
UNIVERSIDAD DE EL SALVADOR
FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA



Universidad de El Salvador
Hacia la libertad por la cultura

**“MODELOS MULTIVARIANTES DE SERIES TEMPORALES
APLICADO A TASAS DE INTERÉS BANCARIAS”**

TRABAJO DE GRADUACIÓN PRESENTADO POR:

BERDUGO RAUDA, JORGE BALTAZAR.

RODRIGUEZ SUAREZ, SALVADOR EMMANUEL.

ASESOR:

DR. JOSÉ NERYS FUNES.

PARA OPTAR AL GRADO DE:

LICENCIADO EN ESTADÍSTICA.

NOVIEMBRE DE 2009.

SAN SALVADOR

EL SALVADOR

CENTRO AMERICA

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA
ESCUELA DE MATEMÁTICA

TRABAJO DE GRADUACIÓN

**“MODELOS MULTIVARIANTES DE SERIES TEMPORALES
APLICADO A TASAS DE INTERÉS BANCARIAS”**

PRESENTADO POR:

BERDUGO RAUDA, JORGE BALTAZAR.
RODRIGUEZ SUAREZ, SALVADOR EMMANUEL.

ASESOR:

DR. JOSÉ NERYS FUNES.

CIUDAD UNIVERSITARIA, SAN SALVADOR, NOVIEMBRE DE 2009.

UNIVERSIDAD DE EL SALVADOR

RECTOR: **ING. RUFINO ANTONI QUEZADA SÁNCHEZ.**

SECRETARIO GENERAL: **LIC. DOUGLAS VLADIMIR ALFARO CHÁVEZ.**

FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICA

DECANO: **DR. RAFAEL ANTONIO GÓMEZ ESCOTO.**

SECRETARIO: **LIC. MARÍA TRINIDAD DE CASTRO.**

ESCUELA DE MATEMÁTICA

DIRECTOR (INTERINO): **LIC. CARLOS MAURICIO CANJURA LINARES.**

SECRETARIO: **LIC. ERNESTO AMÉRICO HIDALGO.**

CIUDAD UNIVERSITARIA, NOVIEMBRE DE 2009.

Contenido

1. Introducción	6
2. Objetivos	9
Capítulo I. Modelos Univariantes y Regresión Dinámica de series temporales.	10
1.1. Introducción	11
1.2. Modelos Univariantes de Series Temporales	12
1.2.1. Definiciones importantes de series temporales	13
1.2.2. Modelos ARIMA(p,d,q)	16
1.2.3. Modelos ARIMA Estacionales	20
1.3. Regresión dinámica entre variables estacionarias	21
1.3.1. Introducción	21
1.3.2. Relaciones entre dos series estacionarias	23
1.3.2.1 Función de covarianzas cruzadas	23
1.3.3. Modelo dinámico entre dos series estacionarias	27
1.3.4. Construcción de modelos de regresión dinámica en variables estacionarias	31
1.3.4.1. Identificación	31
1.3.4.2 Estimación y selección del modelo	33
1.3.5 Contrastes diagnósticos	39
1.3.5.1 Contrastes sobre los parámetros	39
1.3.5.2 Contrastes sobre los residuos	40
1.3.6 Relaciones espurias entre variables dinámicas	41
1.3.6.1 Relaciones sin dinámica	41
1.3.6.2 Perturbaciones autocorreladas	42
1.3.7 El modelo con varias variables explicativas	45
1.3.8 Predicción	46
Capitulo II. Modelos Multivariantes de Series Temporales	48
2.1. Introducción	49

2.2. Procesos vectoriales estacionarios y sus autocorrelaciones	51
2.2.1. Procesos vectoriales estacionarios	51
2.2.2. Matrices de autocorrelación simple y parcial	52
2.2.3. Estimación de las matrices simples y parciales	53
2.3. Modelos ARMA vectoriales	54
2.4. Proceso VAR (1)	56
2.4.1 Proceso VAR (p)	58
2.5. Proceso VMA (q)	58
2.6. Procesos VARMA (p,q)	59
2.7. Construcción de modelos VARMA para series estacionarias	60
2.7.1 Identificación	61
2.7.2 Estimación	62
2.7.3 Diagnósis	63
2.7.4. Contrastes en los modelos	64
2.7.4.1 Contraste de Dickey Fuller y Dickey Fuller Aumentado	64
Capitulo III. Aplicación de los modelos Univariantes, Regresión Dinámica y Multivariantes de Series Temporales a tasas de interés Bancarias	68
3.1. Introducción	69
3.2. Análisis de Series Temporales Univariante	70
3.2.1 Análisis de Estacionariedad y estacionalidad	70
3.2.2 Identificación de los posibles modelos	72
3.2.3 Estimación de los parámetros de los modelos identificados, diagnóstico de modelos y predicción	75
3.2.3.1 Ajuste de un modelo ARIMA a la serie TIBP180	76
3.2.3.2 Ajuste de un modelo ARIMA a las series restantes	79
3.3. Análisis de Modelos de Regresión Dinámica	80
3.3.1 Identificación de los posibles modelos de regresión dinámica	80
3.3.1.1 Análisis de función de correlación cruzada	80
3.3.2 Estimación de los parámetros de los modelos identificados, diagnóstico de modelos y predicción	84
3.3.2.1 Ajuste de un modelo de regresión dinámica a la serie TIBP180	84

3.4. Análisis de Modelos Multivariantes de Series Temporales	89
3.4.1 Identificación del modelo para las series temporales TIBP180 y LIBOR 180	89
3.4.2 Estimación y diagnosis de un modelo VAR para TIBP180 y LIBOR 180	92
3.4.3 Predicciones para el modelo VAR de la serie TIBP180 y LIBOR 180	97
4. Conclusiones	99
5. Recomendaciones	100
6. Bibliografía	101

1. Introducción.

El estudio de las series temporales data de 1662 cuando J. Graunt hace una interpretación de los movimientos monográficos y de la conducta social, a partir de las cantidades brutas de nacimientos y muertes por causas, en Londres. La fuente de datos utilizada fue las tablas semanales de compañía de Sacristanes Parroquiales, estas tablas contenían la cantidad de muertes por causas, y la cantidad de nacimientos ocurrida en cada parroquia londinense, durante el periodo de 1604 a 1661. Se cree que este fue el primer análisis realizado de series temporales. En los años posteriores se continúa el desarrollo de las tablas de vida, dejando abandonado el modelaje de series temporales a través de funciones matemáticas.

En 1920, W. M. Pearson propone las ideas básicas del enfoque clásico en el tratamiento de las series temporales. Este método se desarrolló para el análisis de series económicas. En 1932, Collins estima la tasa semanal de muertes esperadas como un promedio de tasas de muerte observadas para esa semana durante los años no epidémicos anteriores. En 1962, T. C. Eickoff utiliza una ecuación de autorregresión, sobre una serie de tasas semanales de muertes, con la omisión de las tasas correspondientes a las semanas epidémicas, para determinar la tasa de muertes usual esperada para cada semana del año.

Luego en los años 60 los británicos G. E. P. Box y G. Jenkins fueron los encargados de formalizar el estudio y análisis de las series temporales, ya que con sus investigaciones sobre el problema de la predicción y control de series industriales, demostraron que los métodos de alisado exponencial eran óptimos para un tipo especial de proceso ARIMA. Fruto de sus investigaciones es su célebre libro "Time Series Analysis: Forecasting and Control" que marca un hito en el análisis de las series temporales al presentar una metodología unificada para estudiar series estacionarias y no estacionarias, estacionales o no, y aplicar estos modelos en la práctica. Una de sus contribuciones fundamentales fue la introducción de los procesos integrados, que se convierten en estacionarios tomando diferencias. Además, estos autores desarrollaron los fundamentos estadísticos de los sistemas de control.

El análisis de series de tiempo se concentra en el comportamiento histórico de los datos, buscando identificar características en la secuencia que resulten consistentes y permitan su proyección al futuro mediante una función matemática. Las características que suelen estar presentes en una serie de tiempo son cuatro: tendencia, ciclos, estacionalidad y variación aleatoria. Cada uno de estos factores genera un "movimiento" característico a lo largo del tiempo.

En general, una serie temporal se refiere a un **periodo natural** que tan solo es una parte de la historia de un proceso estocástico o del que procede dicha serie. No obstante si las circunstancias sociales o naturales del período muestral al que se refiere la serie considerada se mantienen relativamente estables después de dicho período, entonces se espera que las conclusiones obtenidas del análisis sean aplicables también a momentos posteriores al menos a corto plazo.

El estudio, análisis y aplicación de series temporales es muy extenso, sin embargo en nuestro país solo se cuenta con estudios e investigaciones de series temporales univariantes. Entre los estudios de series temporales que se han realizado en el país están:

- En el 2007, se realizó la tesis de Maestría en Estadística titulada “Ajuste de un modelo ARIMA a la emanación de gas CO_2 del volcán de San Vicente”, en la cual se presentó un análisis y aplicación de series temporales univariantes utilizando la metodología Box-Jenkins a la serie **emanación de gas CO_2** . La tesis se estructuró en 3 capítulos, en el primer capítulo se abordó la descripción, depuración y análisis de los datos de emanaciones de CO_2 . En el segundo capítulo se hizo una síntesis de los diferentes modelos ARIMA univariantes y sus principales características de cada modelo. Y finalmente en el tercer capítulo se ajustó un modelo ARMA(2,2) a la emanación del gas CO_2 del volcán de San Vicente.
- En el 2008, se realizó la tesis de Licenciatura en Estadística titulada “Análisis de Series Temporales con Outliers e intervenciones y sus aplicaciones”, en la cual se presentó un análisis y aplicación de series temporales univariantes con datos atípicos (outliers) e intervenciones que es el caso donde una serie presenta movimientos bruscos o de importancia. La aplicación se realizó a la serie temporal de **Índice del Volumen de la Actividad Económica (IVAE) en Comida, Restaurantes y Hoteles**, la base de datos fue obtenida del Banco Central de Reserva consta de 206 datos con periodicidad mensual desde enero de 1990 a febrero de 2007, donde se ajustaron modelos ARIMA utilizando la metodología de Box-Jenkins, como resultado se obtuvo que el modelo que mejor se ajusta a la serie es el $IMA(1,2) \times IMA(1,1)_{12}$, luego a este modelo se le incorporó las intervenciones y se obtuvo como resultado que el modelo ARIMA con intervenciones que mejor se ajusta a la serie es $I(1) \times ARI(1,1)_{12}$.
- En el 2007, se comenzó a realizar la tesis de Licenciatura en Estadística titulada “Modelo para Series Temporales de Valores Enteros”, la cual tiene como fin modelar series de tiempo de datos enteros y pequeños, de forma análoga a las series temporales para datos continuos (ARIMA), el principal objetivo de los modelos de series temporales para valores enteros es pronosticar un suceso con base a un comportamiento de un fenómeno en estudio previamente observado. La aplicación de estos modelos se realizará en el área de salud, con una base de datos proporcionada por el Hospital Amatepec. Las series de tiempo con las que se trabajará son: **número de pacientes confirmados con la enfermedad de dengue en los años 2004 y 2005, partos atendidos en el 2007**. Actualmente la tesis sigue en proceso.
- En el 2008, se realizó la tesis de Maestría en Estadística titulada “Ajuste de un modelo ARIMA para la precipitación diaria en la zona oriental de El Salvador”, en la cual se presentó un análisis y aplicación de las series temporales univariantes a la serie de tiempo **precipitación diaria en El Salvador**. La medición de esta serie temporal contó con la ayuda del personal del servicio Meteorológico Nacional (SMN) y del Servicio Nacional de Estudios Territoriales (SNET). Las mediciones se enfocaron en los meses de abril a octubre del año 2003, en 19 estaciones en la zona oriental del país. Se ajustó un modelo ARI (1,1) a esta serie utilizando la metodología de Box-Jenkins, con análisis de datos atípicos e intervención.
- En el 2009, se publica un artículo titulado “Análisis econométrico del impacto del incremento del desempleo en EEUU sobre las remesas familiares en El Salvador”, en el cual se presenta un análisis y aplicación de los modelos de regresión dinámica donde se relacionan las series temporales **remesas**,

desempleo hispano, índice de volumen de la actividad económica (IVAE) en El Salvador. Se ajustó un modelo econométrico de evolución de las remesas familiares en El Salvador, en función de un componente tendencial y autorregresivo, de la tasa de desempleo hispano en EEUU, del IVAE y dos variables ficticias estacionales correspondientes a los meses de mayo y diciembre. Se tomaron dos escenarios en el primero la tasa de desempleo hispano de EEUU sube de 9.2% en diciembre de 2008 a 12% en diciembre de 2009, y permanece a ese nivel por dos años más, mientras que el IVAE mantiene su nivel de noviembre de 2008 hasta diciembre de 2011, los resultados obtenidos para este escenario fueron que las remesas caen 6% en 2009, 3.8% en 2010 y suben 2% en 2011. En el segundo escenario es igual al primero, excepto que se asume una reducción de 1% mensual del IVAE para el periodo diciembre 2008 a diciembre 2011, los resultados obtenidos para el segundo escenario fueron que las remesas caen 6% en 2009, 2.7% en 2010 y suben 3.9% en 2011. En este trabajo no se incluyen los modelos ARIMA, a pesar de ser los modelos más confiables para el análisis de Series Temporales Multivariantes.

En nuestra investigación haremos un estudio, análisis y sistematización de temas innovadores en el país como son los modelos de regresión dinámica y los modelos multivariantes de series temporales; se investigará la dependencia de la tasa de interés nacional de las tasas de interés internacionales, para ello, se utilizará la base de datos “Tasas de interés bancario” que contiene las variables: **FED (Tasa de interés que los bancos se cargan unos a otros por créditos overnight con fondos estatales), LIBOR-180 (London Interbanking Offered Rate: tasa promedio de interés, día a día, en el mercado interbancario de Londres), PRIME (Tasa de interés básica que aplican los principales bancos en EEUU para préstamos corporativos), TIBP-180 (Tasa de interés básica pasiva ofrecida por bancos a sus depositantes), TPH1 (Tasa de Préstamo hasta un 1 año plazo nacional), Inflación (Índice de crecimiento de precios), IVAE (Índice de volumen de la actividad económica nacional).**

El problema que se plantea es encontrar un modelo que explique el comportamiento de una serie temporal en función del resto de series temporales consideradas en el estudio. El abordaje es el siguiente:

- 1- Estudio de los modelos de series temporales unidimensionales, modelos de regresión dinámica para series temporales estacionarias y modelos vectoriales o multivariantes de series temporales.
- 2- Ajustar un modelo multivariante para las series temporales contenidas en la base de datos “Tasas de interés bancario”.

En este estudio se establece la relación entre las diferentes tasas de interés bancario, por ejemplo, ver que tanto influye la variable **LIBOR-180** (Tasa promedio de interés, día a día, en el mercado interbancario de Londres) sobre la **TIBP-180** (tasa de interés básica pasiva ofrecida por bancos a sus depositantes nacional), en otras palabras queremos ver si la tasa de interés impuesta internacionalmente por el banco de Londres influye en las tasas de interés que imponen los bancos nacionales, o si las tasas de interés son actualizadas a nivel nacional por un cambio de nivel internacional. También se realizará una aplicación de series temporales multivariantes, con el fin de generar modelos que nos puedan ayudar a definir un comportamiento y predicción de dichas series temporales en nuestro país.

2. Objetivos.

2.1. General.

- Sistematizar la teoría estadística de los modelos multivariantes de series temporales y su aplicación a las tasas de interés bancario nacional e internacional.

2.2. Específicos.

- Estudiar la teoría estadística de las series temporales univariantes y multivariantes.
- Estudiar la regresión dinámica entre variables estacionarias, así como en las no estacionarias.
- Sistematizar la teoría de la regresión dinámica y de series temporales multivariantes.
- Establecer modelos explicativos sobre cada una de las series temporales de tasas de interés bancarias nacional e internacional.
- Ajustar un modelo multivariante de series temporales, para predecir la explicación de las tasas de interés internacionales sobre las tasas de interés nacionales.

Capítulo I.

Modelos univariantes y

Regresión Dinámica de

Series Temporales.

1.1. Introducción.

En las últimas décadas hemos podido percibir la aparición de modelos y métodos estadísticos de gran importancia para el trabajo con datos de series temporales. La necesidad de disponer de predicciones lo más precisas posibles, así como el interés en conocer la dinámica de las distintas variables, ha originado que los métodos de análisis de series temporales hayan pasado a ocupar un lugar central en el estudio de disciplinas muy diversas.

La metodología actual para analizar las series temporales es en estadística el desarrollo de varias líneas de trabajo en distintos campos científicos. Los campos en el que se desarrollan las series temporales son cinco principalmente. El primero tiene sus raíces en los estudios de series astronómicas y climáticas, la cuál es la razón del origen de los procesos estocásticos estacionarios, desarrollada por los matemáticos Kolmogorov, Wiener y Cramer en la primera mitad del siglo XX. El segundo desarrollo es el método de alisado la cual se desarrolló para prever series de producción y ventas entre 1960-1970 en donde ya se aprovecharon los cálculos de las primeras computadoras. El tercero, la teoría de predicción y control de sistemas lineales, desarrollada en ingeniería en los años 70 y estimulada por el desarrollo de la ingeniería aeronáutica y espacial. El cuarto, es la teoría de procesos no estacionarios y no lineales, desarrollada por estadísticos y econométricos en los últimos 20 años del siglo XX. Finalmente, el último campo son los modelos multivariantes y métodos de reducción de la dimensión en sistemas dinámicos, que se encuentra todavía en fase de desarrollo.

Los métodos disponibles en la actualidad para el análisis de series temporales son investigaciones de matemáticos, estadísticos, ingenieros, físicos y economistas; los cuáles tuvieron su auge durante el siglo XX para resolver problemas de predicción y control de variables y sistemas dinámicos.

Box y Jenkins en su texto publicado en 1970, sistematizaron el conocimiento existente sobre procesos estacionarios y lo generalizaron permitiendo procesos evolutivos, causados por la presencia de raíces unitarias autorregresivas. Esta caracterización del nivel evolutivo de las Series Temporales es todavía la dominante en muchas ciencias, como la Economía con la Teoría Econométrica de la Cointegración. Una de las contribuciones fundamentales de Box y Jenkins fue la introducción de los procesos integrados, que se convierten en estacionarios tomando diferencias. La teoría de estos procesos se desarrolla posteriormente en las décadas de los 80 y 90.

Un trabajo pionero sobre las series temporales multivariantes fue debido a Quenouille que introdujo un método para estimar el sesgo de un estimador (posteriormente recibiría el nombre de jackknife) mediante su cálculo en conjuntos que se forman a partir de la muestra original omitiendo una observación. Quenouille con sus métodos de remuestreo y datos omitidos en series temporales desarrolló técnicas muy avanzadas en este tipo de análisis para su época. En 1980 Tiao y Box proponen usar un método general para construir modelos ARIMA vectoriales a través de los modelos de función de transferencia, como son conocidos en ingeniería, modelos econométricos dinámicos en economía o regresión dinámica, que permiten medir como se transmiten los efectos entre variables.

Este capítulo, se divide en dos partes, en la primera se presentan las definiciones necesarias e importantes para el completo entendimiento de los modelos de series

temporales univariantes, así como las propiedades más importantes para su análisis, como lo son: la función de autocorrelación, el correlograma simple, y correlograma parcial. Los modelos que estudiaremos son: **AR, MA, ARMA, y ARIMA**. Y en la segunda parte estudiaremos la relación entre series temporales estacionarias con el fin de construir modelos de regresión dinámica, donde la metodología a utilizar para la construcción de dichos modelos se dividirá en 3 etapas que son: identificación, estimación y diagnosis.

1.2. Modelos Univariantes de Series Temporales.

Introducción.

A comienzo de los años 70, G.E.P Box, profesor de Estadística de la Universidad de Wisconsin, y G.M. Jenkins, profesor de Ingeniería de Sistemas de la Universidad de Lancaster, introdujeron una pequeña revolución en el enfoque del análisis de series temporales, con el propósito de establecer mejores mecanismos de pronóstico y control. Su libro (1976) se convirtió rápidamente en un clásico, y sus procedimientos se utilizan ampliamente desde entonces en diferentes ramas de la ciencia, conociéndose como modelos ARIMA y también como **modelos Box-Jenkins**.

Para este tipo de modelos, el primer paso consiste en convertir nuestra serie de observaciones en una serie estacionaria, que es aquella en la que ni la media, ni la varianza, ni las autocorrelaciones dependen del tiempo. Una vez “estabilizada” la serie mediante las transformaciones adecuadas, se procede a estudiar la presencia de regularidades en la serie, para identificar un posible modelo matemático. Para ello se calculan la función de autocorrelación simple y parcial, y se compara su forma con un catalogo de patrones gráficos, que son típicos de los diferentes modelos propuestos, seleccionando el modelo que más se adecue a la forma de las funciones de autocorrelación que hemos obtenido con nuestros datos.

Una vez elegida la forma del modelo, se estiman los coeficientes del mismo, y finalmente se procede a efectuar un análisis de los residuos (diferencia entre el valor realmente observado y el valor previsto por el modelo), con el fin de comprobar si el ajuste del modelo a nuestros datos es adecuado. Si no lo fuera repetimos el proceso buscando otros modelos, pero si el modelo supera la etapa de validación se comprueba si los pronósticos son adecuados con nuestros datos, volviendo a comenzar la búsqueda de otro modelo si no fuera el caso. Puede por tanto tratarse de un proceso iterativo de mejora del modelo.

En el modelo, cada valor tomado por la variable en un instante dado, esta influido por los valores de la variable en momentos anteriores, y se expresa como una relación lineal, función de:

1. Valores recientes de la variable.
2. Ruidos en valores recientes de la variable.
3. Valores remotos de la variable.
4. Ruidos en valores remotos de la variable.

El esquema general de un modelo ARMA es:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q}$$

Los modelos ARMA, están constituidos por una combinación de p términos AR (proceso autorregresivo), y q términos MA (proceso de medias móviles). La parte AR modela la influencia de los valores anteriores de la serie (y_{t-1} hacia atrás), y la parte MA modela la influencia del ruido en valores anteriores de la serie (u_{t-1} hacia atrás), junto con el término u_t que corresponde al ruido esperado en el mismo momento t en el que se estima el nuevo valor de la variable Y .

Una de las ventajas de estos modelos es su gran simplicidad (suma de términos), frente a los modelos propuestos en la formulación clásica.

La letra **I** que aparece en el nombre del modelo completo **ARIMA**, corresponde al proceso último a realizar, una vez definido el tipo de modelo y estimados los coeficientes de éste, ya que entonces hay que restablecer las características originales de la serie de datos, que fue transformada para inducir estacionariedad. A ese proceso inverso se denomina en general **Integración** y aporta esa letra que completa el nombre.

1.2.1. Definiciones relacionadas con series temporales.

Procesos Estocásticos¹.

Un **proceso estocástico** es una **secuencia de variables aleatorias**, ordenadas y equidistantes cronológicamente, referidas a una (proceso univariante o escalar) o a varias (proceso multivariante o vectorial) características de una unidad observable en diferentes momentos. Algunas representaciones son las siguientes $\dots, Y_{-1}, Y_0, Y_1, Y_2, \dots$; $(Y_t : t = 0 \pm 1, \pm 2, \dots)$; (Y_t) , donde Y_t es una variable aleatoria escalar referida a la unidad observable considerada en el momento t .

Procesos estocásticos estacionarios².

Un proceso estocástico (Y_t) es estacionario cuando las propiedades estadísticas de cualquier secuencia finita $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ ($n \geq 1$) de componentes (Y_t) son semejantes a las de la secuencia $Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}$ para cualquier número entero $h = \pm 1, \pm 2, \dots$

Existen dos tipos de estacionariedad, **la estacionariedad en sentido estricto o fuerte** y **estacionariedad en sentido amplio o débil**.

¹ Mauricio, José Alberto. "Introducción al Análisis de Series Temporales", 2007, Universidad Complutense Madrid.

² Mauricio, José Alberto. "Introducción al Análisis de Series Temporales", 2007, Universidad Complutense Madrid.

- **Estacionariedad en sentido estricto o fuerte:**

Diremos que un proceso es estacionario (en sentido estricto) cuando su función de distribución conjunta es invariante respecto de un desplazamiento en el tiempo.

- **Estacionariedad en sentido amplio o débil:**

A partir del concepto de estacionariedad en sentido estricto, se derivan tres propiedades. Estas son:

1) Las esperanzas matemáticas de las variables aleatorias no dependen del tiempo; es decir son constantes:

$$E(y_t) = E(y_{t+k}) \quad \forall m$$

o bien:

$$\mu_t = \mu \quad \forall t.$$

2) Las varianzas tampoco dependen del tiempo y son finitas, es decir:

$$v(y_t) = v(y_{t+k}) < \infty \quad \forall k$$

o bien:

$$\sigma_t^2 = \sigma^2 \quad \forall t.$$

3) Las covarianzas entre dos periodos de tiempo distintos solamente dependen del lapso de tiempo transcurrido entre estos dos periodos. Esto es:

$$cov(y_t, y_{t'}) = cov(y_{t+k}, y_{t'+k}) < \infty \quad \forall k$$

Procesos de ruido blanco³.

El proceso de ruido blanco es un proceso estacionario muy importante el cual cumple con tres condiciones que son:

a) Su esperanza es siempre constante e igual a cero: $E[y_t] = 0$, $t = 1, 2, \dots$

b) Su varianza es constante: $Var(y_t) = \sigma^2$, $t = 1, 2, \dots$

c) Las variables del proceso están incorreladas para todos los retardos:

$$Cov(y_t, y_{t-k}) = 0, \quad k = \pm 1, \pm 2, \dots$$

³ Peña, Daniel. "Análisis de Series temporales", 2005, Alianza Editorial.

Si este proceso además de cumplir con las 3 condiciones anteriores y cada variable aleatoria está distribuida normalmente se le llama proceso de **ruido blanco gaussiano**.

Procesos estocásticos no estacionarios⁴.

Un proceso estocástico (Y_t) es **no estacionario** cuando las propiedades estadísticas de al menos una secuencia finita $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$ ($n \geq 1$) de componentes de (Y_t) , son diferentes de la secuencia $Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}$ para al menos un número entero $h > 0$.

En general, una condición suficiente para que un proceso sea no estacionario es que la esperanza marginal de algunos de sus componentes sea distinta de la de otros.

Serie temporal⁵.

Una **serie temporal** es una secuencia de N observaciones (datos) ordenadas y equidistantes cronológicamente sobre una característica (serie **univariante** o **escalar**) o sobre varias características (serie **multivariante** o **vectorial**) de una unidad observable en diferentes momentos.

Estimadores de las autocovarianzas y autocorrelaciones.

