruzada  $\rho_{ii}(l)$  son las funciones de autocorrelación de  $X_{it}$ .
2. El elemento fuera de la diagonal  $\rho_{ij}(0)$  mide la relación lineal de concurrencia entre  $X_{it}$  y  $X_{jt}$ .
3. Para  $l > 0$ , el elemento fuera de la diagonal  $\rho_{ij}(l)$  mide la dependencia lineal de  $X_{it}$  con respecto a  $X_{jt-l}$ .

Por lo tanto, si  $\rho_{ij}(l) = 0$  para todo  $l > 0$ ,  $X_{it}$  no depende linealmente de ningún valor del pasado  $X_{jt-l}$ .

### Resumen e interpretación.

En general, la relación lineal entre dos series de tiempo  $X_{it}$  y  $X_{jt}$  puede resumirse en la siguiente forma:

1.  $X_{it}$  y  $X_{jt}$  no tienen relación lineal si  $\rho_{ij}(l) = \rho_{ji}(l) = 0$ ,  $\forall l \geq 0$ .
2.  $X_{it}$  y  $X_{jt}$  están al mismo tiempo correlacionadas si  $\rho_{ij}(0) \neq 0$ .
3.  $X_{it}$  y  $X_{jt}$  no tienen relación de avance-retardo si  $\rho_{ij}(l) = 0$  y  $\rho_{ji}(l) = 0$ ,  $\forall l > 0$ . En este caso, se dice que las series son desacopladas.
4. Existe una *relación unidireccional* desde  $X_{it}$  hacia  $X_{jt}$  si  $\rho_{ij}(l) = 0$ ,  $\forall l > 0$ , pero  $\rho_{ji}(v) \neq 0$  para algún  $v > 0$ . En este caso,  $X_{it}$  no depende de ningún valor del pasado de  $X_{jt}$ , pero  $X_{jt}$  depende de algún valor del pasado de  $X_{it}$ .
5. Existe una *relación de retroalimentación* entre  $X_{it}$  y  $X_{jt}$  si  $\rho_{ij}(l) \neq 0$  para algún  $l > 0$  y  $\rho_{ji}(v) \neq 0$  para algún  $v > 0$ .

Las formulaciones anteriores son suficientes para analizar la dependencia lineal entre series temporales. Un enfoque más informativo para estudiar las relaciones entre las series temporales es construir un modelo multivariante para las series, porque un modelo correctamente especificado considera simultáneamente el número de series y las correlaciones cruzadas de las mismas.

En la práctica se utilizan los estimadores de las matrices antes mencionadas; en particular para  $\Gamma(l)$ :

$$\hat{\Gamma}(l) = \frac{1}{T} \sum_{t=l+1}^T (X_t - \bar{X})(X_{t-l} - \bar{X})', \quad l \geq 0$$

donde,

T: número de observaciones

$$\bar{X} = \frac{\left( \sum_{t=1}^T X_t \right)}{T} : \text{vector de medias muestrales}$$

y para  $\rho(l)$  :

$$\hat{\rho}(l) = \hat{D}^{-1} \hat{\Gamma}_X(l) \hat{D}^{-1}, \quad l \geq 0$$

donde,

$\hat{D}$  : Es la matriz diagonal de orden  $(k * k)$  que contiene las desviaciones estándar muestrales del vector  $X_t$  en la diagonal.

**Ejemplo 2.1.** Se consideran tres series de datos económicos de un país sudamericano: el producto interno bruto (PIB), denotada por  $(X_{1t})$ ; el consumo interno (CI), denotada por  $(X_{2t})$  y la demanda final interna (DFI), denotada por  $(X_{3t})$ . Se dispone de 56 datos trimestrales, desde noviembre de 2010 hasta junio de 2015 (Ver Anexo D.1). Para efectos de comparaciones se trabajará únicamente con los primeros 50 datos y se guardarán los 6 restantes para comparar con predicciones posteriores (enero 2015 - junio de 2015). Se desea estimar las matrices de correlaciones cruzadas de las series.

Resolución.

Las matrices de correlación cruzada se las construyen de manera manual, considerando cada escenario de posibles combinaciones entre las variables; así, en este caso, se obtiene:

a) Estadísticos descriptivos de  $x_{1t}$ ,  $X_{2t}$  y  $X_{3t}$ .