Si la media del proceso fuese conocida, el estimador de las autocovarianzas de orden h es:

$$\tilde{\gamma}_h = \frac{1}{T-h} \sum_{i=1}^T (y_t - \mu)(y_{t-h} - \mu)$$

Sin embargo, cuando μ es desconocida y se sustituye en la fórmula anterior por su estimador, \bar{y} , el estimador resultante no es centrado. Un estimador alternativo, que tiene mejores propiedades es:

$$\hat{\gamma}_h = \frac{1}{T} \sum_{t=k+1}^T (y_t - \bar{y})(y_{t-h} - \bar{y}),$$

Aunque también es un estimador sesgado de la autocovarianza poblacional, tiene menor error cuadrático de estimación que el anterior. En particular la varianza del proceso se estima mediante $\hat{\gamma}_0$.

⁴ Mauricio, José Alberto. "Introducción al Análisis de Series Temporales", 2007, Universidad Complutense Madrid.

⁵ Mauricio, José Alberto. "Introducción al Análisis de Series Temporales", 2007, Universidad Complutense Madrid.

Autocorrelación parcial⁶.

El concepto de autocorrelación parcial entre y_t e y_{t-k} es similar al de coeficiente de autocorrelación entre dichos valores de la serie temporal, salvo porque mide dicha correlación ajustada por el efecto de los retardos intermedios. Es decir al obtener la correlación entre y_t e y_{t-k} se elimina la influencia que, sobre ambas, tienen y_{t-1} , $y_{t-2}, \dots, y_{t-k+1}$.

La función de autocorrelación (FAS) y la función de autocorrelación parcial (FAP) son herramientas gráficas de mucha importancia en el análisis univariante de series temporales ya que a partir de ellas se puede concebir en primera instancia la idea de un modelo específico.

1.2.2. Modelos ARIMA(p,d,q)⁷.

Los modelos *ARIMA* están constituidos por una combinación de p términos *AR* (proceso autorregresivo), y q términos *MA* (proceso de medias móviles). Donde la parte *AR* modela influencia de los valores anteriores de la serie (y_{t-1} hacia atrás) y la parte *MA* modela la influencia del ruido en valores anteriores de la serie (u_{t-1} hacia atrás), junto con el término u_t que corresponde al ruido esperado en el mismo momento t en el que se estima el nuevo valor de la variable y_t . Una de las ventajas de estos modelos es su gran simplicidad (sumas de términos), frente a los modelos propuestos en la formulación clásica.

Un modelo mixto que tiene parte autorregresiva y parte de medias móviles *ARMA(p,q)* se representa como:

$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \delta + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q}$, puede expresarse simplifícadamente, usando los polinomios de retardo como: $\phi(L)y_t = \delta + \theta(L)u_t$. Dado que este modelo contiene una parte *AR* y una parte *MA*, deberá cumplir las condiciones de estacionariedad como de invertibilidad. O sea que diremos que es estacionario si su parte *AR* lo es y diremos que es invertible si su parte *MA* lo es.

Este modelo es estacionario e invertible si tanto las raíces de $\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0$ como las de $\theta(L) = 1 - \theta_1 L - \dots - \theta_q L^q = 0$ caen fuera del círculo unitario.

Una condición necesaria pero no suficiente, para que un modelo mixto autorregresivo-medias móviles de cualquier orden sea estacionario e invertible es la siguiente:

⁶ Peña, Daniel. "Análisis de Series temporales", 2005, Alianza Editorial.

⁷ Peña, Daniel. "Análisis de Series temporales", 2005, Alianza Editorial.

$$\theta_1 + \theta_2 + \dots + \theta_q < 1$$

$$\phi_1 + \phi_2 + \dots + \phi_p < 1$$

Adicionalmente a la estacionariedad e invertibilidad supondremos que las p raíces de $\phi(L) = 0$ y las raíces de $\theta(L) = 0$ no se cancelan, esto es, no hay raíces comunes. Bajo estas condiciones, la esperanza matemática y la varianza del modelo $ARMA(p,q)$ se definen así:

Esperanza matemática:

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p}$$

Varianza:

$$\gamma_0 = \frac{\sigma_u^2(1 + \theta_1^2 + \dots + \theta_q^2 - 2\phi_2\theta_2 - \dots - 2\phi_h\theta_h)}{1 - \phi_1^2 - \dots - \phi_p^2}$$

Función de autocovarianzas:

$$\gamma_j = \begin{cases} \phi_1\gamma_0 + \phi_2\gamma_1 + \dots + \phi_p\gamma_{p-1} - \theta_1\sigma_u^2 - \theta_2E(\tilde{y}_{t-1}u_{t-2}) - \dots - \theta_qE(\tilde{y}_{t-1}u_{t-q}) & \text{para } j = 1 \\ \phi_1\gamma_1 + \phi_2\gamma_0 + \dots + \phi_p\gamma_{p-2} - \theta_2\sigma_u^2 - \theta_3E(\tilde{y}_{t-1}u_{t-3}) - \dots - \theta_qE(\tilde{y}_{t-1}u_{t-q}) & \text{para } j = 2 \\ \dots\dots\dots & \\ \phi_1\gamma_{q-1} + \phi_2\gamma_{q-2} + \dots + \phi_p\gamma_{p-q} - \theta_q\sigma_u^2 & \text{para } j = q \\ \phi_1\gamma_{j-1} + \phi_2\gamma_{j-2} + \dots + \phi_p\gamma_{p-q} & \text{para } j > q \end{cases}$$

Función de autocorrelaciones:

$$\rho_j = \begin{cases} \phi_1 + \phi_2\rho_1 + \dots + \phi_p\rho_{p-1} + \frac{1}{\gamma_0}[-\theta_1\sigma_u^2 - \theta_2E(\tilde{y}_{t-2}u_{t-2}) - \dots - \theta_qE(\tilde{y}_{t-1}u_{t-q})] & \text{para } j = 1 \\ \phi_1\rho_1 + \phi_2\rho_2 + \dots + \phi_p\rho_{p-2} + \frac{1}{\gamma_0}[-\theta_2\sigma_u^2 - \theta_3E(\tilde{y}_{t-2}u_{t-3}) - \dots - \theta_qE(\tilde{y}_{t-2}u_{t-q})] & \text{para } j = 2 \\ \dots\dots\dots & \\ \phi_1\rho_{q-1} + \phi_2\rho_{q-2} + \dots + \phi_p\rho_{p-q} + \frac{1}{\gamma_0}\theta_q\sigma_u^2 & \text{para } j = q \\ \phi_1\rho_{j-1} + \phi_2\rho_{j-2} + \dots + \phi_p\rho_{j-q} & \text{para } j > q \end{cases}$$

Propiedades del modelo $ARMA(p,q)$:

Función de autocorrelación: La función de autocorrelación no se anula, si bien tiende a cero a medida que aumentan los desfases temporales considerados. Consecuentemente, la memoria de este modelo es finita. La presencia del término de medias móviles en el modelo $ARMA(p,q)$ afecta la determinación de las q primeras autocorrelaciones $(\rho_1, \rho_2, \dots, \rho_q)$ estando el resto de autocorrelaciones afectadas únicamente por la parte de medias móviles del modelo y, a continuación, su forma coincide con los del modelo $AR(p)$, es decir, un comportamiento amortiguado hacia cero sin llegar a anularse, si las raíces reales, o comportamiento sinusoidal, pero también amortiguado hacia cero, si las raíces son complejas.

Correlograma Simple: la función de autocorrelación parcial no se anula, como consecuencia de que el modelo $ARMA(p,q)$ contiene el modelo $AR(p)$ como un caso especial, y la función de autocorrelación parcial del modelo $AR(p)$ nunca se anula. La representación gráfica de la autocorrelación, tendrá q autocorrelaciones, en los que influye tanto la parte autorregresiva como el de las medias móviles del modelo, y después de las q autocorrelaciones solo influye el modelo AR.

Correlograma Parcial: la función de autocorrelación parcial no se anula, como consecuencia de que el modelo $ARMA(p,q)$ contiene el modelo $MA(q)$ como un caso especial. La representación gráfica de la función de autocorrelación parcial tendrá también un comportamiento amortiguado hacia cero.

Los modelos $ARMA$ permiten una máxima flexibilidad de representación con muy pocos parámetros y rara vez los órdenes p y q sobrepasan el valor de 2.

Modelos $ARIMA$:

La letra **I** que aparece en el nombre completo **ARIMA** corresponde al proceso último a realizar, una vez definido el tipo del modelo y estimados los coeficientes de éste, ya que entonces hay que restablecer las características originales de la serie de datos que fue transformada para inducir la estacionariedad. A ese proceso inverso se denomina en general Integración y aporta esa letra que completa el nombre.

Dada una serie temporal que no cumple con los requisitos de estacionariedad en media y en varianza. Para obtener un modelo $ARIMA$ será necesario aplicar ciertas transformaciones para convertir dicha serie en estacionaria.

Estas transformaciones se concretan en diferenciaciones sucesivas en la serie (la determinación del parámetro d).

Si diferenciamos una vez y_t , se obtiene un modelo transformado de la serie:

$$w_t = y_t - y_{t-1}$$

donde:

$$w_t = \Delta y_t = y_t - y_{t-1} = (1-L)y_t; \text{ donde } Ly_t = y_{t-1}.$$

Al diferenciar la serie temporal se pasa de y_t al proceso w_t . Ahora considérese el caso inverso de obtener y_t a partir del proceso w_t . Por sustituciones sucesivas se obtiene que

$$y_t = w_t + y_{t-1} = w_t + w_{t-1} + y_{t-2} = w_t + w_{t-1} + w_{t-2} + y_{t-3} = \dots = w_t + w_{t-1} + w_{t-2} + w_{t-3} \dots$$

obteniéndose así:

$$y_t = w_t + w_{t-1} + w_{t-2} + w_{t-3} \dots$$

Como puede notarse, que el proceso y_t se obtiene sumando, o lo que es lo mismo “integrando” el proceso w_t .

A este proceso integrado, y_t se le denomina proceso autorregresivo-medias móviles integrado o más abreviadamente $ARIMA(p,d,q)$, si tomando diferencias de orden d se obtiene un proceso estacionario w_t del tipo $ARMA(p,q)$.

La forma general del modelo $ARIMA(p,d,q)$ es:

$$w_t = \phi_1 w_{t-1} + \dots + \phi_p w_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q}$$

El modelo $ARIMA(p,d,q)$ representa un modelo estocástico general. Así, si $p=d=0$ tenemos un modelo $ARIMA(0,0,q)$. Por otro lado si $q=d=0$ el modelo $ARIMA(p,0,0)$ es el mismo $AR(p)$. Además si $d=0$ se tiene el modelo $ARIMA(p,0,q)$ o bien $ARMA(p,q)$.

La mayoría de las series temporales presentan tendencia creciente o decreciente. La eliminación de esta tendencia (no estacionaria en media) de la serie se puede conseguir mediante las diferenciaciones implícitas en los modelos $ARIMA$. También puede observarse en la varianza, es decir, la dispersión de las observaciones respecto a la media no es constante a lo largo del tiempo. Este tipo de tendencia no se elimina mediante el proceso de diferenciación, siendo lo adecuado transformar la serie aplicando logaritmos naturales.

Esta posibilidad de transformar la serie puede concretarse de la forma más general mediante la transformada de *Box-Cox*. De esta forma el modelo $ARIMA$ más general puede anunciarse como:

$$\phi(L)\Delta^d y_t^\lambda = \delta + \theta(L)u_t$$

Donde μ es la media de y_t^λ , siendo:

$$y_t^\lambda = \begin{cases} \frac{y_t^\lambda - 1}{\lambda} & \text{para } \lambda \neq 0 \\ \ln y_t & \text{para } \lambda = 0 \end{cases}$$

Restringiéndonos a los casos $\lambda = 1$ y $\lambda = 0$, tendremos respectivamente la serie sin transformar (la serie original) o la serie transformada aplicando logaritmos neperianos.

1.2.3. Modelos ARIMA Estacionales.

Hemos visto que podemos convertir series no estacionarias en estacionarias tomando diferencias regulares, es decir, entre periodos consecutivos. En la sección anterior hemos visto también que podemos eliminar la estacionalidad mediante diferencias estacionales. Uniendo ambos resultados, concluimos que podemos convertir una serie con estacionalidad en estacionaria mediante la transformación:

$$w_t = \nabla_s^D \nabla^d z_t,$$

donde D es el número de diferencias estacionales (si hay estacionalidad casi siempre $D=1$, y si no hay estacionalidad $D=0$) y d el número de diferencias regulares ($d \leq 3$).

Cuando exista dependencia estacional podemos generalizar el modelo *ARMA* para series estacionarias incorporando además de la dependencia *regular*, que es la asociada a los intervalos de la dependencia *estacional*, que es la asociada a las observaciones separadas por s periodos. Vamos a estudiar cómo modelar estos dos tipos de dependencias.

La primera solución es incorporar la dependencia estacional a la regular, añadiendo a los operadores *AR* o *MA* en el operador B términos B^s , para representar la dependencia entre observaciones separadas por s periodos muy largos en la parte *AR* y *MA*. Por ejemplo los datos mensuales, $s=12$, si un mes está relacionado con el mismo mes en tres años anteriores necesitamos un *AR* o *MA* de orden 36 para representar esta dependencia. Un enfoque más simple, y que funciona bien en la práctica, es modelar de forma separada la dependencia regular y la estacional, y construir el modelo incorporando ambas de forma multiplicativa. Se obtiene así el modelo *ARIMA estacional multiplicativo* que tiene la forma:

$$\Phi(B^s)\phi_p(B)\nabla_s^D\nabla^d z_t = \theta_q(B)\Theta_Q(B^s)a_t, \quad (1.1)$$

Donde $\Phi_p(B^s) = (1 - \Phi_1 B^s - \dots - \Phi_p B^{sp})$ es el operador *AR* estacional de orden P , $\phi_p = (1 - \phi_1 B - \dots - \phi_p B^p)$ es el operador *AR* regular de orden p , $\nabla_s^D = (1 - B^s)^D$ representa las diferencias estacionales, y $\nabla^d = (1 - B)^d$ las diferencias regulares, $\Theta_Q(B^s) = (1 - \Theta_1 B^s - \dots - \Theta_Q B^{sQ})$ es el operador media móvil estacional de orden Q , $\theta_q = (1 - \theta_1 B - \dots - \theta_q B^q)$ es el operador media móvil regular de orden q y a_t es un proceso de ruido blanco. Esta clase de modelos, introducidos por Box y Jenkins en (1976) representan bien muchas series estacionales que encontramos en la práctica y se escriben de forma simplificada como el modelo *ARIMA*(P, D, Q)_s $x(p, d, q)$.

1.3. Regresión dinámica entre variables estacionarias.

1.3.1. Introducción.

Un modelo de regresión dinámica es una representación de la relación existente entre dos o más series temporales. El modelo de regresión lineal simple supone dos variables x_t , y_t , que siguen procesos de ruido blanco y expresa la media condicionada de la variable dependiente y_t , en función de los valores de la explicativa, x_t , mediante:

$$y_t = \beta_0 + \beta_1 x_t + u_t \quad (1.2)$$

donde las variables u_t que no son observadas, son las innovaciones del modelo que siguen también un proceso de ruido blanco. Al intentar aplicar este modelo a variables que no son ruido blanco pueden tener dependencia temporal y nos encontremos con tres problemas principales.

El primero es que la ecuación (1.2) supone que la relación entre las dos variables es instantánea, cuando entre variables dinámicas la relación puede ser más compleja y transmitirse con ciertos retardos: el valor de y_t puede depender no de x_t sino de x_{t-h} , con un retardo h desconocido, o depender de todas las variables $(x_t, x_{t-1}, \dots, x_{t-h})$. Por ejemplo, los lectores habrán seguramente advertido que si tratamos de aumentar la temperatura del agua caliente, la respuesta del sistema suele ser gradual, de manera que la temperatura de salida del agua tarda unos segundos en alcanzar el nivel deseado. Esto ocurre especialmente cuando existe cierta distancia entre la ducha y la caldera que calienta el agua. Si llamamos y_t a la temperatura del agua que sale de la ducha en el momento t y x_t a la apertura de la llave en ese instante, un aumento de x_t en $t = t_0$ produce efectos en la respuesta y_t en los segundos, aumentando la temperatura gradualmente durante h hasta un cierto nivel.

En consecuencia y_t depende no sólo de la apertura de la llave en t , si no de la apertura en los segundos previos de x_t , es decir, de las variables x_{t-1}, \dots, x_{t-h} . Como segundo ejemplo, si y_t representa los pasajeros transportados en un vuelo el día t y x_t es el precio del billete en ese día, un aumento del precio hoy tendrá poco efecto sobre los pasajeros transportados hoy, ya que la gran mayoría habrán comprado su billete antes de la subida. El incremento del precio, x_t , afectará a los pasajeros futuros y_{t+1}, y_{t+2}, \dots ocasionando una disminución de la demanda en los próximos días. Esto es equivalente a decir que la variable y_{t+h} dependerá no solo x_{t+h} sino también de los valores previos x_{t+h-1}, \dots, x_t . El efecto dinámico será tanto más complejo cuánto más intervalos de tiempo sean necesarios para la transmisión completa del efecto. Por ejemplo, si trabajamos con datos anuales, todos los efectos que se transmitan en pocos días o meses aparecerán como efectos instantáneos, es decir todo el efecto de una subida de precios en el año t se observará en la respuesta ese año, ya que todos los efectos quedan incluidos en el periodo de observación de la serie.

El segundo problema de (1.2) es que este modelo supone implícitamente que la relación va de la variable x_t a la y_t , pero que la variable y_t no influye sobre la x_t . Decimos que la variable y_t es endógena, influenciada por x_t , pero que x_t es exógena o explicativa, lo que implica que no hay relación entre y_{t-h} y x_t para cualquier $h > 0$. Por ejemplo, la cantidad de lluvia influye sobre el crecimiento de una planta, pero no es esperable que el crecimiento de una planta afecte la lluvia futura. Aunque en muchas aplicaciones físicas o técnicas el carácter exógeno de la variable x_t es evidente por el mecanismo generador de los datos. Por ejemplo, los gastos en publicidad, x_t , es probable que influyan sobre las ventas futuras y_{t+h} sino que exista realimentación o causalidad bidireccional entre ambas variables de manera que y_t también influye sobre x_{t+m} . En este capítulo supondremos que conocemos la dirección de causalidad, y por tanto, que el modelo de regresión dinámica es adecuado para las variables. Cuando sospechamos que la variable y_t puede también influir sobre la x_{t+m} tendremos que utilizar los métodos de causalidad bidireccional.

El tercer problema de la ecuación (1.2) es suponer que la parte de la variable respuesta no explicada por la variable independiente x_t , es un proceso u_t de variables independientes. Esta hipótesis es natural para datos estáticos, pero no suele verificarse con datos dinámicos. Si pensamos de nuevo en el ejemplo de la temperatura del agua en la ducha, la ecuación (1.2) indica que cuando x_t (apertura de la llave) permanece constante, la temperatura del agua debe variar aleatoriamente de forma imprevisible. Sin embargo nuestra experiencia diaria, nos indica que, afortunadamente la temperatura evoluciona siempre con inercia, de manera que la temperatura en un instante, y_t será próxima en el instante anterior, y_{t-1} y ésta a y_{t-2} , etc. Lo que supone relación entre las innovaciones. De la misma forma si, el número de pasajeros transportados en un vuelo depende del día de la semana, las innovaciones estarán autocorreladas, por el efecto de la estacionalidad semanal.

En éste capítulo vamos a estudiar modelos para relacionar series estacionarias. Los modelos que vamos a estudiar se conocen en las ciencias físicas, naturales y de ingeniería como *modelos de función de transferencia* y en las ciencias económicas como *modelos econométricos dinámicos*. Describen como se transmiten los efectos desde una variable x_t a otra y_t cuando no existe realimentación o causalidad bidireccional y es un instrumento muy utilizado en todos los campos científicos para evaluar respuestas dinámicas. Si las variables explicativas son controlables estos modelos permiten simular y evaluar políticas alternativas. Si no lo son ofrecen la posibilidad de estudiar como ciertos “escenarios” (definidos por posibles evoluciones de la variable explicativa) que afectan a la variable respuesta.

La utilidad de estos modelos para la predicción depende del intervalo entre observaciones. Si es grande por ejemplo datos anuales, la relación entre la variable explicativa y la respuesta no suele contener retardos y para prever y_t es necesario conocer x_t , en el caso más general en que la variable x_t es estocástica, para preverla es necesario haber establecido su modelo ARIMA univariante y con esto es posible mejorar las predicciones futuras del modelo de regresión dinámico establecido para y_t , y el modelo puede ser muy adecuado para la predicción condicional de la respuesta en

función de posibles valores de la variable explicativa. La situación es distinta cuando el periodo de observación es corto, por ejemplo diario o cada minuto, donde es frecuente que existan retardos en la relación entre las variables, de manera que x_{t-h} afecta a y_t ($h > 0$). En este caso diremos que la variable x_t es un indicador avanzado de la respuesta y como el regresor que necesitamos para prever la respuesta en el instante t es conocido, el modelo de regresión dinámica suele proporcionar mejores predicciones de y_t que el univariante.

1.3.2. Relaciones entre dos series estacionarias.

1.3.2.1 Función de covarianzas cruzadas.

Supongamos que tenemos dos procesos estacionarios x_t , y_t , con medias μ_x , μ_y . Vamos a introducir medidas de su dependencia lineal para distintos retardos. Se define la *función de autocovarianzas cruzadas*, mediante:

$$\gamma_{xy}(t, t+h) = E[(x_t - \mu_x)(y_{t+h} - \mu_y)]. \quad (1.3)$$

Esta función depende del orden en que se toman las variables y de los instantes en que se consideran cada una de ellas. La definición establece que la variable que aparece en primer lugar en el subíndice, x , se supone en t y la segunda, y , en $t+h$. Por tanto $\gamma_{xy}(t, t+h)$ mide la relación lineal entre las dos variables x_t y y_{t+h} . El valor de h puede ser positivo o negativo. Se verifica que $\gamma_{xy}(t, t+h)$ es idéntica a $\gamma_{yx}(t+h, t)$ y diremos que dos procesos estacionarios x_t , y_t , son *conjuntamente estacionarios* si cada uno de ellos es estacionario y las covarianzas cruzadas sólo dependen del retardo entre las variables, h , y no del instante inicial considerado. Entonces, escribimos la función de covarianzas cruzadas como:

$$\gamma_{xy}(t, t+h) = \gamma_{xy}(h), \quad h = 0, \pm 1, \pm 2, \dots,$$

En adelante, para simplificar la exposición, diremos normalmente que dos variables son estacionarias para indicar que son conjuntamente estacionarias. La función de covarianzas cruzadas para procesos estacionarios no varía si intercambiamos las variables y el signo del retardo simultáneamente:

$$\gamma_{xy}(h) = \gamma_{yx}(-h)$$

En efecto:

$$\begin{aligned} \gamma_{xy}(h) &= E[(x_t - \mu_x)(y_{t+h} - \mu_y)] \\ &= E[(x_{t-h} - \mu_x)(y_t - \mu_y)] = \gamma_{yx}(-h) \end{aligned}$$

ya que, por la estacionariedad, la covarianza sólo depende del retardo.

Concluimos que la función $\gamma_{xy}(h)$, ($h = 0, \pm 1, \pm 2, \dots$) resume toda la dependencia lineal entre ambas variables y no es necesario calcular la función γ_{yx} .

La función $\gamma_{xy}(h)$ no es simétrica respecto al retardo h , y se verifican las propiedades siguientes:

1. Para $h > 0$ los coeficientes $\gamma_{xy}(h)$ representa cómo los valores de x_t influye en los valores futuros y_{t+h} , lo que interpretamos diciendo que la parte positiva de esta función representa la relación “causal” de x_t hacia y_{t+h} .
2. Para $h < 0$ los coeficientes $\gamma_{xy}(h)$ representa cómo los valores de y_t influye en los valores futuros x_{t+h} , lo que interpretamos diciendo que la parte negativa de esta función representa la relación “causal” de y_t hacia x_{t+h} .

Diremos que existe relación contemporánea o instantánea entre dos variables si $\gamma_{xy}(0) \neq 0$. De la misma forma que las autocovarianzas permiten identificar el orden del modelo *ARMA*, podría pensarse en utilizar la función de covarianzas cruzadas para identificar si la relación es en una dirección o bidireccional y también obtener el número de retardos distintos de cero en la relación entre dos variables. Sin embargo no es así, ya que esta función tiene muchas limitaciones, como se ilustra a continuación. Supongamos que la relación entre dos variables estacionarias de media cero es:

$$y_t = v_0 x_t + v_1 x_{t-1} + a_{t+h}. \quad (1.4)$$

Multiplicando por x_{t-h} y tomando esperanzas, resulta:

$$\begin{aligned} \gamma_{xy}(h) &= E(x_{t-h} y_t) = v_0 E(x_{t-h} x_t) + v_1 E(x_{t-h} x_{t-1}) + E(x_{t-h} a_t) \\ &= v_0 \gamma_x(h) + v_1 \gamma_x(h-1). \end{aligned} \quad (1.5)$$

Vemos que la función de covarianzas cruzadas es una mezcla de los coeficientes v_i de la relación entre las dos variables y de la estructura de autocovarianzas de x_t , con lo que será complicada interpretar. Supongamos que la variable x_t sigue un proceso *AR(1)*, de manera que $\gamma_x(h) \neq 0$. Entonces si existe correlación contemporánea, $v_0 \neq 0$, las correlaciones cruzadas $\gamma_{xy}(h)$ serán no nulas para $h > 0$, indicando relación desde x_{t-h} hacia y_t , aunque según (1.4) ésta relación existe para $h=0,1$. Además, también serán no nulas para $h < 0$, indicando una relación desde y_{t-h} hacia x_t , que no existe. Por ésta razón no podemos utilizar las covarianzas cruzadas para determinar ni la dirección ni la relación, ni el número de retardos de la ecuación. Sin embargo, si la serie x_t fuese ruido blanco, entonces este problema desaparece, ya que $\gamma_x(h) = 0$ para $h \neq 0$, y las únicas covarianzas cruzadas no nulas son $\gamma_{xy}(0)$ y $\gamma_{xy}(1)$ con valores proporcionales a los coeficientes v_i de la relación.

En resumen, si $h < 0$, todas las covarianzas cruzadas $\gamma_{xy}(h)$ son cero, podemos concluir que no hay relación desde y_{t-h} hacia x_t , pero lo contrario no es cierto: es posible que la causalidad sea unidireccional, y que encontremos coeficientes distintos de cero en esta función. Sólo cuando x_t es ruido blanco podemos interpretar simplemente esta función. Estas propiedades hacen que las covarianzas cruzadas no se utilicen para identificar el modelo, aunque tienen otra utilidad para su diagnóstico, como veremos posteriormente.

Función de autocorrelación cruzada.

Definimos la función de correlación cruzada, $\rho_{xy}(h)$, de dos procesos estocásticos conjuntamente estacionarios x_t, y_t , por:

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sigma_x \sigma_y} \quad (h = 0, \pm 1, \pm 2, \dots)$$

Esta función es la estandarización de $\gamma_{xy}(h)$ y, en consecuencia, tiene propiedades análogas. Se verifica que:

1. $|\rho_{xy}(h)| \leq 1$;
2. $\rho_{xy}(h) = \rho_{yx}(-h)$, esto se cumple por la estacionariedad de las variables, ya que la covarianza sólo depende del retardo $\gamma_{xy}(h) = \gamma_{yx}(-h)$. Pero en general $\rho_{xy}(h) \neq \rho_{yx}(h)$, y la función no es simétrica alrededor del origen. Para $h > 0$ mide la relación desde x hacia y , mientras que para $h < 0$ la relación que mide de y hacia x .

La función de correlación cruzada, al ser la estandarización de la covarianza tiene los mismos problemas que ésta para identificar los retardos que existen en la relación entre dos variables: mezcla los coeficientes de la relación con los de autocorrelación.

Estimación de las funciones de covarianzas y correlaciones cruzadas.

Sabemos que las autocovarianzas teóricas de un proceso y_t se estiman mediante las muestrales, definidas por:

$$\hat{\gamma}_y(h) = \frac{1}{T} \sum_{t=1}^{T-h} (y_t - \bar{y})(y_{t+h} - \bar{y}).$$

De la misma forma, todas las series temporales, x_t, y_t , estimaremos sus covarianzas cruzadas mediante las muestrales :

$$\hat{\gamma}_{xy}(h) = \frac{1}{T} \sum_{t=1}^{T-h} (x_t - \bar{x})(y_{t+h} - \bar{y}),$$

y las correlaciones cruzadas muestrales por:

$$r_{xy}(h) = \frac{\hat{\gamma}_{xy}(h)}{s_x s_y},$$

donde $s_x = \hat{\gamma}_x^{1/2}(0)$, $s_y = \hat{\gamma}_y^{1/2}(0)$.

Las propiedades muestrales de las funciones de correlación cruzadas son complicadas. Sin embargo, podemos utilizar esta función para contrastar la causalidad unidireccional de manera simple en el caso particular en que uno de los dos procesos es *ruido blanco*. En este caso particular, puede demostrarse que si ambos procesos están incorrelados, las varianzas de los coeficientes de autocorrelación cruzada muestrales entre un proceso y_t y un ruido blanco a_t son:

$$\text{Var}[r_{ay}(h)] \approx \frac{1}{T-h}. \quad (1.6)$$

Es decir, podemos aproximar las desviaciones típicas de las estimaciones de los coeficientes de correlación cruzada de manera idéntica a las de las autocorrelaciones mediante $1/\sqrt{T}$. Por tanto, si x_t es ruido blanco podemos construir bandas $2/\sqrt{T}$ sobre las autocorrelaciones y si $r_{ay}(h) = 0$ para $h < 0$ concluimos que no hay relación desde y_t hacia x_{t+h} . La parte de ésta función para $h > 0$ nos permitirá investigar si existe o no relación desde x_{t-h} hacia y_t . Si la variable x_t no es ruido blanco, este contraste puede aplicarse utilizando los residuos del modelo univariante de x_t .

Para interpretar la función de correlación cruzada hay que tener en cuenta que las estimaciones para distintos retardos están correladas, y para valores moderados de h las covarianzas entre los coeficientes de correlación muestrales $r_{ya}(h)$ y $r_{ya}(h+k)$ pueden calcularse aproximadamente por:

$$\text{Cov}[r_{ya}(h), r_{ya}(h+k)] \approx \frac{\rho_y(k)}{T-h},$$

y utilizando (1.6) se obtiene que:

$$\rho[r_{xa}(h)r_{xa}(h+k)] \approx \rho_x(k), \quad (1.7)$$

es decir, aunque y_t y a_t no estén correlados entre sí (por tanto $\rho_{ya}(h) \equiv 0$), los coeficientes estimados $r_{ya}(h)$ sí lo estarán. La función de correlación cruzada estimada variará alrededor de cero con varianza dada por (1.6) pero, según (1.7), x_t y a_t al estar sus coeficientes relacionados de la misma forma que las autocorrelaciones de y_t , tenderá a reproducir el correlograma de y_t .

Esta propiedad hace que la función de correlación cruzada no sea fácil de interpretar, incluso en el caso más simple donde una de las series es ruido blanco existe independencia.

1.3.3. Modelo dinámico entre dos series estacionarias.

Los modelos que se estudiarán en éste capítulo suponen que existe la relación unidireccional de la serie x hacia la serie y pero no al contrario. Una relación unidireccional de x_t hacia y_t entre dos procesos estocásticos estacionarios de media cero puede representarse mediante:

$$y_t = v_0x_t + v_1x_{t-1} + v_2x_{t-2} + \dots + n_t \quad (1.8)$$

donde los coeficientes v_i de la *función de transferencia* describen la relación dinámica entre las dos series y el proceso n_t , que es también estacionario, recoge el efecto de todas las otras variables que pueden tener efecto sobre y_t , y se le denomina el *proceso de perturbación* o de *inercia*, o simplemente *perturbación*, de la relación. Esta ecuación puede escribirse de forma más compacta utilizando el operador retardo, de la siguiente manera:

$$y_t = v(B)x_t + n_t \quad (1.9)$$

donde

$$v(B) = v_0 + v_1B + v_2B^2 + \dots$$

se denomina la *función de transferencia* y a los coeficientes v_i se les denomina función de respuesta a impulsos.

La ganancia de la función de transferencia es:

$$g = v(1) = \sum_{i=0}^{\infty} v_i$$

y representa el efecto a largo plazo que experimenta y_t cuando x_t aumenta en una unidad y permanece constante a continuación.

El proceso de perturbación, n_t , seguirá un proceso *ARMA* estacionario:

$$\phi(B)n_t = \theta(B)a_t, \quad (1.10)$$

donde a_t es un proceso de ruido blanco. El modelo formado por las ecuaciones (1.8) y (1.9) se conoce como modelo de regresión dinámica o modelo de función de transferencia entre las dos variables.

Representación de la función de transferencia como cociente de polinomios.

La expresión (1.8) de la función de transferencia no es útil para ser estimada en la práctica. Si el proceso x_t tiene alta autocorrelación, las variables $x_t, x_{t-1}, x_{t-2}, \dots$ estarán muy correladas entre sí y será muy difícil separar sus efectos por el problema de la multicolinealidad. Además si la relación es larga tendremos que estimar muchos parámetros, con lo que el problema se complica.

Una forma alternativa de escribir la función de transferencia, es utilizar la idea de aproximar una función infinita como cociente de polinomios y, como hicimos en los modelos *ARMA*, representar la función mediante:

$$v(B) = \frac{w_m(B)}{\delta_a(B)} B^b,$$

Donde el término B^b tiene en cuenta que la relación puede establecerse con un retardo inicial b y:

$$w_m(B) = w_0 + w_1 B + \dots + w_m B^m,$$

Es el numerador de la función de transferencia que juega un papel similar a la parte *MA* de un modelo *ARMA*, y:

$$\delta_a(B) = 1 - \delta_1 B + \dots + \delta_a B^a,$$

es el denominador, que juega el papel de la parte *AR*. Los factores de esta ecuación deben ser menores que la unidad para que el sistema proporcione una modificación finita de la x_t . Es decir, si escribimos $\delta_a(B) = (1 - \delta_1 B) \dots (1 - \delta_a B)$, entonces $|\delta_i| < 1$.

En la práctica, los órdenes m y a son pequeños, no mayores que 3. La ganancia de la función de transferencia se calcula como antes:

$$g = v(1) = \frac{w_m(1)}{\delta_a(1)} = \frac{\sum_{j=0}^m w_j}{1 - \sum_{i=1}^a \delta_i}$$

Una ventaja de esta representación es separar la función de transferencia el efecto inicial de la pauta de decrecimiento, y al ser estos dos factores independientes podemos estimarlos mejor. Consideremos el caso más simple:

$$v(B) = \frac{w_0}{(1 - \delta B)} = w_0 (1 + \delta B + \delta^2 B^2 + \dots) \quad (1.11)$$

donde $|\delta| < 1$. Esta función representa una relación donde si x_t aumenta en una unidad y luego vuelve a su valor inicial, la respuesta tiene un aumento instantáneo de w_0 , seguido de un aumento en el siguiente periodo de $w_0 \delta$, y en los periodos siguientes $w_0 \delta^k$, con decrecimiento con tasa $|\delta| < 1$, hasta acumular el efecto total que es la ganancia $g = w_0 / (1 - \delta)$.

Por ejemplo, si $\delta = 0.7$ y $w_0 = 2$, un impulso en x_t , es decir un incremento en t de una unidad que desaparece en el instante $t+1$ produce aumentos en y_t de 2, en y_{t+1} de $(2)(.7) = 1.4$, en y_{t+2} de 0.98 y en y_{t+h} de $(2)(.7)^k$. Si sumamos todos estos efectos, obtenemos la ganancia, $2/0.3 = 6.667$.

Si la verdadera relación es del tipo (1.11) con δ no muy pequeño con la representación (1.8) tendríamos que estimar muchos coeficientes v_i y además independientes, con lo que podremos hacerlo con precisión.

Otro modelo muy útil en la práctica es:

$$v(B) = \frac{w_0 + w_1 B}{(1 - \delta B)} = w_0 + (w_1 + w_0 \delta)B + \delta(w_1 + w_0 \delta)B^2 + \dots \quad (1.12)$$

donde partimos que el efecto inicial sea distinto del que aparece en el decrecimiento gobernado por δ .

Representación de la función de transferencia como efectos a corto y a largo plazo.

Un método alternativo de escribir la función de transferencia es separando la respuesta a largo plazo, o ganancia, de los efectos a corto plazo. Toda función de transferencia de orden máximo k puede escribirse como:

$$v_k(B) = g + \sum_{i=0}^{k-1} \alpha_i \nabla B^i \quad (1.13)$$

donde: $\nabla = 1 - B$.

Utilizando esta representación, en lugar de la ecuación (1.9), podemos escribir:

$$y_t = g x_t + \sum_{i=0}^{k-1} \alpha_i \nabla x_{t-i} + n_t, \text{ donde } \nabla x_{t-i} = (1-B)x_{t-i} \quad (1.14)$$

Que separa el efecto a largo plazo, medido por $g = v(1)$, de los efectos a corto, medidos por los coeficientes α_i . Observemos que, en ambos casos, tenemos, $k+1$ coeficientes. Los v_i ($i=0, \dots, k$) en la izquierda y en la derecha los k coeficientes α_i ($i=0, \dots, k-1$) más el g . La ventaja de la formulación (1.14) frente a (1.9) es que las variables ∇x_{t-i} estarán mucho menos correladas entre sí que las x_t , con lo que sus efectos se estimarán mejor. Por ejemplo, supongamos que x_t sigue un proceso $AR(1)$ con alto coeficiente. Entonces ∇x_t es casi ruido blanco y los términos ∇x_{t-i} estarán muy pocos correlados entre sí y con x_t . La estimación de (1.14) será mucho más eficiente que la de (1.9).

Otra forma alternativa de separar los efectos a corto y a largo plazo es escribiendo la relación como:

$$y_t = g x_t + \sum_{i=0}^k \beta_i \nabla^i x_t + n_t, \text{ donde } \nabla^i x_t = (1-B)^i x_t \quad (1.15)$$

donde, de nuevo, hemos separado el efecto a largo plazo, g , de las fluctuaciones a más corto plazo dadas por las diferencias $\nabla^i x_t$. Se demuestra a continuación, que como en el caso anterior, cualquier polinomio en el operador de retardo de orden k puede escribirse como:

$$v_k(B) = v(1) + \sum_{i=0}^k \beta_i \nabla^i, \text{ donde } \nabla^i = (1-B)^i \quad (1.16)$$

Verificación de (1.16):

Dado un operador $v_k(B) = v_0 + v_1 B + \dots + v_k B^k$ de orden finito k , vamos a demostrar que siempre puede escribirse como:

$$v_k(B) = v(1) + \sum_{i=0}^{k-1} \alpha_i \nabla \beta^i, \text{ donde } \nabla = (1-B)$$

y también como (1.16).

La demostración de esta propiedad es inmediata igualando términos con los dos miembros. Por ejemplo para $k=2$, tenemos para la primera ecuación:

$$v_0 + v_1 B + v_2 B^2 = (v_0 + v_1 + v_2) + \alpha_1 (1-B) + \alpha_2 B(1-B)$$

de donde resulta, igualando las potencias más altas, en B^2 :

$$\alpha_2 = -v_2.$$

Igualando ahora las potencias de B en ambos miembros, tenemos $v_1 = -\alpha_1 + \alpha_2$, que conduce a:

$$\alpha_1 = -v_1 - v_2$$

Apliquemos el mismo método en el segundo caso para obtener los coeficientes β . Entonces:

$$v_0 + v_1 B + v_2 B^2 = (v_0 + v_1 + v_2) + \beta_1 (1-B) + \beta_2 (1-2B+B^2).$$

Es inmediato que ahora $\beta_2 = v_2$, y como tenemos v_1 en el primer miembro y $-\beta_1 - 2\beta_2$ en el segundo resulta:

$$\beta_1 = -(v_1 + 2v_2).$$

Puede demostrarse por inducción que, el resultado general es:

$$\beta_i = \frac{(-1)^i}{i!} v_k^{(i)}(1),$$

Donde $v_k^{(i)}(1)$ es la derivada i -ésima de $v_k(B)$ respecto a B particularizada luego para $B=1$.

Por ejemplo, si $k=2$ e $i=1$, entonces $v_2^{(1)}(B) = v_1 + 2v_2B$, y también $v_2^{(1)}(1) = v_1 + 2v_2$. Sustituyendo estas expresiones se obtiene que $\beta_1 = (-1)(v_1 + 2v_2)$, como obtuvimos directamente. La fórmula es también válida para $i=0$ y entonces $\beta_0 = v_k^{(0)}(1) = v_k(1)$, con lo que se verifica la expresión (1.16).

La expresión (1.16) indica que en la ecuación (1.15) el coeficiente de x_t es la ganancia, y los coeficientes de $\nabla^i x_t$ representan efectos a corto plazo. Los coeficientes β_i representan los efectos sobre la respuesta de los incrementos de x_t o velocidad de cambio de x_t , de la aceleración en el cambio de x_t (medida por $\nabla^2 x_{t-1}$), etc. Es fácil comprobar que si x_t tiene alta autocorrelación las variables $x_t, \nabla x_{t-1}, \nabla^2 x_{t-1}$, tienen menos autocorrelación que las variables x_t, x_{t-1}, x_{t-2} . Por tanto la ecuación (1.15) resulta mucho más conveniente que la (1.9) cuando la serie x_t tenga alta autocorrelación.

Las dos representaciones (1.14) y (1.15) tienen el mismo objetivo: separar los efectos a corto plazo de los efectos a largo plazo y escribir la relación utilizando variables explicativas tan poco correladas como sea posible para estimar mejor los efectos. Cuál de estas dos representaciones sea más eficaz depende de la estructura de la variable x_t .

1.3.4. Construcción de modelos de regresión dinámica en variables estacionarias.

Las etapas para el ajuste de un modelo de regresión dinámica entre dos variables estacionarias contempla tres fases igual que en la estimación de cualquier otro modelo las cuales son: identificación, estimación y diagnóstico.

1.3.4.1. Identificación.

Supone decidir la estructura de la función de transferencia y del modelo para la perturbación, que será $ARMA(p_n, q_n)$.

Podemos utilizar como modelo de la perturbación el de la variable y_t , suponiendo $p_n = p_y$ y $q_n = q_y$. La justificación de esta elección es que, si no hay relación entre las variables y los coeficientes de la función de transferencia son cero, entonces y_t seguirá el proceso de la perturbación.

En general, el modelo final para n_t será más simple que el de y_t , ya que es muy posible que x_t explique gran parte de la dinámica de la respuesta, y una posibilidad alternativa es tomar un modelo $AR(1) \times AR(1)_s$, para captar la autocorrelación regular y la estacional.

El método para identificar la función de transferencia depende del grado de autocorrelación en la variable x_t . Si esta variable sigue un proceso de ruido blanco o tiene una estructura de correlación muy pequeña, podemos estimar el modelo:

$$y_t = \beta_0 + \sum_{i=0}^k v_i x_{t-i} + n_t, \quad (1.17)$$

y estudiar la estructura de los coeficientes v_1, \dots, v_k para decidir el orden del numerador y del denominador de la función de transferencia. El orden k se elige teniendo en cuenta el tamaño muestral y procurando incluir todos los retardos posibles donde pueden existir efectos. En general $k < T/4$. Por ejemplo, con datos anuales k será pequeño $k \leq 2$, pero con datos horarios puede ser alto. Pondremos siempre en la ecuación un modelo con autocorrelación para n_t el cual se utilizará para evitar los problemas de regresión espuria que se comentan posteriormente.

Esta formulación de la ecuación de transferencia (1.17), no es conveniente cuando la serie x_t tenga fuerte autocorrelación, ya que entonces los coeficientes v_i serán difíciles de estimar por la alta multicolinealidad entre las variables explicativas. Es mejor estimar directamente modelos del tipo:

$$y_t = \beta_0 + \frac{w_0 + w_1 B}{(1 - \delta B)} B^b x_t + n_t. \quad (1.18)$$

y comprobar si encontramos coeficientes significativos que nos indican la forma de la respuesta. Si la correlación de la x_t es fuerte, una solución simple es estimar un modelo que nos separe los efectos a largo y a corto plazo, del tipo (1.14) y (1.15). Con los coeficientes significativos determinamos qué forma de la función de transferencia parece más adecuada, y si es posible, la escribimos como cociente de polinomios, como indica la ecuación (1.18).

En esta fase de exploración no conviene seguir reglas estrictas al eliminar coeficientes, y mantener aquellos que sean grandes aunque no sean formalmente significativos. Si se estiman varios coeficientes ($k > 2$) siempre existirá cierta multicolinealidad en la estimación y además los errores estándar que utilizamos sólo serán aproximados, ya que el modelo es también una aproximación al correcto. Por tanto recomendamos mantener todos los coeficientes con valores de t mayores que uno. Escribiendo el modelo estimado en la forma $v(B)$ podemos decidir la estructura de la función de transferencia de forma similar a como se hizo en el correlograma o *FAS* de un proceso *ARMA*. En Concreto:

- a) Si los primeros b coeficientes \hat{v}_i parecen iguales a cero, entonces la función empieza en el retardo b .
- b) Si hay m coeficientes aparentemente significativos sin una pauta característica de variación, tomar ese grado para el numerador de la función.
- c) Si se observa una estructura de decrecimiento geométrico en los coeficientes, incluir un término de orden uno en el denominador, $a=1$. Si el decrecimiento es sinusoidal suponer $a=2$.

De esta manera se establece la forma provisional para la función:

$$\frac{w_m(B)}{\delta_a(B)} B^b$$

que pasaremos a estimar eficientemente en la etapa siguiente.

1.3.4.2 Estimación y selección del modelo.

Una vez establecida la función de transferencia y el modelo para la perturbación y la hipótesis de normalidad, la función de verosimilitud puede escribirse siguiendo el mismo procedimiento que en el caso univariante. Por ejemplo si el modelo para el ruido es $AR(p)$ y en la función de transferencia $\alpha=0$, llamando $h=\max(p,m+b)$, $y = (y_1, \dots, y_T)'$, $y_h = (y_1, \dots, y_h)$, el logaritmo de la verosimilitud o función soporte puede escribirse como:

$$L(y) = \sum_{t=h+1}^T L(y_t | y_{t-1}, \dots, y_1) + L(y_h), \quad (1.19)$$

Esto es porque si se considera el proceso $AR(1)$ de media μ , $y_t = c + \phi y_{t-1} + \alpha_t$, con $c = \mu(1-\phi)$ y $f(y_1)$ tiene una distribución normal. En este caso, $E(y_1) = \mu$ y $Var(y_1) = \frac{\sigma^2}{1-\phi^2}$.

Con lo anterior $y_{10} = \mu$ y $v_{10} = (1-\phi^2)^{-1}$. Para y_2 los momentos de la distribución condicionada a y_1 son $y_{2|1} = E(y_2 | y_1) = c + \phi y_1$ y $Var(y_2 | y_1) = E[(y_2 - c - \phi y_1)^2] = E[\alpha_2^2] = \sigma^2$ con lo que $v_{2|1} = 1$.

De la misma forma se demuestra:

$$y_{t|t-1} = E(y_t | y_{t-1}) = c + \phi y_{t-1}, \quad t = 2, \dots, T.$$

$$Var(y_t | y_{t-1}) = \sigma^2 v_{t|t-1} = \sigma^2, \quad t = 2, \dots, T.$$

En consecuencia, la función de verosimilitud es:

$$f(y_T) = f(y_1) \prod_{t=2}^T \sigma^{-1} (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^T (y_t - c - \phi y_{t-1})^2 \right\}$$

Tomando logaritmos y utilizando que $f(y_1)$ es normal, con parámetros $E(y_1) = \mu$ y $Var(y_1) = \frac{\sigma^2}{1-\phi^2}$, resulta la función soporte:

$$L(\phi, \sigma^2 | y_T) = -\frac{T}{2} \ln(\sigma^2) + \frac{1}{2} \ln(1 - \phi^2) - \frac{(1 - \phi^2)(y_1 - \mu)^2}{2\sigma^2} - \frac{1}{2\sigma^2} \sum_{t=2}^T (y_t - c - \phi y_{t-1})^2$$

La expresión anterior muestra que, si prescindimos del primer término, la función soporte tiene la expresión habitual de un modelo lineal. Si condicionamos a la primera observación tenemos que:

$$f(y_2, y_3, \dots, y_T | y_1) = \prod_{t=2}^T \sigma^{-1} (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=2}^T (y_t - c - \phi y_{t-1})^2 \right\}$$

La función de verosimilitud condicionada asociada a esta función es:

$$L(\phi, \sigma^2 | y_1) = -\frac{(T-1)}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^T (y_t - c - \phi y_{t-1})^2$$

En consecuencia, todas las distribuciones condicionadas para $t = p+1, \dots, T$ son normales, con media igual a la predicción a un paso y varianza σ^2 . Y su función de verosimilitud es:

$$f(y_1, y_2, y_3, \dots, y_T) = \prod_{t=p+1}^T \sigma^{-1} (2\pi)^{-1/2} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{t=p+1}^T (y_t - \mu - \sum_{i=1}^p \phi_i (y_{t-i} - \mu))^2 \right\}$$

La función soporte condicionada a la expresión anterior es:

$$L_c(\mu, \phi, \sigma^2 | y_1, \dots, y_p) = -\frac{(T-p)}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=p+1}^T (y_t - \mu - \sum_{i=1}^p \phi_i (y_{t-i} - \mu))^2$$

Ahora prescindiendo de constantes y despreciando el segundo término en la ecuación (1.19), la función soporte condicional para el vector de parámetros β que incluye todos los del modelo, será:

$$L(\beta | y_h) = \frac{T-h}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=h+1}^T a_t^2$$

donde a_t son los residuos. La estimación condicional equivale, como en el caso univariante, a la minimización de la suma de cuadrados generalizados. Para ilustrar el procedimiento vamos a suponer el caso más simple:

$$y_t = \beta_0 + v_0 x_t + n_t, \quad (1.20)$$

donde la perturbación sigue el proceso $AR(1)$

$$n_t = \rho n_{t-1} + a_t,$$

y a_t es ruido blanco. Si conociésemos el valor de ρ , multiplicando (1.20) por el operador $(1 - \rho B)$, tendríamos que:

$$(1 - \rho B)y_t = (1 - \rho B)\beta_0 + v_0(1 - \rho B)x_t + a_t,$$

y definiendo las nuevas variables, $\tilde{y}_t = y_t - \rho y_{t-1}$, $\tilde{x}_t = x_t - \rho x_{t-1}$, podríamos aplicar mínimos cuadrados a la regresión:

$$\tilde{y}_t = c - v_0 \tilde{x}_t + a_t,$$

Entre \tilde{x}_t e \tilde{y}_t y obtener v_0 . Esto equivale a aplicar a las variables la transformación:

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_T \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \cdots & 0 & 0 \\ \vdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}$$

Que, es aproximadamente el resultado de aplicar el método de mínimos cuadrados generalizados y se justifica a continuación.

Mínimos Cuadrados Generalizados.

Para explicar el procedimiento escribimos el modelo de regresión lineal en su forma general:

$$Y = X\beta + N$$

donde Y es un vector de dimensión $T \times 1$ de la variable dependiente, X es una matriz $T \times p$ de variables explicativas y β es un vector $p \times 1$ de parámetros. La primera columna de X contiene un vector de unos y las siguientes los valores de las variables dependientes, que incluyen valores actuales y retardos de cada una de las variables. Veremos que el estimador mínimo cuadrático es:

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

Sin embargo, con series temporales es esperable que las perturbaciones, N , no sean ruido blanco y tengan una estructura de dependencia definida por su matriz de varianzas covarianzas:

$$E[NN'] = \sigma^2 G = \sigma^2 \begin{pmatrix} 1 & Q_1 & \cdots & Q_{T-1} \\ Q_1 & 1 & \cdots & Q_{T-2} \\ \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ Q_{T-1} & Q_{T-2} & \cdots & 1 \end{pmatrix}$$

Supongamos que la matriz G fuese conocida. Entonces podríamos transformar el problema en otro con nuevas perturbaciones incorreladas y aplicar después mínimos cuadrados a la ecuación transformada. Toda la matriz definida positiva admite una

representación como $G = LL'$; donde la matriz L es triangular, que se conoce como descomposición de Cholesky, y podemos transformar el modelo mediante:

$$L^{-1}Y = L^{-1}X\beta + L^{-1}N$$

y llamando $Y_G = L^{-1}Y$ y $X_G = L^{-1}X$ a las variables transformadas y $E = L^{-1}N$ al nuevo de perturbaciones, que tendrán matriz de varianzas covarianzas:

$$\begin{aligned} E[EE'] &= E[L^{-1}NN'L^{-1}] = L^{-1}E[NN']L^{-1} = L^{-1}\sigma^2GL^{-1} \\ &= \sigma^2L^{-1}GL^{-1} = \sigma^2L^{-1}LL'L^{-1} = \sigma^2I \end{aligned}$$

y tendrá perturbaciones incorreladas. Aplicando mínimos cuadrados a las variables Y_G , X_G , el estimador para β , será:

$$\hat{\beta}_G = (X_G'X_G)^{-1}X_G'Y_G$$

y sustituyendo las variables transformadas por las originales:

$$\hat{\beta}_G = (X'G^{-1}X)^{-1}X'G^{-1}Y$$

que llamaremos **estimador de mínimos cuadrados generalizados o simplemente estimador MCG**.