	Media	Mediana	Máximo	Mínimo	Desv. Est.	Asimetría	Curtosis
$X_{1t}$	96,52	89,79	165,31	49,79	34,15	0,35	1,92
$X_{2t}$	78,10	73,72	126,01	43,05	24,89	0,31	1,84
$X_{3t}$	98,71	92,34	166,66	50,51	34,48	0,30	1,87

b) Matrices de correlación cruzada

	retardo 1			retardo 2			retardo 3		
$X_{1t}$	0,94	0,94	0,94	0,87	0,88	0,87	0,81	0,82	0,81
$X_{2t}$	0,93	0,94	0,94	0,87	0,88	0,87	0,80	0,81	0,81
$X_{3t}$	0,93	0,94	0,94	0,87	0,88	0,87	0,80	0,81	0,81

c) Representación simplificada

+	+	+
+	+	+
+	+	+

+	+	+
+	+	+
+	+	+

+	+	+
+	+	+
+	+	+

Tabla 2.1: Resumen de estadísticas y matrices de correlación cruzada para  $X_{1t}$ ,  $X_{2t}$  y  $X_{3t}$

Para representar a las matrices de correlación cruzada, se utiliza la forma gráfica simplificada , que utiliza el hecho que  $2/\sqrt{T}$  (0,28 en este caso) es el valor crítico de la correlación muestral con nivel de significación del 5 %, bajo la suposición que  $X_t$  es un ruido blanco:

- “+” representa a los coeficientes de correlación que son mayores o iguales a  $2/\sqrt{T}$ .
- “-” representa a los coeficientes de correlación que son menores o iguales que  $-2/\sqrt{T}$ .
- “.” Representa a los coeficientes que se encuentran entre a  $-2/\sqrt{T}$  y  $2/\sqrt{T}$ .

Es fácil ver que las correlaciones cruzadas son significativas en los primeros tres retardos. En algunos paquetes estadísticos se puede encontrar el cálculo de las matrices de correlación cruzada. En Eviews, por ejemplo, se presenta la siguiente salida:

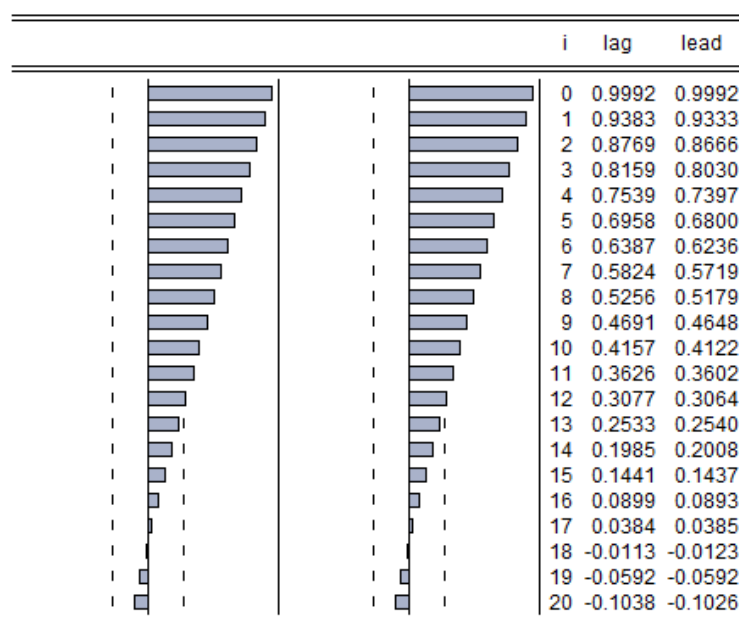


Figura 2.1: Correlaciones cruzadas entre  $X_{1t}$  y  $X_{2t}$

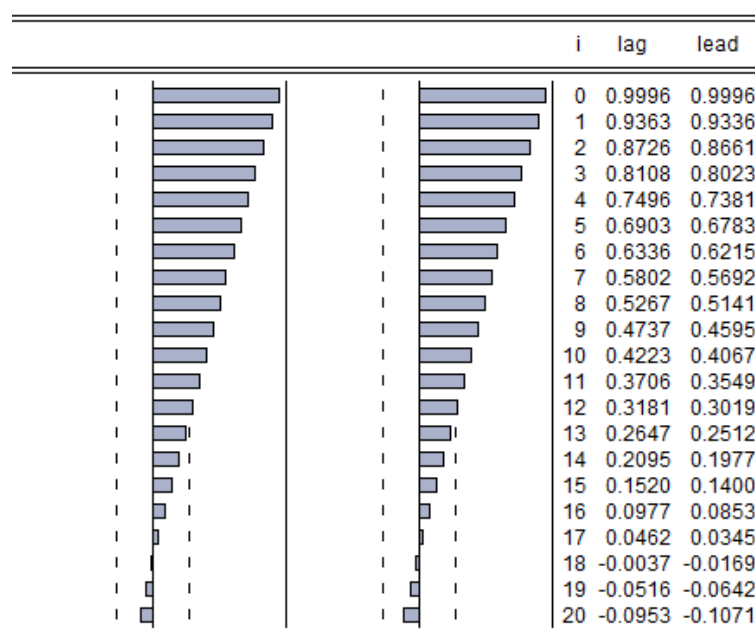
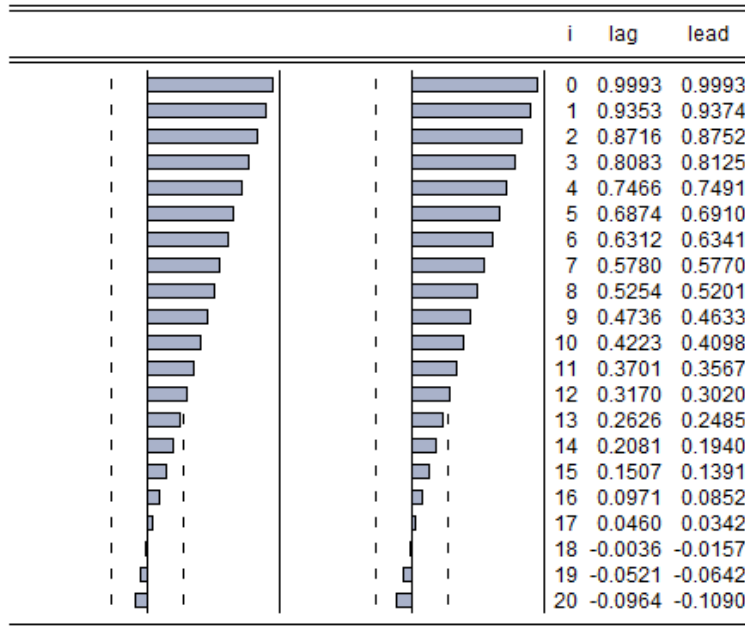


Figura 2.2: Correlaciones cruzadas entre  $X_{1t}$  y  $X_{3t}$





**Figura 2.3:** Correlaciones cruzadas entre  $X_{2t}$  y  $X_{3t}$

Como se puede observar, los valores calculados por el paquete son aquellos que están en la diagonal secundaria de las matrices calculadas manualmente. Para poder completar la matriz, se puede, ver a partir de la fórmula de cálculo que las  $\hat{\rho}_{ii}(l)$  corresponden a las autocorrelaciones simples de orden  $l$  de cada serie univariante dentro de  $X_t$ .

## 2.3 Modelos de Vectores Autoregresivos (VAR)

Este tipo de modelos no pertenecen a los modelos estocásticos desarrollados por Box y Jenkins; sin embargo, la representación VAR se puede considerar como la generalización de los modelos autoregresivos al caso multivariante.

### 2.3.1 El caso bivariante

Una representación VAR bivariante es aquella que consideran dos variables  $X_{1t}$  y  $X_{2t}$ . Cada una de ellas se expresa en función de sus propios valores del pasado y de los del presente y del pasado de la otra variable. Por ejemplo, se va a representar el modelo VAR bivariante de orden  $p = 3$  [VAR (3)]; se escribe:

$$X_{1t} = v_1 + \sum_{i=1}^3 b_{1i}X_{1t-i} + \sum_{i=1}^3 c_{1i}X_{2t-i} - d_1X_{2t} + u_{1t}$$

$$X_{2t} = v_2 + \sum_{i=1}^3 b_{2i}X_{1t-i} + \sum_{i=1}^3 c_{2i}X_{2t-i} - d_2X_{1t} + u_{2t}$$

Las variables  $X_{1t}$  y  $X_{2t}$  son estacionarias; las perturbaciones  $u_{1t}$  y  $u_{2t}$  son ruidos blancos de varianzas constantes y no correlacionados. Se puede ver inmediatamente la gran can-

tividad de parámetros a estimar (aquí 16 coeficientes), con los problemas típicos de pérdida de grados de libertad. Hay que tomar en cuenta que  $X_{1t}$  tiene un efecto inmediato en  $X_{2t}$  y recíprocamente. Este sistema inicial se denomina *forma estructural* de la representación VAR. Su *forma matricial*, se expresa como:

$$BX_t = v + \sum_{i=1}^3 \tilde{A}_i X_{t-i} + u_t$$

con:

$$B = \begin{bmatrix} 1 & d_1 \\ d_2 & 1 \end{bmatrix} \quad X_t = \begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \quad \tilde{A}_i = \begin{bmatrix} b_{1i} & c_{1i} \\ b_{2i} & c_{2i} \end{bmatrix} \quad u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