Por ejemplo, supongamos que el ruido sigue el proceso *ARMA*:

$$N = \Psi A$$

donde:

$$\Psi = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 \\ \psi_1 & 1 & \dots & 0 & 0 \\ \dots & 0 & \dots & 1 & 0 \\ \psi_T & \psi_k & \dots & \psi_1 & 1 \end{pmatrix}$$

es la matriz triangular de coeficientes de la representación $MA(\infty)$ del proceso y $A = (a_1, \dots, a_T)'$ es un proceso ruido blanco. Entonces:

$$\Pi N = A$$

donde $\Pi = \Psi^{-1}$ es la matriz autoregresiva. La matriz de covarianzas de N será:

$$E[NN'] = \Psi\Psi'\sigma^2$$

y la inversa de Ψ es Π . Por tanto, la transformación requerida para un proceso *AR(1)* utiliza su matriz de coeficientes *AR* que tendrá unos en la diagonal, el coeficiente ϕ en la siguiente subdiagonal y el resto serán ceros. Ya que en este caso nuestra matriz $G = \Psi\Psi'$, donde Ψ es conocida y es diagonal, entonces podemos reescribir nuestro modelo como:

$$\begin{aligned}\Psi^{-1}Y &= \Psi^{-1}X\beta + \Psi^{-1}N \\ \rightarrow Y_G &= X_G\beta + \Pi N = X_G\beta + A\end{aligned}$$

Donde tendremos nuevas variables transformadas y A será ruido blanco.

Y aplicando el MCG a este nuevo modelo tendremos:

$$\begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_T \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ -\rho & 1 & \cdots & 0 & 0 \\ \vdots & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & -\rho & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}$$

El análisis anterior es aproximado para los primeros valores, ya que la ecuación $N = \Psi A$ no genera exactamente el proceso, sino solo después de los primeros retardos. Por ejemplo, para un $MA(1)$ todos los Ψ_j son cero para $j \geq 2$ y la ecuación que resulta para $u_1 = a_1$ y solo a partir de $j = 2$ obtenemos $u_2 = a_2 + \psi a_1$. Por tanto, la matriz de varianzas y covarianzas teórica de un $MA(1)$ no es exactamente $\Psi\Psi'\sigma^2$, pero sí aproximadamente, si despreciamos la primera observación.

En general, el proceso de las perturbaciones es desconocido, pero el procedimiento anterior puede aplicarse iterativamente: (1) estimados por mínimos cuadrados una regresión entre ambas variables y obtenemos los residuos $\hat{n}_t = y_t - \hat{\beta}_0 + v_0 x_t$; (2) con estos residuos estimados el $AR(1)$, $\hat{n}_t = \hat{\rho}\hat{n}_{t-1}$; (3) con el coeficiente estimado transformamos las variables en $\tilde{y}_t = y_t - \hat{\rho}y_{t-1}$ y $\tilde{x}_t = x_t - \hat{\rho}x_{t-1}$, y con las nuevas variables repetimos (1), (2) y (3) hasta que los residuos de la regresión sean ruido blanco.

Este proceso puede generalizarse como sigue. Supongamos el modelo:

$$y_t = v(B)x_t + \frac{1}{\pi(B)}a_t, \quad (1.21)$$

Entonces la estimación puede llevarse a cabo en las siguientes etapas:

1. Estimar una ecuación entre y_t y x_t , obtener los residuos, n_t , y estimar el modelo $\hat{n}_t^{(0)} = \hat{\pi}^{(0)}(B)a_t^{(0)}$ con estos residuos. Alternativamente, podemos suponer como modelo para n_t el modelo univariante de y_t . Tomar $i = 0$, y definir $y_t^{(0)} = y_t$, $x_t^{(0)} = x_t$, y $\hat{n}_t^{(0)} = \hat{\pi}^{(0)}(B)a_t^{(0)}$, que es el modelo inicial para la perturbación.
2. Hacer $i = i + 1$. Aplicar el modelo $\hat{\pi}^{(i-1)}(B)$ a la ecuación (1.21) para que tenga residuos incorrelados, mediante:

$$\hat{\pi}^{(i-1)}(B)y_t = v(B)\hat{\pi}^{(i-1)}(B)x_t + \left[\hat{\pi}^{(i-1)}(B) / \pi(B) \right] a_t$$

Llamar:

$$y_t^{(i)} = \hat{\pi}^{(i-1)}(B)y_t^{(i-1)} = y_t^{(i-1)} - \hat{\pi}_1^{(i-1)}y_{t-1}^{(i-1)} - \dots - \hat{\pi}_k^{(i-1)}y_{t-k}^{(i-1)},$$

y, $x_t^{(i)} = \hat{\pi}^{(i-1)}(B)x_t^{(i-1)}$. Suponer que $[\hat{\pi}^{(i-1)}(B)/\pi(B)] \approx 1$ y estimar los parámetros de la función de transferencia $v(B)$ por regresión $y_t^{(i)}$ y $x_t^{(i)}$ por mínimos cuadrados. Sea $y_t^{(i)} = \hat{v}^{(i)}(B)x_t^{(i)}$ la regresión estimada. Con estos valores obtener el proceso de inercia $\hat{n}_t^{(i)} = y_t^{(i)} - \hat{v}^{(i)}(B)x_t^{(i)}$.

3. Estimar los parámetros de la nueva serie del proceso de inercia. Sea $\hat{n}_t^{(i)} = \hat{\pi}^{(i)}(B)\alpha_t^{(i)}$ el modelo estimado. Si el modelo $\hat{\pi}^{(i)}(B)$ es prácticamente el mismo obtenido para $\hat{\pi}^{(i-1)}(B)$, finalizar la estimación. En otro caso, hacer $i = i + 1$ e ir a (2).

El algoritmo itera las dos etapas (2) y (3) hasta la convergencia. Si en la etapa (2) los parámetros de la función de transferencia no pueden obtenerse mediante regresión, porque tenemos términos en el denominador, esta etapa se sustituye por:

- 2*. Calcular recursivamente los residuos del modelo, $\hat{\alpha}_t$, obteniendo primero el proceso de ruido y , a continuación, sus residuos. Por ejemplo, si $b = 0$, $m = 0$, $a = 1$ y $p = 1$, $q = 0$, y tenemos el modelo:

$$y_t = \frac{w_0}{(1 - \delta B)}x_t + \frac{w_0}{(1 - \phi B)}\alpha_t$$

entonces:

$$\alpha_t = y_t - \phi y_{t-1} + w_0 x_t + (\delta + \phi)x_{t-1} + (\delta^2 + \phi^2 + \delta\phi)x_{t-2} + \dots$$

que permite obtener los resultados a partir de unos valores iniciales. Sea β el vector de parámetros del modelo y $\hat{\beta}_0$ el valor inicial, utilizado en los cálculos anteriores. Sea:

$$S(\beta) = \sum \hat{\alpha}_t^2(\beta)$$

y $S(\hat{\beta}_0)$ el valor inicial de esta función. Calcularemos un nuevo valor de los parámetros que disminuya la suma de cuadrados mediante un algoritmo tipo Gauss-Newton e iteraremos hasta obtener convergencia.

Como en el caso univariante, las varianzas y covarianzas de los estimadores se obtendrán a partir de la matriz de segundas derivadas en el máximo. La maximización de la función de verosimilitud exacta se efectúa con estos mismos principios. Normalmente se estiman varios modelos modificando los parámetros y se elige el más adecuado con el criterio *BIC* (*Criterio de Información Bayesiana*).

$$BIC = T \ln \sigma_{MV}^2 + k \ln T.$$

Donde T es el tamaño muestral, σ_{MV}^2 el estimador MV de la varianza y k el número de parámetros.

El criterio BIC se utiliza en la selección de modelos para elegir el mejor entre un conjunto de modelos admisibles. Un modelo es mejor que otro si tiene un valor BIC menor. El BIC se basa en una verosimilitud integrada en la teoría Bayesiana. Tomando en cuenta también la parsimonia la cual consiste en seleccionar el modelo con menor número de parámetros.

1.3.5 Contrastes diagnósticos.

Una vez seleccionado un modelo, antes de aceptarlo conviene efectuar contrastes diagnósticos para comprobar que no presenta deficiencias detectables que pueden sugerir reformulaciones. Clasificaremos los contrastes que pueden aplicarse en: (a) sobre los parámetros estimados, y (b) sobre los residuos.

1.3.5.1 Contrastes sobre los parámetros.

Debe comprobarse que la ganancia concuerda con el comportamiento esperado del sistema. Conviene también examinar:

1. Si el modelo puede simplificarse eliminando operadores con valores próximos en el numerador y en el denominador.
2. Si las raíces de los polinomios autorregresivos cumplen las condiciones de estabilidad o estacionariedad.
3. Si los efectos de los distintos operadores del ruido y de la función de transferencia ofrecen una interpretación coherente con el comportamiento cualitativo del sistema que estamos modelando.

Comenzaremos comprobando que no tenemos raíces comunes que pueden simplificarse, o en un cociente de polinomios o en ambos miembros de la ecuación. Por ejemplo, el modelo:

$$y_t = \frac{5 - 2.8B}{1 - 0.6B} x_t + \frac{1 - 0.3B}{1 - 1.1B + 0.2B^2} a_t,$$

Puede simplificarse, cancelando raíces muy próximas, como $(1 - 0.56B)$ en el numerador y $(1 - 0.6B)$ en el denominador de la función de transferencia y $(1 - 0.3B)$ y $(1 - 0.31B)$ en el ruido, a:

$$y_t = 5x_t + \frac{1}{1 - 0.31B} a_t.$$

La presencia de raíces comunes o próximas puede a veces mostrarse por correlaciones altas entre los parámetros estimados, lo que sugeriría buscar una forma de parametrizar el modelo de separar mejor los distintos efectos.

Comprobaremos también que todos los parámetros incluidos son significativamente distintos de cero, comparando el valor obtenido con la desviación estándar estimada. Hay que recordar que estos contrastes son sólo válidos asintóticamente, por lo que conviene fijar el nivel de significación en función de la interpretación lógica del modelo, especialmente para los parámetros de la función de transferencia: si su signo y magnitud coinciden con lo esperado conviene mantenerlos, aunque el valor de su estadístico t asociado no sea exactamente significativo al 95%.

1.3.5.2 Contrastes sobre los residuos.

Consisten como siempre, en comprobar si los residuos son una secuencia de variables aleatorias normales con varianza constante e independientes.

Esto supone realizar contrastes para confirmar:

1. Normalidad.
2. Que su media es cero.
3. Que están incorrelados entre sí.

Estos contrastes son análogos al caso univariante.

Para diseñar contrastes de errores de especificación, veamos las consecuencias de una especificación incorrecta. Supongamos que el modelo verdadero es:

$$y_t = v(B)x_t + \psi(B)a_t,$$

pero erróneamente estimamos:

$$y_t = \hat{v}_1(B)x_t + \hat{\psi}_1(B)\hat{a}_t,$$

concluimos que los residuos estimados \hat{a}_t verificarán el modelo:

$$\hat{a}_t = \frac{v(B) - \hat{v}_1(B)}{\hat{\psi}_1(B)}x_t + \frac{\psi(B)}{\hat{\psi}_1(B)}a_t.$$

En esta expresión podemos encontrarnos en los casos siguientes:

1. Hemos especificado mal tanto la función de transferencia como el modelo para la perturbación, es decir, $\psi(B) \neq \hat{\psi}_1(B)$ y $v(B) \neq \hat{v}_1(B)$. Entonces, los residuos \hat{a}_t estarán tanto autocorrelados entre sí, ya que tienen estructura dinámica, como correlados con los valores de x_t . La autocorrelación de los residuos se detectará con su correlograma y la dependencia con la variable x_t estudiando la función de correlación cruzada entre los residuos del modelo y la variable explicativa.
2. Si la función de transferencia es incorrecta, aunque el modelo del ruido sea adecuado, también observaremos correlación cruzada entre los residuos y x_t y autocorrelación entre los residuos. En efecto, si la función de transferencia no es correcta pero está bien especificado el modelo del ruido, los residuos pueden escribirse como:

$$\hat{a}_t = \frac{v(B) - \hat{v}_1(B)}{\hat{\psi}_1(B)} x_t + a_t,$$

y esta ecuación implica autocorrelación entre los residuos estimados \hat{a}_t .

3. Si la ecuación de transferencia es correcta, es decir aproximadamente $v(B) = \hat{v}_1(B)$, pero el modelo del proceso de perturbación o de inercia es equivocado $\psi(B) \neq \hat{\psi}_1(B)$ habrá autocorrelación en los residuos \hat{a}_t , pero no existirá correlación cruzada entre \hat{a}_t y x_t .
4. Si tanto la función de transferencia como el modelo del proceso de inercia son correctos, no habrá ni autocorrelación, ni correlación cruzada.

En resumen, la función de correlación cruzada entre los residuos estimados \hat{a}_t y la variable explicativa x_t es muy útil para detectar posibles errores en el modelo. Si encontramos correlaciones cruzadas entre los residuos estimados y la x_t esto sugiere que la función de transferencia no está bien especificada. Además en este caso es posible que necesitemos también reformular el modelo del ruido. Si no hay correlaciones cruzadas entre los residuos estimados y la variable x_t concluimos que la función de transferencia es correcta y miraremos el correlograma de los residuos. Si hay autocorrelación, el modelo de la perturbación debe modificarse. En otro caso el modelo es correcto.

Dado que la función de correlación entre \hat{a}_t y x_t puede no ser muy clara debido a la propia estructura de dependencia de x_t , como ya comentamos en el caso general, es más operativo estudiar la correlación entre \hat{a}_t y a_t , siendo a_t los residuos del modelo univariante de x_t , ya que, al ser el proceso a_t ruido blanco, la imagen será mucha más clara.

1.3.6 Relaciones espurias entre variables dinámicas.

La metodología presentada para construir modelos de regresión entre variables temporales evita el problema que se conoce como regresión espuria y que consiste en que al relacionar con un modelo de regresión clásico de series temporales independientes podemos encontrar fácilmente relaciones significativas.

1.3.6.1 Relaciones sin dinámica.

Consideremos el caso más simple, aunque poco frecuente con series temporales, de un proceso de perturbación, $n_t = a_t$, que sea ruido blanco normal con $E(a_t^2) = \sigma_a^2$ y, por tanto sin dinámica, y una relación simple del tipo:

$$y_t = \beta_0 + \beta_1 x_t + a_t,$$

Donde la serie x_t es estacionaria. Los parámetros β_0 , β_1 y σ^2 , pueden estimarse utilizando la normalidad de las distribuciones condicionadas por máxima verosimilitud, que equivale a mínimos cuadrados. La estimación de la pendiente es:

$$\hat{\beta}_1 = \frac{\hat{\gamma}_{xy}(0)}{s_x^2}, \quad (1.22)$$

Donde $\hat{\gamma}_{xy}(0) = \sum_{t=1}^T (y_t - \bar{y})(x_t - \bar{x})/T$ es la covarianza cruzada muestral de orden cero

definida en la sección anterior y $s_x^2 = \hat{\gamma}_x(0) = \sum_{t=1}^T (x_t - \bar{x})^2/T$, la varianza de la serie

estacionaria x_t . La varianza de la estimación de la pendiente es:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_a^2}{Ts_x^2}. \quad (1.23)$$

Observemos que esta varianza depende del cociente entre la varianza del ruido, σ_a^2 , y la varianza de la señal, que es la varianza del proceso estacionario x_t dada por s_x^2 . Cuanto mayor sea la magnitud de la señal con relación encontrada es significativa necesitamos un estimador de σ^2 . El estimador centrado habitual es:

$$s_R^2 = \frac{\sum \hat{a}_t^2}{T-2}$$

y el contraste de que el parámetro β_1 es distinto de cero utiliza el estadístico:

$$t = \frac{\hat{\beta}_1 s_x \sqrt{T}}{\hat{s}_R} \quad (1.24)$$

Que si $\hat{\beta}_1$ es cero sigue una t de Student con $T-2$ grados de libertad. Rechazaremos que no existe relación cuando este estadístico esté significativamente alejado de cero.

1.3.6.2 Perturbaciones autocorreladas.

Supongamos ahora que la perturbación tiene estructura dinámica, es decir, el proceso n_t de (1.9) es un proceso estacionario general, y la relación es:

$$y_t = \beta_0 + \beta_1 x_t + n_t,$$

Si estimamos este modelo como en el caso anterior, suponiendo que la perturbación es ruido blanco, tendremos un estimador muy ineficiente y sesgado a encontrar relaciones entre procesos independientes. La razón es que la varianza del estimador de la pendiente (1.22) puede ser mucho mayor que la calculada por la fórmula (1.23), por lo que si utilizamos el estadístico (1.24) para el contraste podemos concluir que existe una relación inexistente en realidad.

Se demuestra a continuación que cuando relacionamos variables estacionarias y la perturbación tiene estructura dinámica, la varianza del estimador de mínimos cuadrados (1.22) de la pendiente de la recta es:

$$va(\hat{\beta}_1) = \frac{\sigma_n^2}{Ts_x^2} \left[1 + 2 \sum_{t=1}^{T-1} (T-1)\rho_n(t)E(r_x(t)) \right] \quad (1.25)$$

Donde σ_n^2 es la varianza de la perturbación y $\rho_n(t)$ y $r_x(t)$ son los coeficientes de autocorrelación de orden t de la perturbación y de la variable x , respectivamente.

Demostración:

Supongamos el modelo $y_t = \beta_0 + \beta_1 x_t + n_t$. Sumando para todos los datos y dividiendo por T , tenemos que: $\bar{y} = \beta_0 + \beta_1 \bar{x} + \bar{n}$, y restando estas dos ecuaciones podemos escribir: $y_t = \bar{y} + \beta_1(x_t - \bar{x}) + n_t - \bar{n}$.

Vamos a ver las consecuencias de estimar los parámetros por mínimos cuadrados cuando la perturbación no es ruido blanco. El estimador de la pendiente es:

$$\hat{\beta}_1 = \sum \frac{(x_t - \bar{x})}{Ts_x^2} (y_t - \bar{y})$$

y sustituyendo en la ecuación $y_t = \bar{y} + \beta_1(x_t - \bar{x}) + n_t - \bar{n}$, se obtiene que

$$\hat{\beta}_1 = \beta_1 + \sum \frac{(x_t - \bar{x})}{Ts_x^2} n_t$$

Supongamos que n_t es independiente de x_t y tiene esperanza cero. Tomando esperanzas en esta expresión comprobamos que $E(\hat{\beta}_1) = \beta_1$ y el estimador es centrado. Para calcular su varianza, escribamos:

$$(\hat{\beta}_1 - \beta_1)^2 = \left(\sum \frac{(x_t - \bar{x})}{Ts_x^2} n_t \right)^2.$$

Vamos ahora a tomar esperanzas en esta ecuación condicionando a Ts_x^2 . Entonces:

$$E \left[(\hat{\beta}_1 - \beta_1)^2 \mid Ts_x^2 \right] = \frac{1}{Ts_x^2} E \left(\sum_{t=1}^T \frac{(x_t - \bar{x})^2}{Ts_x^2} n_t^2 + 2 \sum_{t=1}^{T-1} \frac{(x_t - \bar{x})(x_{t+1} - \bar{x})}{Ts_x^2} n_t n_{t+1} + \dots + 2 \sum_{t=1}^{T-k} \frac{(x_t - \bar{x})(x_{t+k} - \bar{x})}{Ts_x^2} n_t n_{t+k} + \dots + 2 \frac{(x_1 - \bar{x})(x_T - \bar{x})}{Ts_x^2} n_1 n_T \right).$$

Esta expresión indica que esperanza condicionada depende de toda la dinámica de ambas variables. Llamando $r_x(k) = \sum_{t=1}^{T-k} (x_t - \bar{x})(x_{t+k} - \bar{x}) / Ts_x^2$ a los coeficientes de

autocorrelación muestral de la variable estacionaria x_t y $\rho_n(k)$ a los coeficientes de autocorrelación teóricos del proceso n_t , como:

$$\begin{aligned} E \left[\sum_{t=1}^{T-k} \frac{(x_t - \bar{x})(x_{t+k} - \bar{x})}{Ts_x^2} n_t n_{t+k} \right] &= \sum_{t=1}^{T-k} E \left[\frac{(x_t - \bar{x})(x_{t+k} - \bar{x})}{Ts_x^2} \right] E(n_t n_{t+k}) \\ &= \sigma_n^2 \rho_n(k) E(r_x(k)), \end{aligned}$$

tenemos que:

$$E \left[\left(\hat{\beta}_1 - \beta_1 \right)^2 \middle| Ts_x^2 \right] = \frac{\sigma_n^2}{Ts_x^2} \left[1 + 2(T-1)\rho_n(1)E(r_x(1)) + \dots + 2\rho_n(T-1)E(r_x(T-1)) \right].$$

Tomando esperanzas ahora respecto a Ts_x^2 y utilizando que $E(y) = E_x E_{y|x}(y)$, obtenemos finalmente que:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_n^2}{Ts_x^2} \left[1 + 2(T-1)\rho_n(1)E(r_{xx}(1)) + \dots + 2\rho_n(T-1)E(r_{xx}(T-1)) \right].$$

Por ejemplo, supongamos para simplificar que ambos procesos siguen procesos $AR(1)$ independientes:

$$x_t = \phi x_{t-1} + \varepsilon_t$$

y

$$n_t = \rho n_{t-1} + \xi_t.$$

Entonces $\rho_n(k) = \rho^k$ y tomando $E(r_x(k)) \approx \phi^k$ tenemos que:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_n^2}{Ts_x^2} \left[1 + 2 \sum_{k=1}^{T-1} (T-k) \rho^k \phi^k \right].$$

Los términos del sumatorio forman una suma progresión geométrica, $a + a^2 + \dots + a^j$, con $a = \rho\phi$, y puede comprobarse que:

$$\sum_{k=1}^{T-1} (T-k)a^k = \frac{1}{(1-a)^2} \left[a(1-a) - (1-a^T) \right]$$

y despreciando para simplificar el término $\rho^T \phi^T$, resulta que:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_n^2}{Ts_x^2} \left[\frac{1 + \rho\phi}{1 - \rho\phi} \right].$$

Por ejemplo, si $\rho = \phi = 0.9$ la varianza real del estimador de mínimos cuadrados es 9.5 veces la calculada sin tener en cuenta la autocorrelación. El estimador es muy ineficiente y podemos fácilmente encontrar relaciones que no existen.

Entonces de la ecuación (1.25) obtenemos las conclusiones siguientes:

1. Si la perturbación es ruido blanco, sustituyendo en (1.25) $\rho_n(t)$ por cero y haciendo $\sigma_n^2 = \sigma_a^2$ se obtiene la expresión (1.23), y el estimador es eficiente.
2. Si tanto la variable explicativa, x_t , como la perturbación, n_t , tiene autocorrelación positiva, la varianza del estimador de la pendiente puede ser mucho mayor que la calculada con la fórmula (1.23) bajo la hipótesis de independencia. En efecto, al calcular la varianza de los residuos tendremos una estimación de σ_n^2 , la varianza del proceso estacionario, pero no tendremos en cuenta el resto de los términos, subestimando la varianza.
3. Un caso especial aparece cuando la variable x_t es ruido blanco incorrelado con la perturbación n_t . Entonces sustituyendo en (1.25) $E(r_x(t))$ por cero tenemos que:

$$\text{var}(\hat{\beta}_1) = \frac{\sigma_a^2}{TS_x^2}.$$

En este caso, estimamos bien la varianza del estimador, por lo que evitamos el riesgo de la regresión espuria, pero utilizamos un estimador ineficiente, que podríamos mejorar estimando por mínimos cuadrados generalizados.

La implicación de este resultado es que al relacionar series temporales con autocorrelación, como es esperable que la perturbación, n_t , también la tenga, no debemos estimar una regresión clásica y utilizar las fórmulas habituales, ya que podemos fácilmente encontrar una relación significativa entre las dos variables independientes. En efecto, el estimador (1.22) tendrá una varianza que puede ser mucho mayor que la dada por la fórmula (1.23), y al contrastar la relación mediante el estadístico t con (1.24), como la varianza está muy subestimada, podemos concluir fácilmente que existe relación entre dos variables aunque esta no exista.

1.3.7 El modelo con varias variables explicativas.

La metodología que hemos descrito en las secciones anteriores se generaliza sin dificultad para p variables explicativas. El modelo será entonces:

$$y_t = \frac{w_1(B)B^{b_1}}{\delta_1(B)} x_{1t} + \dots + \frac{w_p(B)B^{b_p}}{\delta_1(B)} x_{pt} + n_t. \quad (1.26)$$

Con varias series explicativas que pueden estar correladas entre sí en lugar de intentar identificar la función de transferencia para cada una conviene pasar a estimar directamente el modelo conjunto y reformularlo a la vista del resultado. Un procedimiento simple es tomar como modelo inicial para la perturbación el modelo univariante para la serie y_t y suponer estructuras simples iniciales para todas las variables, del tipo:

$$\frac{w_0 + w_1 B}{1 - \delta B}$$

En función de los coeficientes significativos encontrados iremos reformulando el modelo hasta obtener la estructura adecuada.

Cuando las variables explicativas tengan alta autocorrelación, o estén altamente correladas entre sí, no conviene estimar el modelo inicial del tipo:

$$y_t = \sum v_{1i} x_{1,t-i} + \dots + \sum v_{pi} x_{p,t-i} + n_t \quad (1.27)$$

ya que existirá multicolinealidad muy alta y los coeficientes se estimarán muy ineficientemente. Es mejor reescribir las funciones de transferencia separando la ganancia de los efectos a corto plazo como vimos para una variable. Por ejemplo, podemos estimar como modelo inicial:

$$y_t = \sum_{i=1}^p \beta_{oi} x_{it} + \sum_{i=1}^p \beta_{li} \nabla x_{i,t} + n_t$$

que separa los efectos a largo plazo, medidos por β_{oi} , de los a corto plazo, dados por β_{li} . Para las $p_1 \leq p$ variables donde se observen efectos significativos a corto plazo podemos aumentar la dinámica a corto plazo añadiendo un término en la variable $\nabla^2 x_{i,t}$ y así sucesivamente hasta obtener el modelo:

$$y_t = \sum_{i=1}^p \beta_{oi} x_{it} + \sum_{j=1}^{p_1} \beta_{li} \nabla x_{i,t} + \dots + \sum \beta_{im} \nabla^i x_{m,t} + n_t$$

La estimación, selección y diagnosis de estos modelos se realiza con los mismos principios que hemos visto anteriormente.