Para obtener la *forma estándar* de un modelo VAR, se multiplica la ecuación anterior por  $B^{-1}$  (que se supone existe); es decir, se expresa por:

$$X_{1t} = v_1^0 + \sum_{i=1}^3 a_{1i}^1 X_{1t-i} + \sum_{i=1}^3 a_{1i}^2 X_{2t-i} + \vartheta_{1t}$$

$$X_{2t} = v_2^0 + \sum_{i=1}^3 a_{2i}^1 X_{1t-i} + \sum_{i=1}^3 a_{2i}^2 X_{2t-i} + \vartheta_{2t}$$

En esta especificación, los errores  $\vartheta_{1t}$  y  $\vartheta_{2t}$  son funciones de las innovaciones  $u_{1t}$  y  $u_{2t}$ ; en efecto, de  $\vartheta = B^{-1}u$ , se obtiene:

$$\vartheta_{1t} = \frac{(u_{1t} - d_1 u_{2t})}{(1 - d_1 d_2)} \text{ y } \vartheta_{2t} = \frac{(u_{2t} - d_2 u_{1t})}{(1 - d_1 d_2)}$$

Se puede ver que:

$$E(\vartheta_{1t}) = 0; \quad E(\vartheta_{2t}) = 0; \quad E(\vartheta_{1t} \vartheta_{1t-i}) = 0; \quad E(\vartheta_{2t} \vartheta_{2t-i}) = 0$$

Por lo tanto, los elementos de cada familia de errores tienen esperanza nula y son no correlacionados. Además:

$$E(\vartheta_{1t}^2) = \frac{(\sigma_{u_1}^2 + d_1^2 \sigma_{u_2}^2)}{(1 - d_1 d_2)^2}; \quad E(\vartheta_{2t}^2) = \frac{(\sigma_{u_2}^2 + d_2^2 \sigma_{u_1}^2)}{(1 - d_1 d_2)^2}$$

Donde  $\sigma_{u_1}^2$  y  $\sigma_{u_2}^2$  son las varianzas de  $u_1$  y  $u_2$ , respectivamente. Así, la varianza de los errores es constante (independiente del tiempo). Además:

$$E(\vartheta_{1t} \vartheta_{2t}) = -\frac{(d_2 \sigma_{u_1}^2 + d_1 \sigma_{u_2}^2)}{(1 - d_1 d_2)^2}$$

Si  $d_1 = d_2 = 0$ , las variables  $X_{1t}$  y  $X_{2t}$  no tienen ninguna influencia sincrónica entre sí, pues los errores  $\vartheta_{1t}$  y  $\vartheta_{2t}$  serían no correlacionados. En caso contrario, los errores  $\vartheta_{1t}$  y  $\vartheta_{2t}$  estarían correlacionados y por tanto, una variación de uno de estos errores en un instante dado tiene impacto en el otro.

**Proposición 2.1.** *El modelo VAR no permite distinguir entre variables endógenas (variables propias del fenómeno estudiado) y exógenas (variables externas que ayudan a explicar las*

variables endógenas).

### 2.3.2 Representación general de un VAR

**Notación.** Un modelo VAR a  $k$  variables con  $p$  retardos se denota  $VAR(p)$ .

La generalización de la representación VAR a  $k$  variables con  $p$  retardos se escribe en su forma estándar como:

$$X_t = v_0 + A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + u_t$$

donde,

$$X_t = \begin{bmatrix} X_{1,t} \\ X_{2,t} \\ \vdots \\ X_{k,t} \end{bmatrix}; \quad v_0 = \begin{bmatrix} v_1^0 \\ v_2^0 \\ \vdots \\ v_k^0 \end{bmatrix}; \quad A_i = \begin{bmatrix} a_{1i}^1 & \dots & a_{1i}^k \\ \vdots & \ddots & \vdots \\ a_{ki}^1 & \dots & a_{ki}^k \end{bmatrix}; \quad u_t = \begin{bmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{kt} \end{bmatrix}$$

$u_t$  es el vector compuesto por los ruidos blancos de cada una de las  $k$  ecuaciones del modelo.

Se denota por:  $\Sigma_u = E(u_t u_t')$ , la matriz desconocida, de dimensión  $k$ , de varianzas-covarianzas de los errores.

Esta representación puede escribirse mediante el operador de retardo  $B$ , como:

$$(I - A_1 B - A_2 B^2 - \dots - A_p B^p) X_t = v_0 + u_t, \quad \text{o también:} \quad A(B) X_t = v_0 + u_t$$

donde, el operador de retardo  $B$  se define de la siguiente manera:

$$B^i X_t = X_{t-i}, \quad i = 1, 2, \dots$$

$$B^0 X_t = X_t$$

#### 2.3.2.1 Estabilidad de un VAR

Considérese un modelo  $VAR(1)$ :

$$X_t = v_0 + A_1 X_{t-1} + u_t$$

Se dice que un  $VAR(1)$  es estable si todos los valores propios de  $A_1$  son de valor absoluto menor que 1; lo que se puede expresar también por:

$$\det(I_k - A_1 z) \neq 0, \quad \text{para} \quad |z| \leq 1$$

Esto implica que todas las raíces del polinomio característico están fuera del círculo unidad.

### 2.3.2.2 Representación de un proceso VAR(p) en la forma de VAR(1)

Un proceso  $VAR(p)$  se puede escribir como un proceso  $VAR(1)$  si se plantea en la siguiente forma:

$$X_t = A_0 + AX_{t-1} + U_t$$

donde,

$$X_t = \begin{pmatrix} X_t \\ X_{t-1} \\ \vdots \\ X_{t-p+1} \end{pmatrix}_{kp \times 1} \quad X_t = \begin{pmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{kt} \end{pmatrix}_{k \times 1} \quad A_0 = \begin{pmatrix} v_0 \\ 0_{k \times 1} \\ \vdots \\ 0_{k \times 1} \end{pmatrix}_{kp \times 1} \quad U_t = \begin{pmatrix} (u_t)_{k \times 1} \\ 0_{k \times 1} \\ \vdots \\ 0_{k \times 1} \end{pmatrix}_{kp \times 1}$$

$$A = \begin{pmatrix} A_1 & A_2 & \dots & A_{p-1} & A_p \\ I_k & 0 & \dots & 0 & 0 \\ 0 & I_k & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & I_k & 0 \end{pmatrix}_{kp \times kp}$$

Lo importante de esta representación es que para obtener las propiedades de los procesos VAR, es suficiente con probarlas para una  $VAR(1)$

### 2.3.2.3 Procesos VAR(p) estables

Se dice que un  $VAR(p)$  es estable si:

$$\det(I_{kp} - Az) \neq 0, \text{ para } |z| \leq 1$$

Además, se puede demostrar que:

$$\det(I_{kp} - Az) = \det(I - A_1z - A_2z^2 - \dots - A_pz^p)$$

**Observación 2.2.** Se puede demostrar que si un proceso  $VAR(1)$  es estable, entonces es estacionario.

En general, se puede demostrar que un proceso  $VAR(p)$  es estacionario si el polinomio definido a partir de la expresión:  $\det(I - A_1z - A_2z^2 - \dots - A_pz^p)$  tiene sus raíces fuera del círculo unidad del plano complejo; es decir:

$$\det(I - A_1z - A_2z^2 - \dots - A_pz^p) \neq 0 \forall z \text{ tal que } |z| \leq 1$$

**Ejemplo 2.2.** Determine si el siguiente modelo es estacionario.

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} 2 \\ 5 \end{bmatrix} + \begin{bmatrix} 0,8 & 0,9 \\ 0,7 & 0,7 \end{bmatrix} \begin{bmatrix} X_{1t-1} \\ X_{2t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

**Resolución.**

Se tiene que:

$$\det \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0,8 & 0,9 \\ 0,7 & 0,7 \end{bmatrix} z \right) = 1 - 1,5z - 0,07z^2, \quad \text{entonces} \quad \begin{aligned} z_1 &= -16,17 \\ z_2 &= -15,97 \end{aligned}$$

Las dos raíces son superiores a 1 en valor absoluto; por lo tanto, el proceso es estable; lo que implica que es estacionario.