1.3.8 Predicción.

Las predicciones de un modelo de regresión dinámica se calculan con los mismos principios que en el caso univariante. Se demuestra que la previsión que minimiza el error cuadrático medio de los errores es la esperanza condicionada a la información disponible. En consecuencia:

$$\hat{y}_t(k) = E[y_{t+k} | y_t, \dots, y_1; x_t, \dots, x_1]$$

Las previsiones se obtienen recursivamente como en el caso univariante. Escribiendo el modelo como:

$$y_t = v(B)x_t + \psi(B)a_t \quad (1.28)$$

tenemos que, para el horizonte k:

$$y_{t+k} = (v_0 x_{t+k} + v_1 x_{t+k-1} + \dots + v_m x_{t+k-m}) + a_{t+k} + \psi_1 a_{t+k-1} + \dots + \psi_m a_{t+k-m}$$

y tomando esperanzas condicionales a la información en t y suponiendo que conocemos los parámetros y, por tanto, las perturbaciones a_t hasta el origen de la predicción:

$$\hat{y}_t(k) = v_0 \hat{x}_t^U + \dots + v_k x_t + \dots + \psi_k a_t + \dots + \psi_{k+h} a_{t-h}$$

donde las previsiones \hat{x}_t^U , son las de esta variable dada su historia y se obtendrán con el modelo univariante para esta serie x_t . Vemos que, en general, la predicción con un modelo de regresión dinámica requiere conocer no solo la función de transferencia sino también el modelo univariante para y_t .

Para calcular la varianza de la predicción, supongamos que x_t sigue el modelo:

$$x_t = \psi_x(B)\alpha_t \quad (1.29)$$

Sustituyendo (1.29) en (1.28) tenemos:

$$y_t = v^*(B)\alpha_t + \psi(B)a_t$$

donde $v^*(B) = v(B)\psi_x(B)$. La predicción puede entonces escribirse también como:

$$\hat{y}_t(k) = v_k^*\alpha_k + v_{k+1}^*\alpha_{k-1} + \dots + \psi_k a_t + \psi_{k+h} a_{t-1} + \dots$$

y su varianza será:

$$\text{var}(\hat{y}_t(k)) = E(y_{t+k} - \hat{y}_t(k))^2 = \sigma_\alpha^2 \sum_{j=0}^{k-1} v_j^{*2} + \sigma^2 \sum_{j=0}^{k-1} \psi_j^2$$

Capítulo II.

Modelos Multivariantes de Series Temporales.

2.1. Introducción.

A menudo se toman observaciones de manera simultánea en dos o más series de tiempo. Por ejemplo, podemos observar diversas medidas de la actividad económica en un determinado país, a intervalos regulares de tiempo, digamos cada mes. Las variables podrían incluir el índice de precios al por menor, el nivel de desempleo y un índice ponderado de precios de las acciones. Teniendo en cuenta estos datos multivariantes, puede ser conveniente tratar de desarrollar un modelo multivariado para describir las interrelaciones entre las series, en este capítulo se estudiarán este tipo de modelos para hacer previsiones o pronósticos. Con series cronológicas de datos, el proceso de modelización se complica por la necesidad de un modelo, no sólo en la interdependencia entre las series, sino también por la dependencia entre sus componentes.

En este capítulo se presenta una variedad de modelos de series temporales multivariantes y los métodos de predicción que se basan en ellos. Los modelos incluyen la función de transferencia, matrices de autocorrelación simple y parcial, las versiones multivariantes de los modelos AR, MA y ARMA, que son los modelos VAR, VMA y VARMA.

El Ajuste de los modelos de series temporales multivariante aun no es fácil a pesar de las enormes mejoras en los programas informáticos en los últimos años. Hay muchas investigaciones en curso, en parte estimulada por la mejora de los recursos computacionales, pero aún queda mucho por hacer.

Los modelos univariantes pueden ser útiles para muchos propósitos, incluyendo la previsión de un gran número de series temporales, y proporcionar un punto de referencia en estudios comparativos de previsión, pero, parece claro que los modelos multivariantes también tienen mucho que ofrecer para obtener una mejor comprensión de la estructura subyacente de un determinado sistema y con suerte obtener mejores pronósticos. Los modelos multivariantes pueden darnos un mejor ajuste que los modelos univariantes, aunque existen varias razones por las cuales no necesariamente tiene que ser así (aunque, por supuesto, a veces lo hacen). Las razones son las siguientes:

- (i) Hay que estimar más parámetros, con esto existen más posibilidades de variación de la muestra para aumentar la incertidumbre de los parámetros y afectar a los pronósticos.
- (ii) Con más variables a medir, hay más oportunidades para que los errores y valores atípicos puedan entrar silenciosamente.
- (iii) Los datos multivariantes observados no necesariamente puede ser adecuados para ajustar un modelo multivariado.
- (iv) El cálculo de los pronósticos de una variable dependiente puede requerir los valores futuros de las variables explicativas, y estos pueden no estar disponibles en el momento de la previsión en que se deben hacer. Luego, las variables explicativas deben ser predichas de alguna manera antes de que las previsiones de la variable dependiente se hayan encontrado, y esto conduce inevitablemente

a una reducción en la precisión. Si las predicciones de las variables explicativas tienen una exactitud pobre, entonces las predicciones resultantes de la variable dependiente puede tener una peor precisión que las establecidas por su modelo univariante.

(v) El cálculo de las predicciones de múltiples variables depende de tener un modelo multivariante bueno, pero esto no puede ser garantizada. A diferencia de los modelos univariantes, un modelo multivariante se puede establecer de forma incorrecta o puede cambiar durante el período de ajuste o en el futuro. Parece que los modelos de series temporales multivariantes, son mucho más complicados, y son más vulnerables a los errores de especificación, en comparación de los modelos univariantes.

La experiencia en la búsqueda de un modelo multivariado casi corrobora la importancia de obtener información de fondo suficiente como para comprender el contexto y determinar todas las variables explicativas correspondientes. Esto puede ser difícil y es vital para hacer muchas preguntas y ver, por ejemplo, si existen relaciones empíricas anteriormente conocidas entre las variables medidas. Se requiere generalmente un enfoque iterativo para la construcción de estos modelos. Por otro lado siempre se tiene la incertidumbre de la inclusión de variables explicativas incluidas innecesariamente en el modelo, que parecen mejorar el ajuste, pero en realidad llevan a predicciones erróneas, o la omisión de variables cruciales que se necesitan. Dicho de otra manera, es conveniente buscar un modelo parsimonioso (de modo que menos parámetros deben ser estimados), pero garantizando que las variables importantes estén dentro de nuestro modelo. Sin embargo, cualquiera que sea el modelo que se establezca finalmente, la inferencia estadística se realiza normalmente supeditada al modelo ajustado, bajo el supuesto de que el modelo es “verdadero”.

En resumen, en este capítulo vamos a estudiar la modelación conjunta de las relaciones entre los componentes de un vector de series temporales. Veremos que la metodología a utilizar es una generalización de las series temporales univariantes y que podemos utilizar herramientas similares para identificar el modelo, estimarlo y realizar contrastes de hipótesis.

2.2. Procesos vectoriales estacionarios y sus autocorrelaciones.

A menudo varias perturbaciones y fenómenos de ruido simultáneos influyen en un sistema dado. Esto hace necesario introducir procesos estocásticos vectoriales valuados (o valorados).

Un proceso estocástico es una familia de funciones temporales. Cada función la llamaremos: una realización del proceso. Supongamos que $v_1(t), v_2(t), \dots, v_n(t)$, son n procesos estocásticos escalares que pueden ser mutuamente dependientes.

Entonces llamaremos a: $v(t) = [v_1(t), v_2(t), \dots, v_n(t)]^T$, un proceso estocástico vectorial, supondremos que cada componente de $v(t)$ toma valores reales y que $t \geq t_0$ con t_0 dado.

Estudiaremos las propiedades que deben cumplir los procesos vectoriales para ser estacionarios, los cuales son una generalización de los procesos univariantes y para los cuales se utilizarán herramientas similares para describir su estructura, es por ello que se debe tener una relación muy estrecha con la teoría de procesos escalares, ya que son de vital importancia para la comprensión de la teoría vectorial que continuación se presenta.

2.2.1. Procesos vectoriales estacionarios.

Dado un vector de k series temporales $Z_t = (z_{1t}, \dots, z_{kt})'$ definiremos su esperanza como el vector obtenido tomando las esperanzas en cada uno de los componentes. Diremos que este vector de series, o este proceso vectorial, es estacionario si el vector de medias

$$E(Z_t) = \mu$$

es constante, y sus matrices de autocovarianzas definidas por:

$$\begin{aligned} \Gamma(m) &= E[(Z_{t-m} - \mu)(Z_t - \mu)'] \\ &= E[(Z_t - \mu)(Z_{t+m} - \mu)'] \\ &= \{\gamma_{ij}(m)\}, \end{aligned}$$

sólo dependen del retardo. Hemos llamado $\gamma_{ij}(m)$ al elemento ij de la matriz $\Gamma(m)$ de autocovarianzas de orden m .

$$\begin{aligned} \gamma_{ij}(m) &= E[(z_{it} - \mu_i)(z_{jt+m} - \mu_j)] \\ &\neq E[(z_{jt} - \mu_j)(z_{it+m} - \mu_i)] = \gamma_{ji}(m) \end{aligned}$$

Por ejemplo, para un vector bivalente, $Z_t = (z_{1t}, z_{2t})'$ las matrices $\Gamma(m)$ son cuadradas de orden dos y contienen en la diagonal la autocovarianza de orden m de cada componente y, fuera en la diagonal, las covarianzas cruzadas de orden m que se describen en el capítulo I. El término $\gamma_{12}(m)$ de esta matriz mide el efecto lineal de la

primera componente en t sobre la segunda en $t + m$. Análogamente, $\gamma_{21}(m)$ mide el efecto lineal de la segunda componente sobre la primera con retardo m . La covarianza cruzada $\gamma_{ij}(m)$ representa como afecta el componente z_i hoy al componente z_j en m periodos posteriores, es decir mide la relación que va de la variable z_{it} a la z_{jt+m} , mientras que $\gamma_{ji}(m)$ mide la relación de la causalidad del componente j sobre el i , que puede ser muy diferente al $\gamma_{ij}(m)$. Concluimos que las matrices $\Gamma(m)$ incluyen las covarianzas cruzadas entre los componentes del vector y no son, en general, simétricas para $m \neq 0$. En particular, la matriz $\Gamma(0)$ si es simétrica, ya que representa las varianzas de los componentes y las covarianzas instantáneas y será además definida positiva. Para las restantes matrices sus términos verifican, como vimos en el capítulo I, que:

$$\gamma_{ij}(m) = \gamma_{ji}(-m),$$

Por lo que sólo es necesario calcular las matrices $\Gamma(m)$ para retardos positivos, ya que existe la relación:

$$\Gamma(-m) = \Gamma'(m).$$

Por otra parte, se define el proceso de ruido blanco vectorial a_t como un proceso de media cero de variables incorreladas con matriz de varianzas y covarianzas instantáneas, Σ , definida positiva. Observemos que esta definición permite correlaciones contemporáneas entre los componentes del vector de ruido blanco, pero que no pueden existir ni autocorrelaciones en un componente, ni correlaciones cruzadas entre dos que no sean instantáneas. Si además, las variables a_t siguen un proceso normal multivariante, la incorrelación garantiza la independencia de las variables y hablaremos de un proceso de ruido blanco normal.

2.2.2. Matrices de autocorrelación simple y parcial.

Las matrices de autocorrelación del proceso vectorial se definen por:

$$R(m) = \{\rho_{ij}(m)\}$$

donde los términos de esta matriz vienen dados por:

$$\rho_{ij}(m) = \frac{\gamma_{ij}(m)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}},$$

Las matrices de autocorrelación generalizan las autocorrelaciones al caso vectorial y veremos que juegan un papel similar al caso univariante en la construcción del modelo. De la misma forma, podemos introducir las matrices de autocorrelación parcial del proceso. Definimos la matriz de autocorrelación parcial de orden m de un proceso estacionario vectorial:

$$\tilde{Z}_t = Z_t - \mu,$$

de esperanza cero como la matriz $P(m)$ obtenida ajustando por mínimos cuadrados del modelo lineal *multivariante*:

$$\tilde{Z}_t = \Phi_1 \tilde{Z}_{t-1} + \dots + \Phi_{m-1} \tilde{Z}_{t-m+1} + P(m) \tilde{Z}_{t-m} + u_t. \quad (2.1)$$

En este modelo cada serie se explica por una regresión sobre sus m valores pasados y los m pasados de las otras series. Como en el caso escalar, $P(m)$ representa la relación lineal existente entre el vector \tilde{Z}_t y el vector \tilde{Z}_{t-m} cuando eliminamos de ambos el efecto de las variables explicativas intermedias, $\tilde{Z}_{t-1}, \dots, \tilde{Z}_{t-m+1}$. La forma de (2.1) es la de un modelo lineal de regresión multivariante. Por ejemplo, tomemos $m=1$. La primera fila de la matriz $P(1)$, que denotaremos por $p_1'(1)$, se obtiene estimando por mínimos cuadrados la regresión múltiple:

$$\tilde{z}_{1t} = p_1'(1) \tilde{Z}_{t-1} + u_{1t},$$

en donde si $m=1$ se ve en la ecuación (2.1) que conlleva a la ecuación anterior, ya que sólo nos queda la parte de la matriz de autocorrelación con la variable retardada m periodos más el ruido blanco correspondiente.

Transponiendo la ecuación anterior, premultiplicando el resultado por \tilde{Z}'_{t-1} y tomando esperanzas, se tiene:

$$p_1(1) = E(\tilde{Z}'_{t-1} \tilde{Z}_{t-1})^{-1} E(\tilde{Z}'_{t-1} \tilde{z}'_{1t}) = \Gamma(0)^{-1} \gamma_1(1),$$

donde $\gamma_1(1)$ es el vector de covarianzas cruzadas entre el primer componente y todos los demás retardados un período, y coincide con la primera columna de la matriz Γ . En consecuencia, para todas las columnas:

$$P(1)' = \Gamma(0)^{-1} \Gamma(1),$$

Que es similar a la ecuación para estimar el parámetro en un $AR(1)$, pero sustituyendo las varianzas y autocovarianzas de orden uno por las matrices de autocovarianzas.

2.2.3. Estimación de las matrices simples y parciales.

Los elementos de las matrices de autocorrelación se estiman por la siguiente expresión:

$$\hat{\rho}_{ij}(m) = \frac{\sum_{t=1}^{T-m} (z_{it} - \bar{z}_i)(z_{jt+m} - \bar{z}_j)}{\sqrt{\sum_{t=1}^T (z_{it} - \bar{z}_i)^2 \sum_{t=1}^T (z_{jt} - \bar{z}_j)^2}},$$

donde $\bar{z}_i = T^{-1} \sum_{t=1}^n z_{it}$. Como en el caso univariante. Si tenemos un proceso de ruido blanco vectorial, los coeficientes estimados $\hat{\rho}_{ij}(m)$, para $m > 0$, tienen una distribución asintóticamente normal, con media cero y varianza T^{-1} . Por lo tanto consideraremos no

significativamente distintos de cero aquellos coeficientes menores, en valor absoluto, que $2/\sqrt{T}$.

Al aplicar mínimos cuadrados a la ecuación (2.1) se estiman los valores de las matrices de autocorrelación parcial. Llamando $\hat{P}(m)$ a las estimaciones de estas matrices, podemos construir un contraste de que las matrices poblacionales son iguales a cero, para ello, se comienza con $P(1)$ y se contrasta si es igual a cero. Si no lo es, se contrasta que $P(2)$ es igual a cero, condicionada a que $P(1)$ es distinta de cero, y así sucesivamente. Vamos a estudiar este contraste. Sea:

$$e_{t,m} = \tilde{Z}_t - \hat{\Phi}_1 \tilde{Z}_{t-1} - \dots - \hat{\Phi}_{m-1} \tilde{Z}_{t-m+1} - \hat{P}(m) \tilde{Z}_{t-m},$$

los residuos del ajuste al estimar el coeficiente de autocorrelación parcial de orden m , y sea

$$S(m) = \sum_{t=m+1}^T e_{t,m} e'_{t,m},$$

la matriz de suma de cuadrados de estos residuos. El contraste de razón de verosimilitud de que la matriz de correlación parcial de orden m es nula, condicionada a que esta matriz de orden $m-1$ no lo sea, depende de la matriz de varianzas y covarianzas de los residuos y que viene dada por:

$$\Lambda = \frac{|S(m)|}{|S(m-1)|},$$

donde $S(m)$ es la matriz de sumas de cuadrados de los residuos al ajustar un modelo de regresión multivariante con m regresores y $S(m-1)$ la obtenida si el orden máximo incluido es $m-1$. El contraste analiza si la reducción en la suma de cuadrados de los residuos es significativa. Bajo la hipótesis de que $P(m)$ es cero, la distribución de este estadístico puede aproximarse a una X^2 central mediante la transformación:

$$X^2 = -\left(T - \frac{1}{2} - mk\right) \log \Lambda,$$

que se distribuye asintóticamente como una X^2 con k^2 grados de libertad, el número de términos en la matriz $\hat{P}(m)$, lo que nos permite hacer un contraste de que esta matriz es cero.

2.3. Modelos ARMA vectoriales.

La generalización multivariante del teorema de Wold nos indica que todo proceso vectorial estacionario sin componentes deterministas admite una representación:

$$\tilde{Z}_t = \Psi(B) a_t = \sum_{i=1}^{\infty} \Psi_i a_{t-i}, \quad (2.2)$$

donde $\tilde{Z}_t = Z_t - \mu$, siendo μ el vector de medias de Z_t , B es el operador de retardo, que aplicado a vectores se define por $BZ_t = Z_{t-1}$, Ψ_i son matrices cuadradas $k \times k$, con $\Psi_0 = I$, la matriz unidad y a_t un vector de ruido blanco de media cero y matriz de covarianzas Σ definida positiva. Análogamente al caso univariante los modelos *ARMA* vectoriales surgen como representaciones libres del operador infinito $\Psi(B)$. Supondremos que esta matriz puede aproximarse por el producto de las matrices finitas:

$$\tilde{Z}_t = \Phi_p(B)^{-1} \Theta_q(B) a_t \quad (2.3)$$

donde:

$$\Phi_p(B) = \Phi_0 - \Phi_1 B - \dots - \Phi_p B^p$$

Es un operador en B y los términos Φ_i matrices cuadradas de orden k que contienen los coeficientes autorregresivos del proceso. Además que las raíces de la ecuación $|\Phi_p(B)| = 0$ deben estar fuera del círculo unitario para que el proceso sea estacionario. El operador de media móvil es:

$$\Theta_q(B) = \Theta_0 - \Theta_1 B - \dots - \Theta_q B^q$$

Y los ceros de la ecuación $|\Theta_q(B)| = 0$ deben estar fuera del círculo unitario para que el proceso sea invertible. Con esto se obtiene la representación *ARMA* multivariante:

$$\Phi_p(B) \tilde{Z}_t = \Theta_q(B) a_t. \quad (2.4)$$

Las matrices de un modelo *ARMA* vectorial deben cumplir determinadas condiciones para que el modelo sea identificable, es decir, exista una equivalencia uno a uno entre la representación (2.4) y la función de covarianzas del proceso. Por ejemplo los procesos:

$$\tilde{Z}_t = a_t - \Theta_1 a_{t-1}$$

y

$$\tilde{Z}_t = M u_t - \Theta_1^* u_{t-1}$$

donde $a_t = M u_t$, $\Theta_1^* = \Theta_1 M^{-1}$ son indistinguibles para cualquier matriz M no singular. Es necesario introducir condiciones para definir la estructura *ARMA* vectorial de manera única.

Las condiciones necesarias y suficientes son las siguientes:

- a. $\Phi_0 = \Theta_0 = I$ (matriz unidad) y Σ cualquiera.
- b. Φ_0 cualquiera, $\Theta_0 = I$ y Σ diagonal.
- c. Θ_0 cualquiera $\Phi_0 = I$ y Σ diagonal.
- d. $\Sigma = I$ y Φ_0 y Θ_0 cualquiera.

Estas formulaciones difieren en donde establecen la correlación contemporánea: en el caso (a) se pone en Σ ; en el caso (b) en la matriz Φ_0 ; en el caso (c) en la matriz Θ_0 ; y en el caso (d) se permite en ambas matrices Φ_0 y Θ_0 , pero se exige una condición de normalización, $\Sigma = I$.

En resumen si permitimos generalidad en la matriz de covarianzas del ruido, Σ , entonces las matrices Φ_0 y Θ_0 deben ser unitarias, mientras que si restringimos la matriz de covarianzas a ser diagonal, una de las matrices Φ_0 o Θ_0 puede ser cualquiera. En adelante supondremos, por analogía con el caso univariante, que $\Phi_0 = \Theta_0 = I$ y que la matriz Σ es definida positiva. A continuación estudiamos algunos casos particulares de estos procesos.

2.4. Proceso VAR (1).

Se define el proceso $VAR(1)$, autorregresivo vectorial de orden uno, mediante:

$$Z_t - \mu = \Phi(Z_{t-1} - \mu) + a_t.$$

En adelante, para simplificar la exposición, supondremos $k = 2$, y $\mu = 0$.

Entonces, podemos escribir el proceso como:

$$\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} B \right) \begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

Esta ecuación matricial da lugar a las dos ecuaciones:

$$z_{1t} = \phi_{11}z_{1t-1} + \phi_{12}z_{2t-1} + a_{1t}$$

y

$$z_{2t} = \phi_{21}z_{1t-1} + \phi_{22}z_{2t-1} + a_{2t}.$$

La condición de estacionariedad es que los ceros de la ecuación:

$$|I - \Phi B| = 0,$$

Estén fuera del círculo unitario. Como:

$$|I - \Phi B| = -B|\Phi - B^{-1}I| = -B|\Phi - \lambda I| = 0$$

Con $\lambda = B^{-1}$, tenemos que el proceso será estacionario si los valores propios de la matriz Φ están dentro del círculo unitario. En particular, si Φ es diagonal, tenemos dos procesos univariantes $AR(1)$ con ruidos correlados y el proceso será estacionario si $|\phi_{ii}| < 1$, que es la misma condición que en el caso univariante.

El modelo $VAR(1)$ determina los modelos univariantes de las series.

Escribiendo:

$$Z_t = (I - \Phi B)^{-1} a_t,$$

Y utilizando que la inversa de una matriz se calcula obteniendo la matriz adjunta y dividiendo por el determinante de la matriz, tenemos que para $k = 2$:

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \frac{1}{\alpha(B)} \begin{bmatrix} 1 - \phi_{22}B & \phi_{12}B \\ \phi_{21}B & 1 - \phi_{11}B \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix},$$

donde $\alpha(B) = ((1 - \phi_{11}B)(1 - \phi_{22}B) - \phi_{12}\phi_{21}B^2)$ que conduce a la siguiente ecuación, por ejemplo, para z_{1t} :

$$((1 - \phi_{11}B)(1 - \phi_{22}B) - \phi_{12}\phi_{21}B^2)z_{1t} = (1 - \phi_{22}B)a_{1t} + \phi_{12}a_{2t-1}.$$

Vemos que cada componente tiene la misma estructura AR , que será un polinomio de orden igual al número de componentes del vector: $AR(2)$ en este caso y un $AR(K)$ en general. La parte MA de cada componente será la suma de un ruido blanco y tantos proceso MA de orden igual al máximo de los órdenes de los sumandos, en el caso bivalente, la parte de la izquierda será $MA(1)$ y los componentes del proceso bivalente serán $ARMA(2, 1)$. Este es el orden máximo del proceso, porque es posible que exista cancelación entre los operadores y el modelo univariante sea de orden menor.

En el caso en que uno de los coeficientes ϕ_{ij} , $i \neq j$, sea cero, por ejemplo él $\phi_{12} = 0$, y Σ sea diagonal, entonces no existe relación desde el componente dos hacia el uno.

La variable z_{1t} se denomina exógena y la variable z_{2t} es la dependiente o endógena, y el modelo $VAR(1)$ puede escribirse como:

$$(1 - \phi_{11}B)z_{1t} = a_{1t} \quad (2.5)$$

$$z_{2t} = \frac{\phi_{21}}{1 - \phi_{22}B} z_{1t-1} + \frac{a_{2t}}{1 - \phi_{22}B}. \quad (2.6)$$

Vemos que el modelo VAR proporciona tanto el modelo de regresión dinámica entre las dos variables como el modelo univariante para la variable exógena.

Las funciones de autocovarianzas de un $VAR(1)$ se obtienen con:

$$\begin{aligned} \Gamma(1) &= E[(Z_t - \mu)(Z_{t+1} - \mu)'] \\ &= E[(Z_t - \mu)(\Phi(Z_t - \mu) + a_t)'] = \Gamma(0)\Phi' \end{aligned}$$

y aplicando la definición se obtiene:

$$\Gamma(m) = \Gamma(m-1)\Phi',$$

Que es la generalización del resultado univariante. De la definición de matrices de correlación parcial es inmediato que este proceso sólo tendrá la matriz de orden uno distinta de cero.

2.4.1 Proceso VAR (p).

El modelo generalizado es:

$$Z_t - \mu = \Phi_1(Z_{t-1} - \mu) + \dots + \Phi_p(Z_{t-p} - \mu) + a_t$$

y diremos que es estacionario si las raíces de la ecuación $|\Phi(B)|=0$ están fuera del círculo unitario. Tomando la transpuesta de esta ecuación y multiplicando por la derecha por $Z_{t-m} - \mu$ se obtiene:

$$(Z_{t-m} - \mu)(Z_t - \mu)' = (Z_{t-m} - \mu)(Z_{t-1} - \mu)' \Phi_1' + \dots + (Z_{t-m} - \mu)(Z_{t-p} - \mu)' \Phi_p' + (Z_{t-m} - \mu)a_t'$$

y tomando esperanzas, concluimos que las matrices de correlación verifican la relación:

$$\Gamma(m) = \sum_{i=1}^p \Gamma(m-i) \Phi_i' \quad m = 1, 2, 3, \dots \quad (2.7)$$

donde $\Gamma(m) = \Gamma'(-m)$, en efecto, para los retardos positivos, hacemos lo anterior, solamente que ahora multiplicamos por la derecha por $Z_{t+m} - \mu$ y transponemos toda la ecuación, y bajo el supuesto de que es un proceso vectorial estacionario, tomando el valor esperado obtenemos:

$$\Gamma'(-m) = E[(Z_{t+m} - \mu)(Z_t - \mu)'] = E[(Z_t - \mu)(Z_{t+m} - \mu)'] = E[(Z_{t-m} - \mu)(Z_t - \mu)'] = \Gamma(m)$$

Para un proceso *AR* general las matrices $\Gamma(m)$ serán distintas de cero para cualquier retardo, y su relación de dependencia temporal es la generalización natural de los procesos *AR* univariantes. Análogamente a dichos procesos un *VAR(P)* tendrá las *p* primeras matrices de correlación parcial distintas de cero y el resto serán nulas.