## 2.4 Representación VARMA de una Serie Multivariante

### 2.4.1 La representación VMA

Un modelo media móvil vectorial de orden  $q$  ( $VMA(q)$  por sus siglas en inglés), tiene la siguiente forma:

$$X_t = m_0 + u_t - M_1 u_{t-1} - \dots - M_q u_{t-q} \quad \text{o} \quad X_t = m_0 + M(B)u_t$$

donde,

$m_0$  : Es un vector de dimensión  $k$  constante.

$M_i$  : Son matrices de dimensión  $k * k$ .

$M(B) = I - M_1 B - \dots - M_q B^q$  es el polinomio matriz MA en términos del operador de retardo  $B$ .

$\{u_t\}_{t \in \mathbb{Z}}$  : Es un ruido blanco multidimensional.

De manera similar al caso univariante, los procesos  $VMA(q)$  son débilmente estacionarios, siempre que la matriz de covarianzas ( $\Sigma_u$ ) de  $u_t$  exista. Si se toma la esperanza de  $X_t$ , se tiene:

$$\mu = E(X_t) = m_0$$

Así, el vector constante  $m_0$  es el vector media de  $X_t$  para un modelo VMA.

Se define  $\tilde{X}_t = X_t - m_0$  como el proceso corregido en media  $VAR(q)$ . Cuando se tiene un proceso  $VMA(q)$  y considerando el hecho de que los  $\{u_t\}$  no están correlacionados, se obtiene:

1.  $Cov(\tilde{X}_t, u_t) = \Sigma_u$
2.  $\Gamma_0 = \Sigma_u + M_1 \Sigma_u M_1' + \dots + M_q \Sigma_u M_q'$
3.  $\Gamma_l = 0$  si  $l > q$
4.  $\Gamma_l = \sum_{j=l}^q M_j \Sigma_u M_{j-l}'$  si  $1 \leq l \leq q$ , donde  $M_0 = -I$

Dado que  $\Gamma_l = 0$  para  $l > q$ , las matrices de correlación cruzada de un proceso  $VMA(q)$  satisfacen:

$$\rho_l = 0, \quad l > q$$

### 2.4.2 Representación lineal de un VAR(p)

Cuando se analizaron las series temporales univariantes, se mostró que bajo ciertas condiciones un proceso  $AR(1)$  se puede representar como un proceso lineal. De la misma manera,

para las series multivariantes se puede representar, en particular, un  $VAR(1)$  como un proceso lineal (se dice que es la representación lineal del proceso). Un modelo con esta forma permite medir el impacto en los valores presentes de una variación de innovaciones (o choques).

Sea  $X_t$  un  $VAR(1)$  estable:

$$X_t = v_0 + A_1 X_{t-1} + u_t$$

Si se realizan sustituciones repetidas en el proceso hasta el  $i$ -ésimo paso, se obtiene:

$$\begin{aligned} X_t &= v_0 + A_1 (v_0 + A_1 X_{t-2} + u_{t-1}) + u_t = (I + A_1) v_0 + A_1^2 X_{t-2} + (A_1 u_{t-1} + u_t) \\ X_t &= v_0 + A_1 \left( (I + A_1) v_0 + A_1^2 X_{t-3} + A_1 u_{t-2} + u_{t-1} \right) + u_t \\ &= (I + A_1 + A_1^2) v_0 + A_1^3 X_{t-3} + (A_1^2 u_{t-2} + A_1 u_{t-1} + u_t) \\ &\vdots \\ X_t &= (I + A_1 + \dots + A_1^i) v_0 + A_1^{i+1} X_{t-i} + \sum_{j=0}^i A_1^j u_{t-j}, \quad i = 0, 1, 2, \dots \end{aligned}$$

Por definición,  $A^0 = I$ .

Como el  $VAR$  es estable, se cumple que:

$$(I + A_1 + \dots + A_1^i) v_0 \rightarrow (I - A_1)^{-1} v_0 \quad \text{si } i \rightarrow \infty$$

Además,  $A_1^{i+1} \rightarrow 0$  rápidamente; así, se lo puede ignorar. Por lo tanto, se obtiene la siguiente representación:

$$X_t = (I - A_1)^{-1} v_0 + \sum_{i=0}^{\infty} A_1^i u_{t-i}$$

La generalización a un proceso  $VAR(p)$  se la realiza aplicando la representación de un  $VAR(p)$  como un  $VAR(1)$ . Así, se obtiene:

$$X_t = \mu + \sum_{i=0}^{\infty} M_i u_{t-i}$$

donde,

$$\begin{aligned} \mu &= (I - A_1 - A_2 - \dots - A_p)^{-1} v_0 \\ M_i &= \sum_{j=1}^{\min(p,i)} A_j M_{i-j} \quad i = 1, 2, \dots \quad \text{y} \quad M_0 = I \end{aligned}$$

Las matrices  $M_i$  aparecen como un “*factor de impacto*”, a través de las cuales se analiza el efecto de un choque a lo largo de todo el proceso.