2.5. Proceso VMA (q).

La expresión de un modelo de media móvil vectorial es:

$$Z_t = (I - \Theta_1 B - \dots - \Theta_q B^q) a_t$$

Por ejemplo si $k=2$ y $q=1$, el modelo *MA(1)* bivalente es:

$$\begin{bmatrix} z_{1t} \\ z_{2t} \end{bmatrix} = \begin{bmatrix} 1 - \theta_{11} & -\theta_{12} B \\ -\theta_{21} B & 1 - \theta_{22} B \end{bmatrix} \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

y cada serie seguirá un modelo univariante que es un MA(1), ya que será la suma de un ruido blanco y un proceso MA(1). La condición de invertibilidad es que los ceros de la ecuación $|I - \Theta B|$ estén fuera del círculo unitario y es fácil comprobar que esta condición equivale para el VMA(1) a que los autovalores de la matriz Θ estén dentro del círculo unitario.

Las matrices de autocovarianzas se calculan con:

$$\begin{aligned}\Gamma(m) &= E[(Z_t - \mu)(Z_{t+m} - \mu)'] \\ &= E\left[\left(a_t - \Theta_1 a_{t-1} - \dots - \Theta_q a_{t-q}\right) x \left(a'_{t+m} - a'_{t+m-1} \Theta'_1 - \dots - a'_{t+m-q} \Theta'_q\right)\right]\end{aligned}$$

Que conducirá al resultado:

$$\begin{aligned}\Gamma(m) &= -\Sigma \Theta'_m + \Theta_1 \Sigma \Theta'_{m+1} + \dots + \Theta_{q-m} \Sigma \Theta'_q \quad m \leq q \\ \Gamma(m) &= 0 \quad m > q\end{aligned} \quad (2.8)$$

Por lo tanto, un proceso MA multivariante tendrá únicamente q matrices de autocovarianzas distintas de cero. De nuevo encontramos el paralelismo entre los modelos univariante y multivariante, que va a darse también para las matrices de autocorrelación parcial. Por ejemplo, consideremos el caso más simple $MA(1)$. En la hipótesis de que Z_t es invertible, la serie seguirá el proceso:

$$(I - \Theta_1 B)^{-1} Z_t = a_t,$$

Y utilizando la serie de Neumann del operador inverso, dada por:

$$(I - \Theta_1 B)^{-1} = I + \Theta_1 B + \Theta_1^2 B^2 + \dots,$$

El $VMA(1)$ puede también escribirse:

$$Z_t = -\Theta_1 Z_{t-1} - \Theta_1^2 Z_{t-2} - \Theta_1^3 Z_{t-3} - \dots + a_t.$$

Este es un proceso $VAR(\infty)$ y hemos generalizado el resultado de que un proceso MA equivale a un AR infinito. En consecuencia, sus matrices de autocorrelación parcial, $P(m)$ serán distintas de cero para cualquier retardo.

2.6. Procesos VARMA (p,q).

Hemos visto que las propiedades de los procesos $ARMA$ univariantes se generalizan inmediatamente para sus representaciones multivariantes.

Un proceso vectorial $VAR(p)$ tiene únicamente p matrices de autocorrelación parcial distintas de cero y un proceso $VMA(q)$ tiene q matrices de autocorrelación simple distintas de cero. En el caso de procesos mixtos $VARMA$ estas matrices, igual que el caso univariante, no tienen un corte en el orden del proceso.

Vamos a obtener la función de covarianzas para el proceso $ARMA$ general:

$$Z_t = \Phi_1 Z_{t-1} + \dots + \Phi_p Z_{t-p} + a_t - \Theta_1 a_{t-1} - \dots - \Theta_q a_{t-q}, \quad (2.9)$$

donde a_t es un proceso de ruido blanco con matriz de covarianzas Σ .

Transponiendo esta ecuación, premultiplicando el resultado por Z_{t-m} y tomando esperanzas, se tiene:

$$\Gamma(m) = \sum_{i=1}^p \Gamma(m-i) \Phi'_i - \sum_{j=0}^q E(Z_{t-m} a'_{t-j} \Theta'_j) \quad \text{con } \Theta_0 = -I \quad (2.10)$$

Para calcular el segundo miembro, observemos que, al ser el proceso estacionario, puede escribirse:

$$Z_t = \Phi(B)^{-1} \Theta(B) a_t = \Psi(B) a_t \quad (2.11)$$

donde a_t es ruido blanco con matriz de covarianzas Σ , implica que:

$$E(Z_{t-m} a'_{t-j}) = E(\Psi(B) a_{t-m} a'_{t-j}) = \begin{cases} 0 & j > m \\ \Sigma & j = m \\ \Psi_{j-m} \Sigma & j < m \end{cases} \quad (2.12)$$

Sustituyendo en la expresión (2.10), obtenemos que:

$$\Gamma(m) = \sum_{i=1}^p \Gamma(m-i) \Phi'_i - \sum_{j=1}^q \Psi_{j-m} \Sigma \Theta'_j, \quad 0 \leq m \leq q$$

$$\Gamma(m) = \sum_{i=1}^p \Gamma(m-i) \Phi'_i, \quad m > q$$

Que generalizan los resultados anteriores para modelos *VAR* y *VMA*.

2.7. Construcción de modelos VARMA para series estacionarias.

Dado un vector de k series estacionarias Z_t se quiere construir un modelo VARMA que describa su estructura de dependencias dinámicas. Supondremos que a priori, no especificamos que algunas variables son endógenas y otras exógenas y que se permite en consecuencia la existencia de realimentación entre todas ellas.

El método que se propone a continuación es similar al caso univariante y consta de las etapas siguientes:

1. Identificar la estructura del modelo, entendiendo por ello el orden del modelo VARMA estacionario.
2. Estimar los parámetros mediante máxima verosimilitud y seleccionar el mejor modelo entre los estimados.
3. Realizar contrastes para comprobar si el modelo seleccionado no tiene deficiencias detectables.

2.7.1 Identificación.

El objetivo de esta etapa es determinar el conjunto de modelos posibles que se ajustan al comportamiento de las variables en estudio. Cuando tengamos muchas series, la identificación requiere, en primer lugar, comprobar que la dimensión del sistema es realmente k , es decir, que tenemos k series independientes y que ninguna de ellas se ha construido como combinación lineal de las demás. Por ejemplo si tenemos un sistema trivariante con ingresos, z_{1t} , costes, z_{2t} , y beneficios, $z_{3t} = z_{1t} - z_{2t}$, no tenemos realmente tres series temporales sino dos.

Supongamos que existen h relaciones contemporáneas entre las series:

$$(Z_t - \bar{Z})'c_i = 0,$$

entonces tendremos únicamente $k-h$ componentes independientes. Esta situación exacta es poco frecuente con pocas series, pero con muchas series podemos encontrarnos con combinaciones lineales que son casi constantes o con muy poca variabilidad. Para contrastar este punto, comenzaremos calculando la matriz $\hat{\Gamma}(0)$. Como:

$$\hat{\Gamma}(0) = \frac{1}{T} \sum (Z_t - \bar{Z})(Z_t - \bar{Z})'$$

tendremos que:

$$\hat{\Gamma}(0)c_i = \frac{1}{T} \sum (Z_t - \bar{Z})(Z_t - \bar{Z})'c_i = 0,$$

lo que significa que la matriz $\hat{\Gamma}(0)$ debe tener $k-h$ autovalores iguales a cero. Si la relación es $(Z_t - \bar{Z})'c_i = u_i$, para h series u_i que tienen media cero y una varianza muy pequeña, tendremos que la matriz $\hat{\Gamma}(0)$ tendrá h autovalores muy pequeños. En lugar de modelar el vector de k series podemos tomar los vectores y valores propios de la matriz $\hat{\Gamma}(0)$ y modelar las $k-h$ combinaciones lineales que tienen alta varianza. Un aspecto a tener en cuenta es que cuando los componentes de las series no van en las mismas dimensiones conviene estandarizar desde el principio cada serie dividiéndola por su desviación típica para evitar problemas de escala. Esto equivale a obtener los componentes de la matriz de correlaciones estimadas, en lugar de las covarianzas.

Para determinar la estructura de matrices $\Phi(B)$ y $\Theta(B)$ de un vector de series estacionarias utilizamos las matrices estimadas de correlación simple y parcial. Si el orden m de las matrices de autocorrelación simple es tal que para $h > m$ las matrices $\hat{R}(h)$ tiene todos los elementos nulos y $|\hat{R}(h)|$ es alto (mayor que 3 ó 4), esto indica un proceso AR . Si el orden h es bajo (1,2 ó 3), la hipótesis más simple es que el proceso es $MA(h)$. Por lo tanto, la función de correlación simple nos permitirá diferenciar, en una primera instancia, entre procesos AR y MA . Las matrices de correlación parcial tienen la propiedad opuesta. Si hay unas pocas distintas de cero, lo que contrastaremos con el test de verosimilitudes, esto indica un proceso AR , mientras que matrices con coeficientes distintos de cero hasta retardos elevados sugieren un proceso MA .

Una información adicional importante para este análisis es el estudio de las matrices de autocorrelación simple de los residuos estimados en cada etapa, que pueden indicar la existencia de un modelo mixto. Por ejemplo, supongamos que el verdadero modelo, es $ARMA(1,1)$:

$$(I - \Phi B)Z_t = (I - \Theta B)a_t,$$

y erróneamente especificamos un $AR(1)$, dado por:

$$(I - \Phi^* B)Z_t = \varepsilon_t,$$

Entonces, los residuos ε_t tendrán estructura multivariante, indicando la posible existencia de un modelo mixto.

2.7.2 Estimación.

Una vez determinados los posibles órdenes p y q de las partes AR y MA , pasaremos a estimar el modelo:

$$(I - \Phi_1 B - \dots - \Phi_p B^p)Z_t = (I - \Theta_1 B - \dots - \Theta_q B^q)a_t.$$

Sea $m = \max(p, q)$. La función de verosimilitud condicionada a los valores iniciales del proceso Z_1, \dots, Z_m , de innovaciones a_{m+1}, \dots, a_T , es:

$$f_c(a_{m+1}, \dots, a_T) = \prod_{t=m+1}^T \frac{1}{|\Sigma|^{1/2} 2\pi^{k/2}} \exp\left\{-\frac{1}{2}(\mathbf{a}'_t \Sigma^{-1} \mathbf{a}_t)\right\},$$

dado que \mathbf{a}_t es $N_k(0, \Sigma)$. Tomando logaritmos y llamando β al vector de parámetros que incluye todos los AR y MA podemos escribir:

$$\log f_c(\beta) = -\frac{T-m}{2} \log |\Sigma| - \frac{1}{2} \sum_{t=m+1}^T (\mathbf{a}'_t \Sigma^{-1} \mathbf{a}_t).$$

En donde el valor de la constante $\frac{1}{2\pi^{k/2}}$ se ha quitado de la función de verosimilitud ya que no es significativa en la misma.

Utilizando que un escalar es igual a su traza, y $tr(ABC) = tr(BCA) = tr(CAB)$, tendremos que:

$$\begin{aligned} tr \sum_{t=m+1}^T \mathbf{a}'_t \Sigma^{-1} \mathbf{a}_t &= \sum_{t=m+1}^T tr(\mathbf{a}'_t \Sigma^{-1} \mathbf{a}_t) \\ &= tr\left(\Sigma^{-1} \sum_{t=m+1}^T \mathbf{a}_t \mathbf{a}'_t\right) \\ &= tr(\Sigma^{-1} S(\beta)). \end{aligned}$$

El procedimiento operativo para maximizar esta función es utilizar un algoritmo no lineal y seguir los mismos principios que el caso univariante. La matriz de covarianzas estimada vendrá dada por:

$$\hat{\Sigma} = \frac{1}{T-p} \sum_{t=m+1}^T \hat{\mathbf{a}}_t \hat{\mathbf{a}}_t'$$

Para decidir entre varios modelos estimados utilizaremos el criterio *BIC*. Si ajustamos un *VAR(p)*, el número de parámetros estimados es pk^2 y si ajustamos un *VARMA(p,q)* es $(p+q)k^2$, con lo que en este caso el criterio es:

$$BIC = \log |\hat{\Sigma}| + \frac{1}{T} (p+q)k^2 \log T.$$

En las experiencias de Monte Carlo, (es una técnica que combina una distribución de probabilidad con una serie de números pseudo-aleatorios para determinar el comportamiento futuro de una variable) el criterio *BIC* suele dar mejores resultados que el *AIC*, que tiende a sobreparametrizar. El criterio *AIC* viene dado por:

$$AIC = \log |\hat{\Sigma}| + \frac{2}{T} (p+q)k^2.$$

2.7.3 Diagnosis.

Los diagnósticos del modelo están diseñados para identificar posibles errores en la formulación del mismo y consisten en el análisis de los residuos estimados $\hat{\mathbf{a}}_t$ para comprobar si siguen un proceso de ruido blanco, lo que implica:

1. Cada componente a_{it} debe ser ruido blanco. Esto lo comprobaremos analizando las *fas* y *fap* de cada componente y utilizando los contrastes estudiados en el caso univariante.
2. Conjuntamente el vector $\hat{\mathbf{a}}_t$ sólo debe tener correlación contemporánea, lo que implica que tanto las matrices de autocorrelación $R_{at}(l)$, como las parciales $P_{at}(l)$, deben tener los elementos no significativamente distintos de cero para $l \geq 1$. Para fijar ideas, supongamos el modelo $\Phi(B)Z_t = \Theta(B)\hat{\mathbf{a}}_t$ y observamos que $\hat{\mathbf{a}}_t$ tiene estructura dinámica, del tipo:

$$\hat{\mathbf{a}}_t = \Psi(B)\mathbf{a}_t,$$

reformularemos el modelo mediante:

$$\Phi(B)Z_t = \Theta(B)\Phi(B)\mathbf{a}_t,$$

y volvemos a la etapa de estimación.

2.7.4. Contrastes en los modelos.

Contrastes de raíces unitarias.

En muchas ocasiones, el orden de integración de un proceso estocástico se puede determinar de manera bastante fiable a la vista del gráfico temporal y del correlograma de una serie generada por el proceso considerado.

En los casos en los que el orden de integración de un proceso no está claro a la vista de dichos instrumentos gráficos, los contrastes formales de raíces unitarias son el **contraste de Dickey Fuller- DF** o el de **Dickey Fuller Aumentado- DFA** de forma que puede ayudar a decidir el orden de integración en un modelo.

En cualquier caso, la determinación fiable del orden de integración de un proceso es importante en relación con la forma correcta de hacer inferencias.

2.7.4.1 Contraste de Dickey Fuller (DF) y Dickey Fuller Aumentado (ADF).

El primero de estos dos contrastes fue propuesto por Dickey Fuller (1979) Para el caso en que el proceso sea un paseo aleatorio bajo la H_0 y un proceso $AR(1)$ estacionario bajo la alternativa. Posteriormente, en el año 1981 lo amplían para el caso en que el proceso siga un esquema $AR(p)$ estacionario bajo la hipótesis alternativa. Esta generalización del anterior se conoce como contraste **Dickey Fuller Aumentado**.

Supongamos que x_t sigue un proceso $AR(p)$ sin término constante:

$$x_t = \sum_{i=1}^p \phi_i x_{t-i} + \varepsilon_t,$$

La ecuación característica del polinomio autorregresivo de x_t es:

$$\lambda^p - \sum_{i=1}^p \phi_i \lambda^{p-i} = 0$$

Siendo $\lambda_1, \lambda_2, \dots, \lambda_p$ las raíces características del proceso. Si $|\lambda_i| < 1$ entonces x_t converge a un proceso estacionario.

El contraste de dicha hipótesis en el caso de un $AR(1)$ se puede plantear mediante la estimación de:

$$x_t = \phi x_{t-1} + \varepsilon_t \quad (2.13)$$

Estableciendo la $H_0 : \lambda_1 = 1$ como $\phi = 1$, es decir, que $x_t \sim I(1)$ es integrada de orden 1, frente a la $H_a : \phi < 1$. Es, por tanto, un contraste a una cola.

Dicho contraste puede efectuarse con dos estadísticos: 1) el sesgo normalizado, $T(\hat{\phi} - 1)$, obtenido a partir de la estimación MCO (**Mínimos Cuadrados Ordinarios**)

y 2) con un contraste de la t de la estimación MCO de ϕ , siendo la H_0 el valor unitario de este parámetro.

En la práctica es habitual utilizar este último, aunque al ser más directa y cómoda el contraste de la hipótesis de significancia de un parámetro, se puede estimar por MCO. El modelo equivalente al anterior es el siguiente:

$$\Delta x_t = \alpha x_{t-1} + \varepsilon_t \quad (2.14)$$

donde $\alpha = (\phi - 1)$.

Por tanto, contrastar la hipótesis nula de existencia de una raíz unitaria ($\phi = 1$) en (2.13) equivale a contrastar $\alpha = 0$ en (2.14).

La hipótesis alternativa sería $H_a : \alpha < 0$. Por tanto, los valores críticos serán negativos, por lo que si se obtiene un valor inferior a esos valores críticos del estadístico de prueba se rechazará la hipótesis nula.

La estimación del parámetro ϕ (o equivalentemente α) es consistente pero sesgada. Sin embargo, la distribución de dicho estimador es distinta según el parámetro; la distribución del estimador es asintóticamente normal si $|\phi| < 1$. Debido a ello no podremos utilizar las tablas de la distribución t de Student, ya que bajo la $H_0 : \phi = 1$, t_α no sigue una distribución estándar. Se deberán utilizar las distribuciones empíricas de este estadístico tabuladas en **Fuller** (1976). En la misma referencia se encuentra tabulada la distribución del sesgo normalizado bajo la misma H_0 .

Además, la distribución del estimador de α no es independiente de la presencia de un término constante y/o de una tendencia determinista en la especificación de la ecuación del contraste. Por lo tanto, se deben considerar separadamente estas posibilidades. Así tendremos:

$$\left. \begin{array}{l} \text{a) } \Delta x_t = \alpha x_{t-1} + \varepsilon_t \\ \text{b) } \Delta x_t = \mu + \alpha x_{t-1} + \varepsilon_t \\ \text{c) } \Delta x_t = \mu + \beta t + \alpha x_{t-1} + \varepsilon_t \end{array} \right\} \quad (2.15)$$

Contrastándose la hipótesis nula $H_0 : \alpha = 0$ con los valores de t asociados a α en cada especificación: τ , τ_μ , τ_T respectivamente, cuya distribución, tanto asintótica como para distintos tamaños muestrales, aparece en Fuller (1976). En Dickey y Fuller (1981) se presentan los valores críticos sobre la significación individual de los parámetros μ y β en las especificaciones b) y c).

La especificación (2.15c) plantea la hipótesis nula de integrabilidad de primer orden frente a una hipótesis alternativa de proceso $AR(1)$ estacionario.

Por otro lado, la expresión (2.15c) es la forma reducida del modelo:

$$\begin{aligned} x_t &= \delta_0 + \delta_1 t + u_t \\ u_t &= \phi u_{t-1} + \varepsilon_t \end{aligned}$$

con $\mu = \delta_0(1-\phi) + \delta_1\phi$ y $\beta = \delta_1(1-\phi)$. Así, bajo la $H_0 : \phi = 1$, tenemos que $\mu = \delta_1$ y $\beta = 0$, por lo que se constata la irrelevancia de este último parámetro bajo la H_0 . El mismo tan solo se introduce en el contraste para hacerlo consistente frente a una hipótesis alternativa de estacionariedad sobre una tendencia determinista. A tal tipo de parámetros se les denomina parámetros molestos. Además las estimaciones de los tres parámetros del modelo (2.15c) estarán fuertemente correladas, al ser combinación de los parámetros estructurales $(\delta_0, \delta_1, \phi)$.

El modelo (2.15b) plantea la hipótesis nula de paseo aleatorio sin constante frente a una alternativa de esquema $AR(1)$ estacionario sin tendencia. Sería la forma reducida de:

$$\begin{aligned}x_t &= \delta_0 + u_t \\u_t &= \phi u_{t-1} + \varepsilon_t\end{aligned}$$

donde $\mu = \delta_0(1-\phi)$. También en este caso el parámetro μ es irrelevante bajo la hipótesis nula. El mismo se introduce para permitir que bajo la hipótesis alternativa el proceso autorregresivo tenga media no nula. Es, por tanto, un parámetro molesto.

Finalmente, el modelo (2.15a) contrasta la hipótesis nula de paseo aleatorio con $x_0 = 0$ frente a una alternativa de proceso autorregresivo estacionario con media nula. Dicha hipótesis nula es poco realista para la mayoría de las series temporales económicas.

Para efectuar el contraste conjunto de parámetros en las ecuaciones (2.15a) y (2.15c), **Dickey y Fuller** (1981) proponen la construcción de un test de la F que nos permite contrastar las hipótesis nulas.

$$\begin{aligned}H_0 : \mu = \alpha = 0 & \quad \text{en (2.15b)} \\H_0 : \mu = \beta = \alpha = 0 & \quad \left. \vphantom{H_0 : \mu = \beta = \alpha = 0} \right\} \text{ en (2.15c)} \\H_0 : \beta = \alpha = 0 & \quad \left. \vphantom{H_0 : \beta = \alpha = 0} \right\}\end{aligned}$$

Dichos contrastes de la F se denominan ϕ_1 , ϕ_2 y ϕ_3 respectivamente. De nuevo el problema se plantea cuando estos estadísticos no siguen distribuciones estándar. Para ello se tabulan valores críticos correspondientes mediante simulaciones, es decir se tienen valores propios para dicho contraste.

Debe tenerse en cuenta que los valores críticos en muestra finita de los distintos contrastes se han tabulado bajo el supuesto de que las perturbaciones se distribuyen normalmente con media cero y varianza constante. En cambio, los valores críticos asintóticos son válidos con carácter general. Por ello es mejor usar valores asintóticos e interpretar los resultados con precaución.

Por otro lado, la inclusión de variables ficticias estacionales en (2.15b) y (2.15c) no cambia la distribución asintótica de los estadísticos debido a que las mismas son del mismo orden que la constante introducida en las regresiones también señalan que las distribuciones asintóticas de los estadísticos no dependen de la asunción de que la varianza del término de perturbación sea constante.

Debido a las distintas implicaciones de los tres modelos sobre el comportamiento de la variable, algunos autores proponen seguir la estrategia de partir del modelo más general:

Examinar la ecuación (2.15c), el modelo más general. Si no rechazamos la hipótesis de raíz unitaria y β es no significativa, entonces examinar la ecuación (2.15b). Si la hipótesis nula de raíz unitaria no es rechazada y μ es no significativa entonces examinar la ecuación (2.15a).

Al plantear este contraste se está suponiendo que ε_t no está autocorrelado. Pero este supuesto no tiene porqué cumplirse, por lo que la inferencia en cualquiera de las tres ecuaciones planteadas se verá afectada. Se han propuesto dos tipos de solución a este problema:

1. Solución paramétrica. Sugerida por **Dickey Fuller** (1981) consiste en la inclusión en el test **DF** de una estructura de retardos de la variable dependiente que nos permita capturar la estructura autorregresiva de ésta, quedando la perturbación lo más incorrelada posible. Este contraste se conoce como test de *dickey Fuller Aumentado* (**ADF**) y consiste en estimar:

$$\Delta x_t = \mu + \beta t + \alpha x_{t-1} + \sum_{i=1}^{p-1} \gamma_i \Delta x_{t-1} + \varepsilon_t$$

con p lo suficientemente grande para garantizar que ε_t sea aproximadamente ruido blanco. La distribución asintótica de los parámetros $\hat{\mu}$, $\hat{\beta}$ y $\hat{\alpha}$ es la misma que en el **DF** e independiente de los parámetros $\hat{\gamma}_i$, los cuales siguen asintóticamente una distribución normal bajo la H_0 . La inclusión de los citados retardos dependerá de su significancia según el estadístico t contrastado.

Debe tenerse en cuenta también que un número excesivo de retardos reducirá la potencia del contraste, mientras que si no se especifican suficientes no se recogerá toda la autocorrelación residual, por lo que los valores críticos tabulados no serán aplicables.

2. Solución no Paramétrica: Propuesta por Philips (1897) y Philips Perron (1998). Sugieren transformar los estadísticos del test de **DF** para hacerlos compatibles con la presencia de autocorrelación y heterocedasticidad (varianza constante a lo largo del tiempo) en el término de perturbación. La idea es utilizar los residuos estimados ε_t en la regresión de **DF** para corregir el estadístico t asociado a los parámetros. De esta forma obtenemos unos nuevos estadísticos $z(t)$, $z(t_\mu)$ y $z(t_\tau)$ que tienen las mismas distribuciones que los estadísticos tabulados en **Fuller**.

Finalmente cabe mencionar que, la sobrediferenciación de una serie temporal llevará normalmente a valores altos y positivos del contraste de **DF** acompañado por un elevado R^2 de la regresión del contraste.

Capítulo III.

Aplicación de los modelos

Univariantes, Regresión

Dinámica y Multivariantes

de Series Temporales a

tasas de interés Bancarias.

3.1. Introducción.

En este capítulo se hará un estudio de la base de datos “Tasas de interés bancario”, la cual es de uso público y ha sido obtenido de la página web del Banco Central de Reserva de El Salvador (BCR), esta base de datos contiene 174 observaciones de 7 variables temporales, las cuales tienen un inicio en enero del año 1995 y finalizan en Junio del año 2009. Las observaciones pertenecen a 5 tasas de interés bancario, las cuales son: **FED (Tasa de interés que los bancos se cargan unos a otros por créditos overnight con fondos estatales)**, **LIBOR-180 (London Interbanking Offered Rate: tasa promedio de interés, día a día, en el mercado interbancario de Londres)**, **PRIME (Tasa de interés básica que aplican los principales bancos en EEUU para préstamos corporativos)**, **TIBP-180 (Tasa de interés básica pasiva nacional ofrecida por bancos a sus depositantes)**, **TPH1 (Tasa de Préstamo hasta un 1 año plazo nacional)**; donde cada tasa está dada en porcentajes. Y las otras dos series temporales son: **IGInflación (Índice de crecimiento de precios o índice general de inflación)** e **IVAE (Índice de volumen de la actividad económica nacional)**.