### Observación 2.3.

1. Así, se obtiene que si un proceso  $VAR(p)$  es estable, tiene una representación lineal estacionaria.
2. No se profundiza sobre la modelación  $VMA(q)$  porque no está implementada en los

programas comerciales usuales.

**Ejemplo 2.3.** Considérese el proceso VMA (1):

$$X_t = \mu + u_t - M_1 u_{t-1} = \mu + u_t - M u_{t-1}$$

donde, por simplicidad, se ha quitado el subíndice de  $M_1$ . Este modelo puede escribirse explícitamente como:

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} - \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix}$$

Se dice que la serie de retardos ( $X_t$ ) solo depende del presente y del pasado de  $\{u_t\}$ . Por lo tanto, el modelo es de memoria finita.

El parámetro  $m_{12}$  denota la dependencia lineal de  $X_{1t}$  con  $u_{2,t-1}$  en la presencia de  $u_{1,t-1}$ . Si  $m_{12} = 0$ ,  $X_{1t}$  no depende de los retardos de  $u_{2t}$  y, entonces tampoco, de los retardos de  $X_{2t}$ . De manera similar, si  $m_{21} = 0$ ,  $X_{2t}$  no depende de los valores pasados de  $X_{1t}$ . Los elementos fuera de la diagonal de  $M$  muestran la dependencia entre las componentes de las series.

Para este ejemplo, se pueden clasificar las relaciones entre  $X_{1t}$  y  $X_{2t}$  así:

1. Son series desacopladas si  $m_{12} = m_{21} = 0$ .
2. Hay una relación dinámica unidireccional de  $X_{1t}$  sobre  $X_{2t}$  si  $m_{12} = 0$ , pero  $m_{21} \neq 0$  y viceversa.
3. Hay una relación de retroalimentación entre  $X_{1t}$  y  $X_{2t}$  si  $m_{12} \neq 0$  y  $m_{21} \neq 0$ .

Finalmente, la correlación actual entre los  $m_{ij}$  (coeficientes estimados para el modelo VMA) es la misma que entre los  $u_{it}$ . La descripción previa se puede generalizar para un modelo  $VMA(q)$ .

### 2.4.3 La representación VARMA

La representación VAR puede generalizarse (es una aplicación multivariante del teorema de descomposición de Wold (1954)), por analogía con los procesos  $ARMA(pq)$ .

$$X_t = A_0 + A_1 X_{t-1} + A_2 X_{t-2} + \dots + A_p X_{t-p} + u_t + M_1 u_{t-1} + M_2 u_{t-2} + \dots + M_q u_{t-q}$$

Se trata de un proceso ARMA multivariante que se denota: VARMA.

Las condiciones de estacionariedad son análogas a las de un proceso ARMA univariante:

- Un proceso VAR es siempre invertible; es lineal (por ende estacionario) cuando es estable.
- Un proceso VMA es siempre estacionario. Es invertible si las raíces del polinomio característico asociado a  $M(z)$  están fuera del círculo unitario complejo.

- Las condiciones de estacionariedad e invertibilidad de un VARMA están dadas, respectivamente, por la parte VAR y la parte VMA del VARMA.

La generalización de los modelos ARMA encuentra nuevos temas que no ocurren en el desarrollo de los modelos VAR y VMA. Uno de ellos es el *problema de identificación*. A diferencia de los modelos ARMA, los modelos VARMA pueden no estar definidos de manera única.

Considere un modelo bivalente VMA(1):

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} - \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix}$$

Es idéntico al modelo bivalente VAR(1):

$$\begin{bmatrix} X_{1t} \\ X_{2t} \end{bmatrix} - \begin{bmatrix} 0 & -2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} = \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}$$

La equivalencia de los modelos se puede examinar fácilmente componente a componente. Es decir, para el modelo VMA(1) se tiene:

$$X_{1t} = u_{1t} - 2u_{2,t-1} \quad y \quad X_{2t} = u_{2t}$$

Por otro lado, para el modelo VAR(1) se tiene:

$$X_{1t} + 2X_{2,t-1} = u_{1t} \quad y \quad X_{2t} = u_{2t}$$

De los modelos se puede ver que:

$$X_{2,t-1} = u_{2,t-1}$$

Luego, los modelos para  $X_{1t}$  son idénticos. Este tipo de problema de identificación es inofensivo porque cualquiera de los modelos puede ser utilizado en una aplicación real. Sin embargo, existen casos en los que esta situación si se convierte en un problema y hay que tener en cuenta muchas restricciones para poder estimar un modelo VARMA.

## 2.5 Formulación de un modelo VAR

Los parámetros de un proceso VAR pueden estimarse solamente en las series temporales estacionarias. Se conoce que muchas series pueden volverse estacionarias a través de un proceso de diferenciación (en el caso de una tendencia determinista o una estacionalidad) o a través de una transformación de las variables (por ejemplo, una transformación logarítmica) en ciertos casos con heteroscedasticidad.

### 2.5.1 Estimación

En el caso de un proceso VAR, las ecuaciones pueden estimarse por MCO independientemente una de la otra (o por un método de máxima verosimilitud).



Sea el modelo  $VAR(p)$  estimado:

$$X_t = \hat{A}_0 + \hat{A}_1 X_{t-1} + \hat{A}_2 X_{t-2} + \dots + \hat{A}_p X_{t-p} + \hat{u}_t$$

siendo,  $\hat{u}_t$  el vector de dimensión  $(k, 1)$  de componentes  $\hat{u}_{1t}, \hat{u}_{2t}, \dots, \hat{u}_{kt}$ .

Se denotará por  $\hat{\Sigma}_{u,p}$  la matriz de varianzas covarianzas estimada de los residuos del modelo. Para cualquier orden  $p$ , se define por:

$$\hat{\Sigma}_{u,p} = \frac{1}{T - kp - 1} \sum_{t=p+1}^T \hat{u}_t (\hat{u}_t)' , \quad p \geq 0$$

### 2.5.2 Determinación del número de retardos

El mayor problema que debe enfrentarse a la hora de estimar los modelos VAR es el de la determinación del número de retardos a incluir en la estimación; suele realizarse en forma cuantitativa, analizando los resultados de la estimación y comparando los resultados obtenidos entre distintos modelos alternativos, ya que no es frecuente encontrar evidencias teóricas al respecto.

Los criterios comúnmente utilizados para la selección entre modelos alternativos son el criterio informativo de Akaike (AIC), el criterio de información bayesiano (BIC), que también se conoce como el criterio de Schwarz (SC) o el criterio de Hanan-Quinn (HQ).

Para el caso de la representación VAR, estos criterios se pueden utilizar para determinar el orden  $p$  del modelo. EL proceso de selección del orden de la representación consiste en estimar todos los modelos VAR para retardos de 0 a  $p_0$  ( $p_0$  es el máximo retardo admisible por la teoría económica o por los datos disponibles y se fija de antemano). Los estadísticos AIC( $p$ ), SC( $p$ ) y HQ( $p$ ) para el caso multivariante tienen las siguientes expresiones:

$$\begin{aligned} AIC(p) &= \ln \left[ \left| \sum_u \right| \right] + \frac{2k^2 p}{T} \\ SC(p) &= \ln \left[ \left| \sum_u \right| \right] + \frac{k^2 p \ln(T)}{T} \\ HQ(p) &= \ln \left[ \left| \sum_u \right| \right] + \frac{2k^2 p \ln(\ln(T))}{T} \end{aligned}$$

donde,

$k$ = número de variables del sistema.

$T$ = número de observaciones.

$p$ = número de retardos.

$\sum_u$  = matriz de varianzas covarianzas de residuos del modelo con retardo  $p$  (fijo).

Otro criterio utilizado para determinar el retardo del modelo, es la razón de máxima verosimilitud. Para utilizar este criterio, es necesario que el vector de las innovaciones tenga una distribución normal; el logaritmo de la función de verosimilitud tiene la siguiente

expresión:

$$l = -\frac{Tp}{2} (1 + \ln(2\pi)) - \frac{T}{2} \ln \left[ \left| \hat{\Sigma}_u \right| \right]$$

# Índice alfabético

## ARCH

Definición, 3

Restricciones, 4

Dependencia lineal multivariante, 13

## EGARCH

Definición, 5

## GARCH

Definición, 4

## IGARCH

Definición, 4

## Matriz

Correlaciones cruzadas, 12

## PARCH

Definición, 6

Relaciones en retardo, 12

## TARCH

Definición, 6