El estudio de las variables antes descritas se ha realizado de acuerdo al siguiente orden:

- a) Análisis de series temporales univariantes, en donde ajustaremos un modelo ARIMA a cada una de las series temporales en estudio, haciendo uso de las funciones de autocorrelación simple (FAS) y parcial (FAP). Se establece el mejor modelo a cada una de las series utilizando el paquete estadístico SPSS.
- b) Relación entre las series temporales con el fin de construir modelos de regresión dinámica, donde la metodología a utilizar para la construcción de dichos modelos se dividirá en 3 etapas que son: identificación, estimación y diagnóstico.
- c) Modelos de series temporales multivariantes, los cuales nos sirven para realizar una modelación conjunta de las relaciones entre los componentes de un vector de series temporales. La metodología a utilizar es una generalización de las series temporales univariantes y que podemos utilizar herramientas similares para identificar el modelo, estimarlo y realizar contrastes diagnósticos.

Los programas o software que se utilizarán para realizar los análisis planteados anteriormente son: **SPSS versión 15.0** y **EViews versión 6.0**.

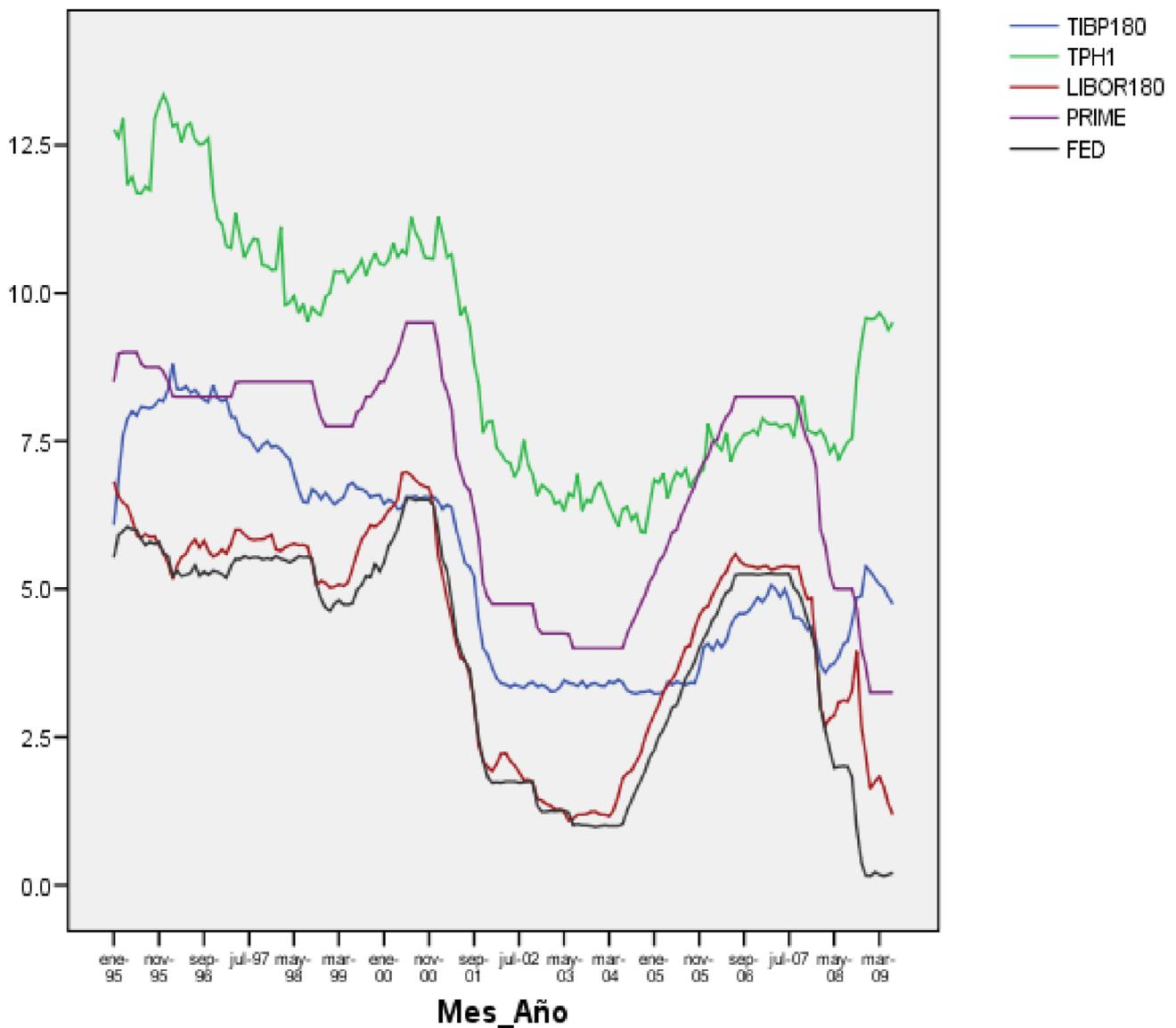
3.2. Análisis de Series Temporales Univariante.

3.2.1 Análisis de Estacionariedad y estacionalidad.

Como habíamos visto en el capítulo uno, el primer paso para establecer un modelo univariante de una serie temporal en estudio, es verificar la Estacionariedad de dicha serie, es decir, que ni su media, ni la varianza, ni las autocorrelaciones dependen del tiempo. En caso de que la serie no sea estacionaria, entonces procedemos a estabilizarla mediante transformaciones adecuadas.

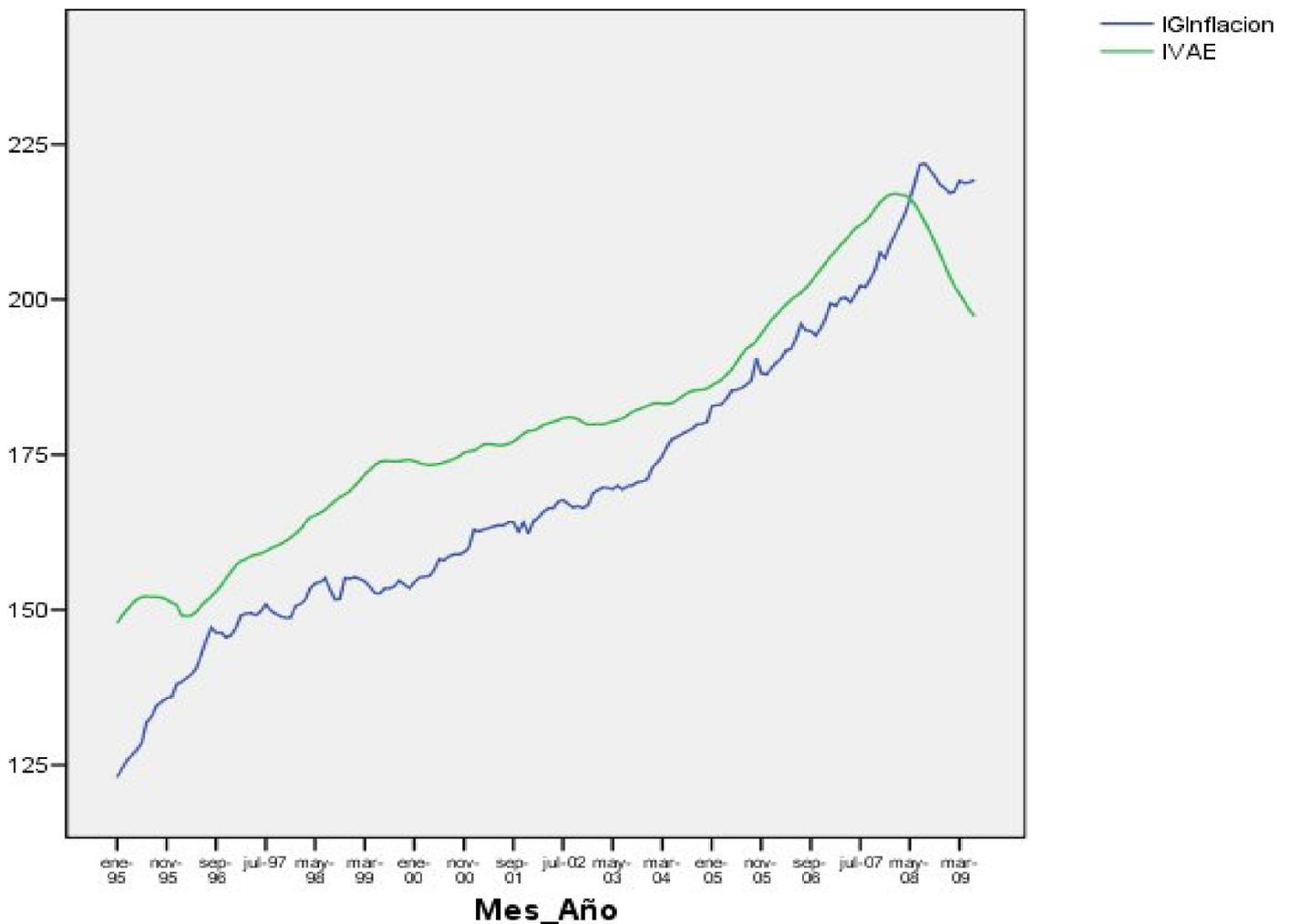
A continuación se presentan los gráficos de secuencias de las 7 series temporales en estudio:

Gráfico 1. Series temporales tasas de interés bancario.



En el gráfico anterior, observamos que las tasas de interés bancarias tienen bastante similitud a lo largo del tiempo, aparentemente, no son estacionarias en media debido a que presentan valores aleatorios con crecimiento y decrecimiento en ciertos periodos; la variabilidad es constante a lo largo del tiempo, por lo que se conjetura la estacionariedad en varianza; esto es esperable, ya que las tasas de interés que establecen los bancos en todo el mundo tienen una cota superior e inferior, por ejemplo, una tasa de interés nunca puede tener un valor negativo o ser cero, ya que los bancos estarían perdiendo sus ganancias, por tanto la serie siempre será positiva; y tampoco puede ir teniendo un crecimiento rápido hacia valores muy grandes, ya que entonces nadie utilizaría los bancos para realizar préstamos ni depósitos con tasas de interés altísimas.

Gráfico 2. Series temporales Índice General de Inflación(IGInflacion) e IVAE.



En el gráfico de secuencia de los 2 índices en estudio, se observa que tienen una clara tendencia de crecimiento, por lo tanto, podemos decir que las series no son estacionarias en media, pero sus valores no varían mucho con respecto al anterior a él, entonces podríamos pensar que los 2 índices son estacionarios en varianza.

A partir de los gráficos de secuencias de las series de tiempo en estudio, podemos suponer que ninguna de las series presenta estacionalidad, ya que en ninguno de los gráficos se observa que la serie se comporte de forma similar cada cierto periodo de tiempo; además supondremos que nuestras series poseen raíces unitarias, ya que no son estacionarias en media, y necesitan una diferenciación, para comprobar esta hipótesis realizamos el test aumentado de Dickey-Fuller, utilizando el programa EViews, se obtuvieron los siguientes resultados:

Tabla 1. Test de Dickey-Fuller aumentado para la serie TIBP180.

Null Hypothesis: TIBP180 has a unit root
 Exogenous: None
 Lag Length: 10 (Automatic based on AIC, MAXLAG=13)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.463255	0.1337
Test critical values:		
1% level	-2.579226	
5% level	-1.942793	
10% level	-1.615408	

*MacKinnon (1996) one-sided p-values.

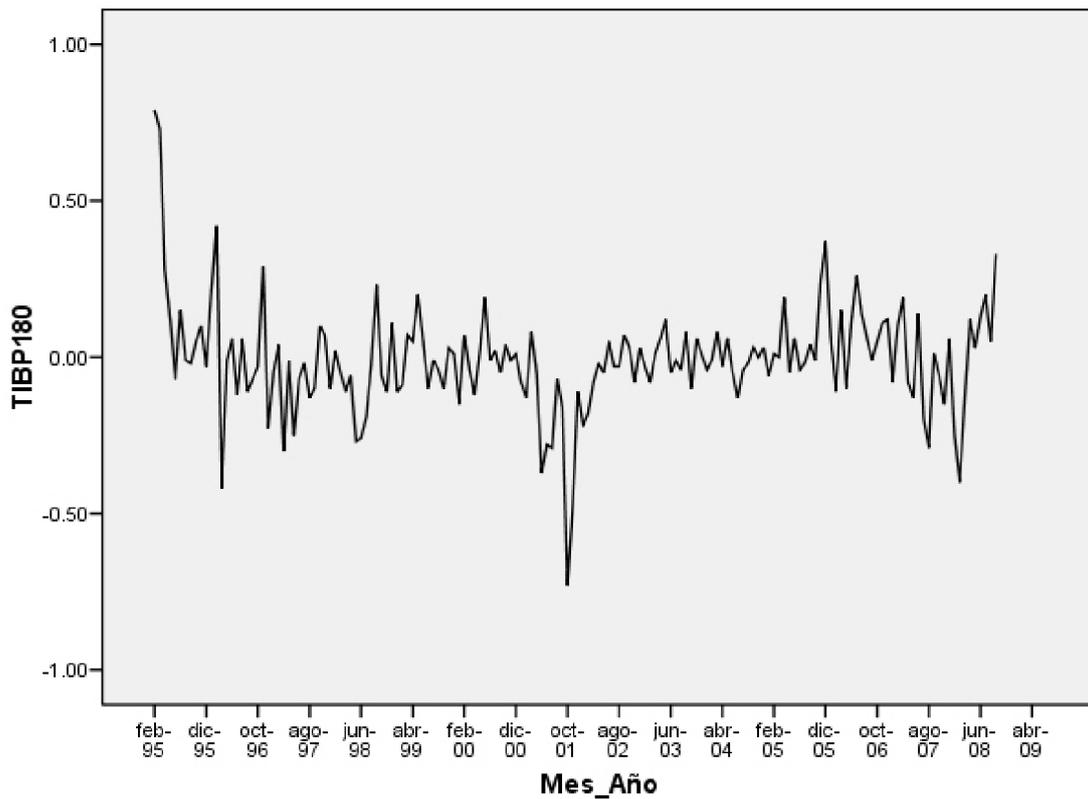
Los resultados del test de raíces unitarias de Dickey-Fuller aumentado nos dice que todas las series en estudio presentan una raíz unitaria, por ejemplo, para la serie temporal TIBP180 se obtuvo un valor del estadístico $t(-1.463255)$ mayor a los valores críticos de la distribución DF $(-2.579226, -1.942793, -1.615408)$, por lo tanto, aceptamos H_0 a un nivel de confianza del 1%, 5% y 10%, es decir que la serie posee una raíz unitaria. Para las demás series temporales en estudio, el resultado es similar, con lo cual podemos decir que todas las series no son estacionarias en media, entonces necesitamos diferenciarlas una vez para transformarlas en estacionarias, pero, en la práctica a veces esto no es necesario, como veremos más adelante.

3.2.2 Identificación de los posibles modelos.

Después de haber realizado el análisis de estacionariedad y estacionalidad de cada una de las series, pasaremos a la identificación de cuál es el mejor modelo ARIMA univariante para cada una de ellas, para efectos de comparación y predicción, nosotros dejaremos el 5% de los datos originales, es decir que ahora todas las series constan de 165 observaciones, que corresponden al periodo de enero de 1995 hasta septiembre del 2008.

Para la identificación de los posibles modelos, se comenzará analizando las funciones de autocorrelación simple y parcial, para las series diferenciadas, ya que según el test de Dickey-Fuller las series presentan raíces unitarias con lo cual no eran estacionarias en media. A continuación presentamos los gráficos de la FAS y la FAP de las series con una diferenciación:

Gráfico 3. Serie temporal TIBP180 con una diferencia.



Transformaciones: diferencia(1)

Gráfico 4. FAS de la serie temporal TIBP180 con una diferencia.

TIBP180

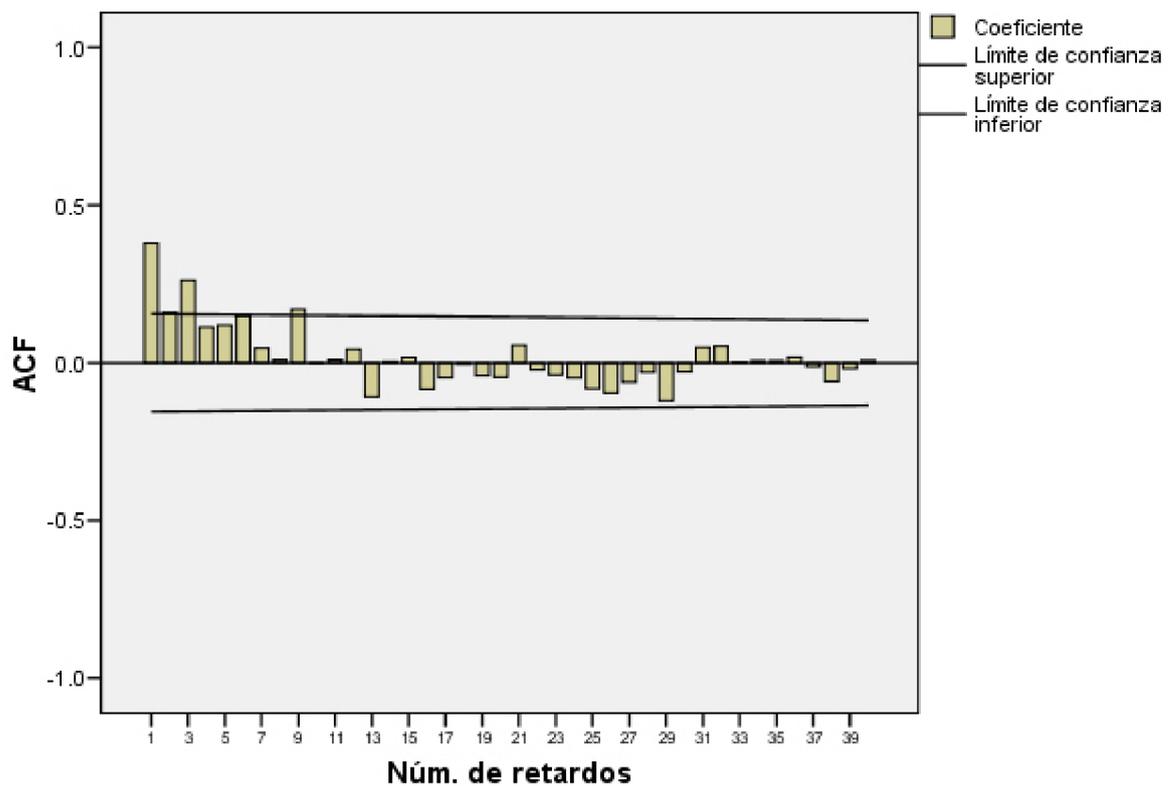
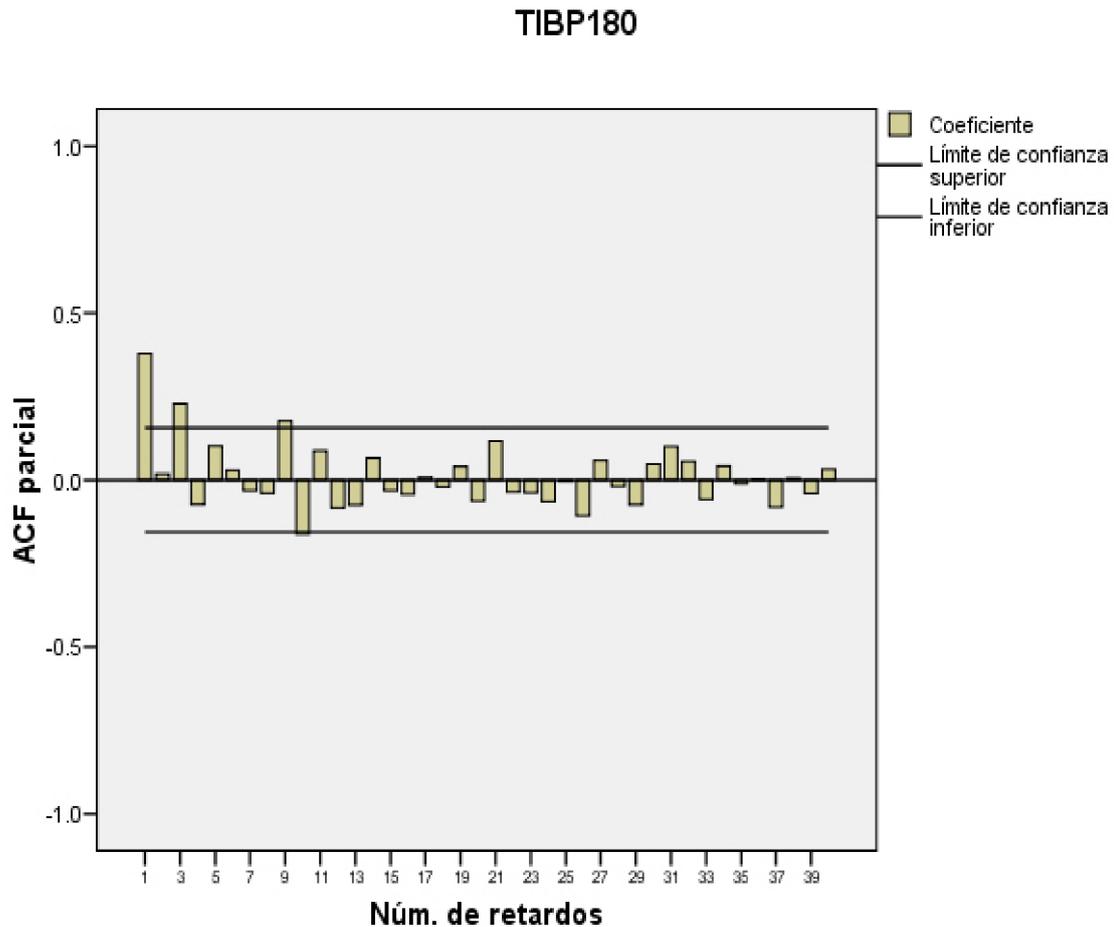


Gráfico 5. FAP de la serie temporal TIBP180 con una diferencia.



Después de realizarle una diferencia a la serie TIBP180, podemos ver en su gráfico de secuencia que ahora es estacionaria en media y varianza, con respecto a su FAS tiene 2 picos significativos en los primeros retardos y 2 en su FAP, con lo que podemos pensar que los posibles modelos son: ARIMA(1,1,0), ARIMA(2,1,0), ARIMA(0,1,1), ARIMA(0,1,2), ARIMA(1,1,1), ARIMA(1,1,2) o ARIMA(2,1,2). Se realizó el mismo análisis para el resto de las series, en donde, para TPH1 y el IGInflacion, sus modelos son obvios, ya que en sus FAS y FAP no presentan picos significativos, por tanto, podemos decir que los posibles modelos son procesos de ruido blanco; las demás series presentan en su FAS un decrecimiento amortiguado hacia cero y en su FAP se observan uno o dos picos significativos, a continuación se presentan los posibles modelos para cada una de ellas:

Tabla 2. Resumen de los posibles modelos ARIMA para ser ajustados a las series TPH1, LIBOR180, PRIME, FED, IGInflacion e IVAE.

	Modelo Identificado	Modelo Reducido
TPH1	ARIMA(1,0,0)	AR(1)
	ARIMA(0,1,0)	I(1)
LIBOR180	ARIMA(1,1,0)	ARI(1,1)
	ARIMA(2,1,0)	ARI(2,1)
PRIME	ARIMA(1,1,0)	ARI(1,1)
	ARIMA(2,1,0)	ARI(2,1)
FED	ARIMA(1,1,0)	ARI(1,1)
	ARIMA(2,1,0)	ARI(2,1)
IGInflacion	ARIMA(1,0,0)	AR(1)
	ARIMA(0,1,0)	I(1)
IVAE	ARIMA(1,1,0)	ARI(1,1)

Como podemos ver en la tabla 2, varias de las series presentan la misma estructura ARIMA en sus posibles modelos a establecer, el proceso a seguir para ajustarlos es parecido, por lo tanto, seguiremos la metodología de comparación de modelos estimados solamente para la serie TIBP180, ya que en su FAS y FAP es la que presenta más picos significativos. Para las series THP1 y IGInflacion por lo visto en sus FAS y FAP sabemos cuál es el mejor modelo ARIMA que se les ajustará, y para el resto de las series solo se presentará el mejor modelo que se ajustó a cada una de ellas. Todas las estimaciones se harán con la ayuda del programa o software SPSS.

3.2.3 Estimación de los parámetros de los modelos identificados, diagnosis de modelos y predicción.

Una vez identificados los tipos de modelos que pueden tener cada una de nuestras series temporales, primero estimaremos los parámetros de los modelos, luego haremos un diagnosis de cuál es el mejor haciendo uso de los estadísticos BIC y R^2 el cual nos indica el porcentaje que explica cada uno de los modelos ajustados a la serie, además utilizaremos los gráficos de la FAS y FAP residuales para saber que tan bien explica el modelo estimado a la serie en estudio, y ver si sus residuos están incorrelados; finalmente haremos predicciones con cada modelo estimado para verificar su idoneidad.

3.2.3.1 Ajuste de un modelo ARIMA a la serie TIBP180.

Estimación de los parámetros de los modelos identificados.

Los modelos identificados anteriormente para esta series eran del tipo ARIMA(p,d,q), en donde d=1, p=1,2, y q=1,2. Después de hacer un análisis previo de los posibles modelos a estimar, realizaremos la comparación solamente para los 2 mejores modelos encontrados de todos los posibles, que son: ARIMA(1,1,0), ARIMA(1,1,2).

Pasaremos a estimar los coeficientes para cada uno de los modelos antes mencionados:

Modelo ARIMA(1,1,0).

Tabla 3. Parámetros del modelo ARIMA(1,1,0) ajustado a la serie TIBP180.

				Estimación	ET	t	Sig.
TIBP180-Modelo_1	TIBP180	Sin transformación	Constante	-.004	.022	-200	.842
			AR Retardo 1	.445	.071	6.242	.000
			Diferencia	1			

Analizando la tabla 3 podemos decir que la estimación de la constante en el modelo no es significativa ya que tiene un p-valor = 0.842 que es mayor a 0.05, por otro lado el primer retardo AR del modelo diferenciado ajustado si es significativo en el modelo ya que su nivel de significancia es menor que 0.05, y el valor de $|\phi_1| = |0.445| < 1$ con lo cual podemos decir que el modelo es estacionario, e invertible ya que una de las características de los modelos AR es que siempre son invertibles, entonces el modelo ajustado para la serie es:

$$w_t = 0.445w_{t-1} + \hat{u}_t$$

$$\rightarrow y_t - y_{t-1} = 0.445y_{t-1} - 0.445y_{t-2} + \hat{u}_t$$

$$y_t = 1.445y_{t-1} - 0.445y_{t-2} + \hat{u}_t$$

$$(1 - 1.445L + 0.445L^2)y_t = \hat{u}_t$$

Modelo ARIMA(1,1,2).

Tabla 4. Parámetros del modelo ARIMA(1,1,2) ajustado a la serie TIBP180.

				Estimación	ET	t	Sig.
TIBP180-Modelo_1	TIBP180	Sin transformación	Constante	.015	.042	.351	.726
			AR Retardo 1	.903	.069	13.046	.000
			Diferencia	1			
			MA Retardo 1	.462	.112	4.118	.000
			Retardo 2	.193	.096	2.009	.046

Observando la tabla 4 podemos ver que la estimación de la constante en el modelo no es significativa ya que tiene un p-valor = 0.726 que es mayor a 0.05, mientras que las estimaciones tanto del retardo 1 de la parte autorregresiva como los dos retardos de la parte de medias móviles son significativos, ya que cada uno tiene un valor de significancia menor que 0.05, además el modelo es estacionario e invertible, pues los parámetros estimados cumplen con: $|\phi_1| = |0.903| < 1$ y $|\theta_1| = |0.462| < 1$, $|\theta_2| = |0.193| < 1$,

$\theta_1 + \theta_2 = 0.462 + 0.193 = 0.655 < 1$, $\theta_1 - \theta_2 = 0.462 - 0.193 = 0.269 < 1$; entonces el modelo ajustado es:

$$w_t = 0.903w_{t-1} + \hat{u}_t - 0.462u_{t-1} - 0.193u_{t-2}$$

$$\rightarrow y_t - y_{t-1} = 0.903y_{t-1} - 0.903y_{t-2} + \hat{u}_t - 0.462u_{t-1} - 0.193u_{t-2}$$

$$y_t = 1.903y_{t-1} - 0.903y_{t-2} + \hat{u}_t - 0.462u_{t-1} - 0.193u_{t-2}$$

Diagnosis de los modelos.

Utilizando el criterio BIC, el estadístico R^2 y los gráficos FAS y FAP residuales, se encontró que el mejor modelo era el ARIMA(1,1,0), a continuación se presentan dichos estadísticos:

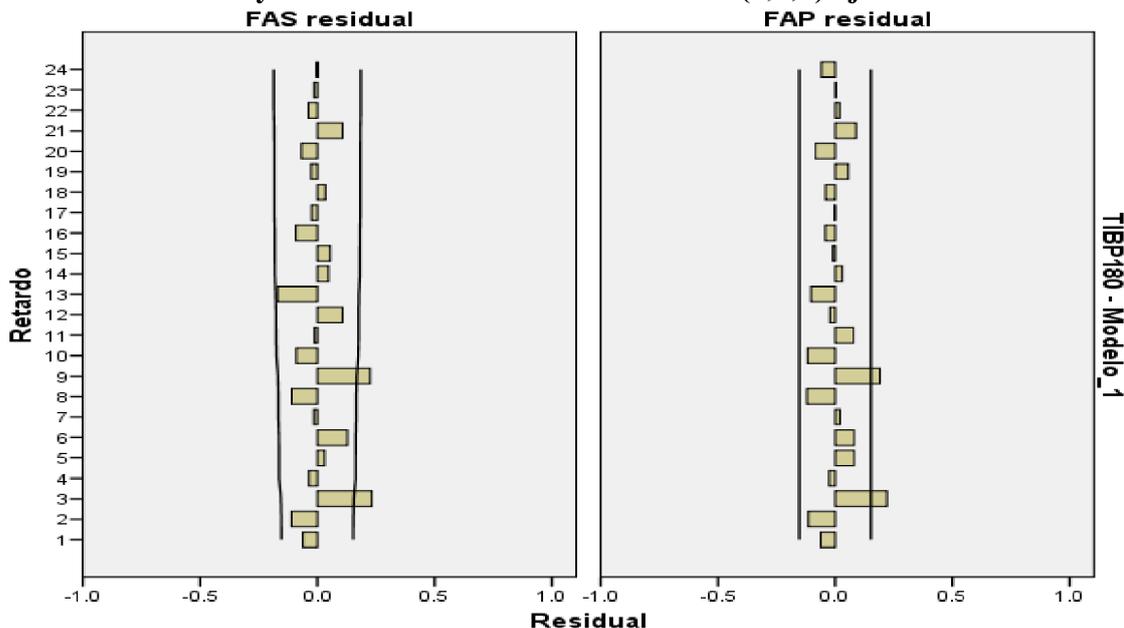
Modelo ARIMA(1,1,0).

Tabla 5. Estadísticos del modelo ARIMA(1,1,0) ajustado a la serie TIBP180.

Modelo	Número de predictores	Estadísticos de ajuste del modelo			Ljung-Box Q(18)			Número de valores atípicos
		R-cuadrado estacionaria	R-cuadrado	BIC normalizado	Estadísticos	GL	Sig.	
TIBP180-Modelo_1	0	.145	.992	-3.584	37.521	17	.003	0

En la tabla 5 podemos ver que el modelo ARIMA(1,1,0) explica el 99.2% de la tasa de interés ofrecida por bancos nacionales a sus depositantes, ya que el valor de su estadístico $R^2 = 0.992$, por otro lado su estadístico BIC es bastante pequeño y negativo con un valor de -3.584, mientras que los estadísticos del modelo ARIMA(1,1,2) son: $R^2 = 0.992$ y BIC = -3.538, por lo tanto, podemos concluir que el mejor modelo que ajusta a la serie TIBP180 es el ARIMA(1,1,0).

Gráfico 6. FAS y FAP residual del modelo ARIMA(1,1,0) ajustado a la serie TIBP180.



En la FAS y FAP residuales podemos observar que el 95% de los retardos no son significativos, ya que solamente dos son significativos, con lo cual podemos decir que los residuos están incorrelados, y se aproximan a un comportamiento de ruido blanco.

Predicción de los modelos.

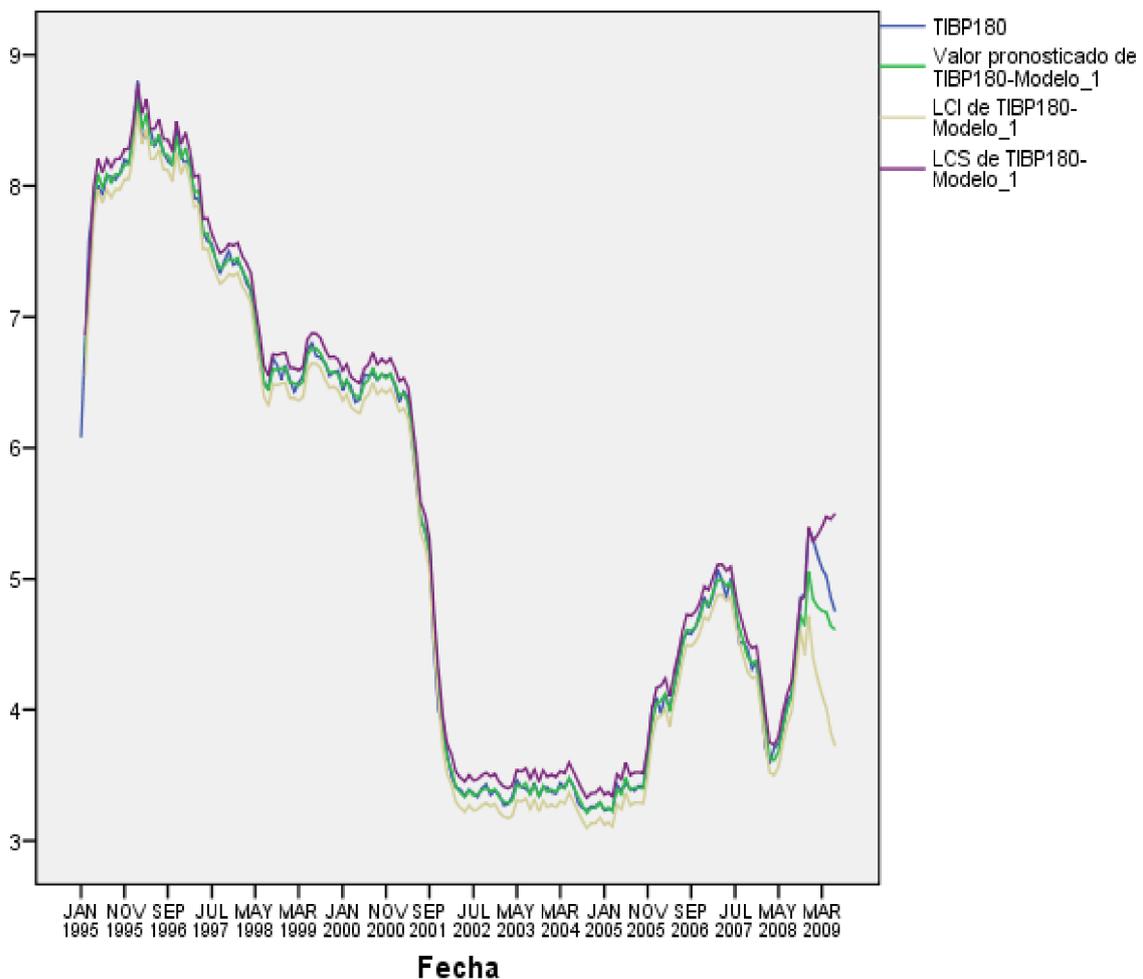
En la etapa anterior de diagnóstico se estableció que el mejor modelo para la serie TIBP180 es un ARIMA(1,1,0), entonces haremos predicciones con este modelo para corroborar que ajusta bien a nuestra serie en estudio, a continuación se presenta la tabla de predicciones, así como su gráfico:

Tabla 6. Predicciones del modelo ARIMA(1,1,0) ajustado a la serie TIBP180.

		Previsión									
Modelo		Oct 2008	Nov 2008	Dic 2008	Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	
	Valor Observado	4.86	4.88	5.38	5.3	5.18	5.08	5.02	4.86	4.75	
TIBP180-Modelo_1	Previsión	4.59	4.66	4.68	4.69	4.69	4.69	4.69	4.68	4.68	
	LCS	4.91	5.21	5.44	5.62	5.78	5.91	6.03	6.14	6.24	
	LCI	4.28	4.10	3.93	3.76	3.61	3.47	3.34	3.22	3.11	

Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles

Gráfico 7. Predicciones del modelo ARIMA(1,1,0) ajustado a la serie TIBP180.



Podemos observar en la tabla y gráfico de predicción, que los valores predichos no distan mucho de los valores originales, y además estos siempre caen dentro de los límites (LCS y LCI) que se han calculado con un 95% de confianza, con lo cual podemos confirmar que este modelo es el que mejor ajusta a la serie TIBP180.

3.2.3.2 Ajuste de un modelo ARIMA a las series restantes.

La metodología que se utilizó para establecer los modelos univariantes de series temporales para las series restantes, fue el mismo que ya se presentó para establecer el mejor modelo ARIMA a la serie TIBP180, como habíamos visto en la tabla 2, los modelos que se ajustaron son bastante similares en su estructura ARIMA, y se presentan a continuación:

Tabla 7. Resumen de los modelos ARIMA ajustados a las series TPH1, LIBOR180, PRIME, FED, IGInflacion e IVAE.

	Modelo Identificado	Modelo Reducido	Expresión
TPH1	ARIMA(1,0,0)	AR(1)	$y_t = 9.815 + 0.991y_{t-1} + \hat{u}_t$ $\rightarrow (1 - 0.991L)y_t = 9.815 + \hat{u}_t$
LIBOR180	ARIMA(1,1,0)	ARI(1,1)	$w_t = 0.608w_{t-1} + \hat{u}_t$ $\rightarrow (1 - 1.608L + 0.608L^2)y_t = \hat{u}_t$
PRIME	ARIMA(1,1,0)	ARI(1,1)	$w_t = 0.628w_{t-1} + \hat{u}_t$ $\rightarrow (1 - 1.628L + 0.628L^2)y_t = \hat{u}_t$
FED	ARIMA(1,1,0)	ARI(1,1)	$w_t = 0.647w_{t-1} + \hat{u}_t$ $\rightarrow (1 - 1.647L + 0.647L^2)y_t = \hat{u}_t$
IGInflacion	ARIMA(0,1,0)	I(1)	$w_t = 0.597 + \hat{u}_t$ $\rightarrow (1 - L)y_t = 0.597 + \hat{u}_t$
IVAE	ARIMA(1,1,0)	ARI(1,1)	$w_t = 0.902w_{t-1} + \hat{u}_t$ $\rightarrow (1 - 1.902L + 0.902L^2)y_t = \hat{u}_t$

3.3. Análisis de Modelos de Regresión Dinámica.

El análisis de los modelos de regresión dinámica se hará con el fin de establecer modelos dinámicos a las tasas de interés bancarias nacionales (TIBP180 y TPH1) en función de las tasas de interés internacional (LIBOR180, PRIME, FED) y los dos índices en estudio (IVAE e IGINflacion), para establecer este tipo de modelos entre variables estacionarias se contemplan tres fases al igual que en la estimación de cualquier otro modelo, las cuales son: **identificación, estimación y diagnóstico.**

La fase de identificación nos ayudará a decidir la estructura de la función de transferencia que utilizarán las variables independientes y el modelo para la perturbación. En la estimación una vez establecida la función de transferencia y el modelo para la perturbación asumiendo hipótesis de normalidad, pasaremos a estimar los parámetros de los modelos identificados, y finalmente se hará un diagnóstico con los estadísticos BIC y R^2 para establecer si el modelo aun se puede mejorar, o tiene variables sobrantes.

3.3.1 Identificación de los posibles modelos de regresión dinámica.

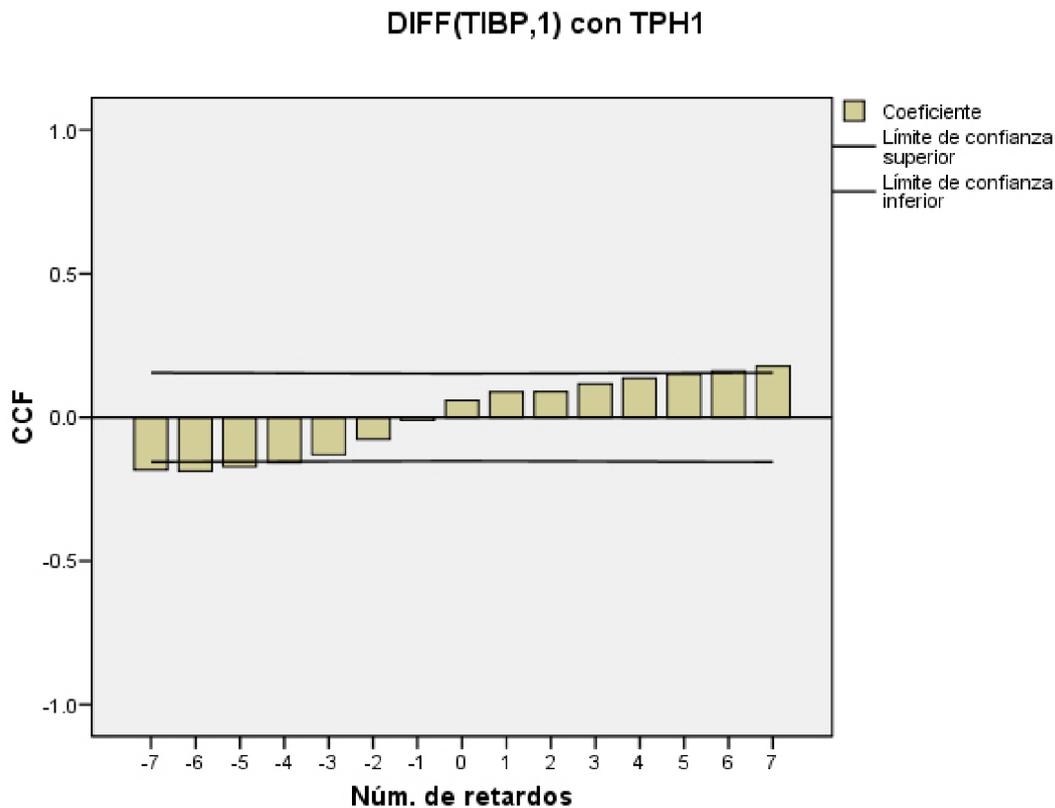
Para comenzar el análisis de los modelos de regresión dinámica, es importante que nosotros tengamos en cuenta la colinealidad que existe entre las series de tiempo en estudio, ya que esto podría suponer que las variables tiene una relación instantánea y simple, lo cual podría afectarnos en los modelos de regresión dinámica que nosotros esperamos probar, entonces previamente haremos un análisis de la función de transferencia entre las series en estudio con la ayuda de la función de correlación cruzada.

3.3.1.1 Análisis de función de correlación cruzada.

El análisis de las Funciones de covarianzas cruzadas entre pares de series estacionarias, depende del orden en que se toman las variables y de los instantes en que se consideran cada una de ellas. La definición⁸ establece que la variable que aparece en primer lugar se supone en el instante t y la segunda en $t+h$, cuando $h > 0$ mide la relación que va desde la primera variable hacia la segunda, mientras que para $h < 0$ la relación que mide es la que va desde la segunda variable hacia la primera, una de las limitaciones de esta función es que no nos permite identificar directamente si la relación es en una dirección o bidireccional, pero si nos puede dar una noción de la relación. Ya que esta función utiliza series estacionarias, nosotros usaremos nuestras series con una diferencia a excepción de la serie TPH1 ya que como vimos en el ajuste de los modelos ARIMA esta serie no necesita ser diferenciada para ser estacionaria, para facilitar el análisis de esta función se llamará a las series de la siguiente manera: $LIBOR180 = w1_t$, $TPH1_t$, $LIBOR180 = w3_t$, $PRIME = w4_t$, $FED = w5_t$, $IGInflacion = w6_t$, e $IVAE = w7_t$, donde el subíndice t indica el instante en el que se encuentra cada una de las variables, los resultados que se presentan a continuación fueron obtenidos con el SPSS:

⁸ Para más detalles del tipo de relación que existe entre variables, puede revisar el apartado 1.3.2.1 Función de covarianzas cruzadas.

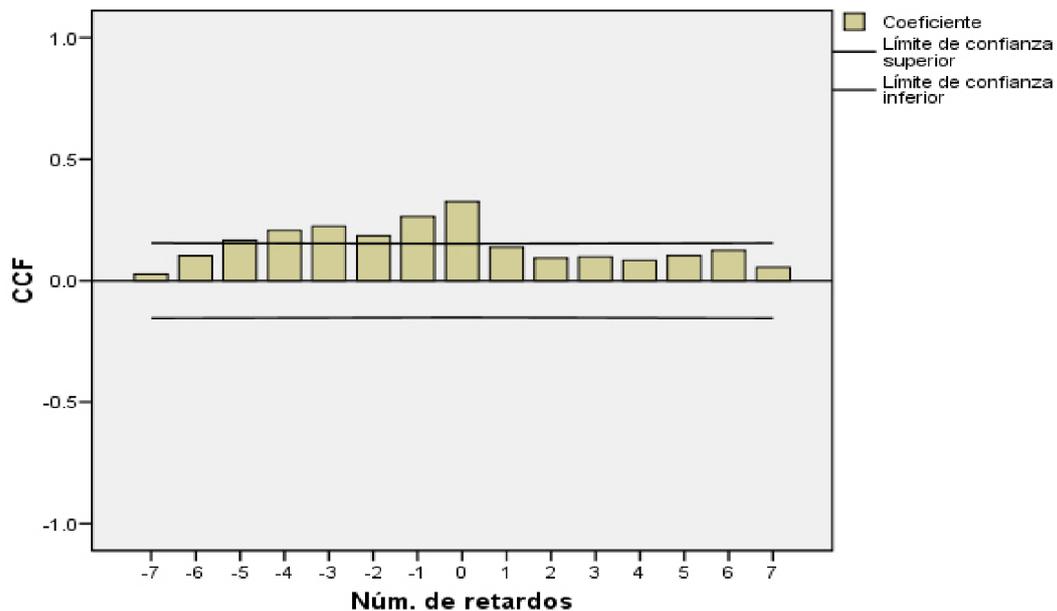
Gráfico 8. Función de Correlación Cruzada entre las series TIBP180 y TPH1.



Analizando el gráfico anterior podemos ver que las series TIBP180 y TPH1 no tienen una relación instantánea, es decir que $w1_t$ no está relacionado con $TPH1_t$, con un valor de $\rho_{TIBP180,TPH1}(h=0) = 0.058$ y es el que posee el menor valor de correlación, además vemos que a medida h aumenta o disminuye, los retardos van creciendo lentamente pero no llegan a ser significativos, ya que, ninguno sale de las bandas de confianza estimadas a un nivel del 95%, por lo que podemos decir que estas dos series no tienen relación de causalidad. Este mismo comportamiento presenta la serie TPH1 en su función de correlación cruzada con el resto de variables, por lo que descartamos que se pueda establecer un buen modelo de regresión dinámica para dicha serie. Por otro lado la serie TIBP180 presenta una relación instantánea con las series LIBOR180, PRIME y FED, es decir que $w1_t$ tiene relación con $w3_t$, $w4_t$, y $w5_t$, de igual forma con los primeros 2 retardos de la serie LIBOR180, o sea $w1_t$ tiene relación con $w3_{t-1}$ y $w3_{t-2}$; mientras que con las series PRIME y FED tiene relación con sus primeros 4 retardos; esto se puede ver más claramente, por ejemplo, con el gráfico de la función de correlación cruzada TIBP180 con PRIME:

Gráfico 9. Función de Correlación Cruzada entre las series TIBP180 y PRIME.

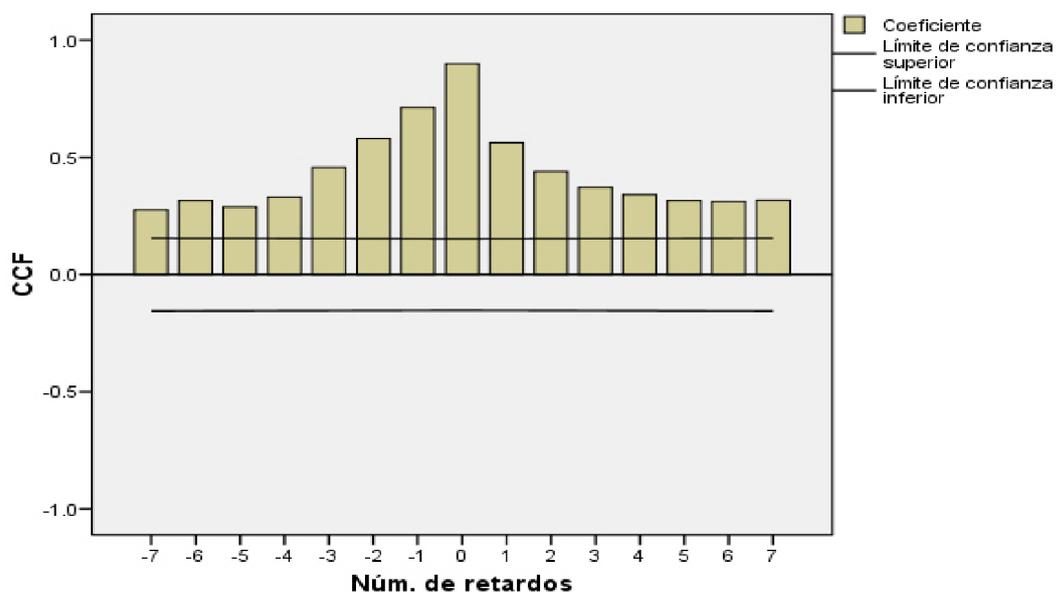
DIFF(TIBP,1) con DIFF(PRIM,1)



Por otra parte, la serie LIBOR180 posee una relación instantánea con las series PRIME y FED, además tiene una relación bidireccional con sus primeros 3 retardos, es decir que $w3_t$ tiene relación con $w4_{t-1}$, $w4_{t-2}$, $w4_{t-3}$, $w5_{t-1}$, $w5_{t-2}$, y $w5_{t-3}$; mientras que $w4_t$ está relacionado con $w3_{t-1}$, $w3_{t-2}$, $w3_{t-3}$ y $w5$ con $w3_{t-1}$, $w3_{t-2}$, $w3_{t-3}$; de igual forma las series PRIME y FED están relacionadas instantáneamente, y de forma bidireccional en sus primeros 10 retardos, esto lo podemos observar claramente en el gráfico de correlación cruzada de estas dos series:

Gráfico 10. Función de Correlación Cruzada entre las series PRIME y FED.

DIFF(PRIM,1) con DIFF(FE,1)



Para los demás pares de series no existe relación causal según la función de autocorrelación cruzada. El estudio de los modelos de regresión dinámica se hará con el fin de establecer modelos dinámicos a las dos tasas de interés bancario nacional TIBP180 y TPH1, después del análisis de las funciones de autocorrelación cruzadas, se descarto establecer un modelo a la TPH1 debido a que no tienen ninguna relación con el resto de las variables, mientras que para la serie TIBP180 vimos que tiene relaciones instantáneas y en los primeros retardos con las series LIBOR180, PRIME y FED, esto nos da la pauta para pensar que podemos establecer modelos de regresión dinámica a esta serie en función de una, dos o las 3 series anteriores. Haciendo un análisis de la función de autocorrelación cruzadas observamos que tienen fuerte colinealidad, sobre todo las series PRIME y FED, por lo tanto, solo ajustaremos modelos con una sola de las variables, los modelos propuestos de regresión dinámica a ajustarle a la serie TIBP180 son:

Tabla 8. Resumen de los posibles modelos de regresión dinámica para ser ajustados a la serie TIBP180.

	Modelo de Regresión Dinámica sin parte autorregresiva en la variable dependiente	Modelo de Regresión Dinámica con parte autorregresiva en la variable dependiente
LIBOR180	$w1_t = \delta + v_0 w3_t + n_t$ $w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$ $w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + v_2 w3_{t-2} + n_t$	$(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + n_t$ $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$ $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + v_2 w3_{t-2} + n_t$
PRIME	$w1_t = \delta + v_0 w4_t + n_t$ $w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + n_t$ $w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + v_2 w4_{t-2} + n_t$ $w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + v_2 w4_{t-2} + v_3 w4_{t-3} + n_t$ $w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + v_2 w4_{t-2} + v_3 w4_{t-3} + v_4 w4_{t-4} + n_t$	
FED	$w1_t = \delta + v_0 w5_t + n_t$ $w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + n_t$ $w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + v_2 w5_{t-2} + n_t$ $w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + v_2 w5_{t-2} + v_3 w5_{t-3} + n_t$ $w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + v_2 w5_{t-2} + v_3 w5_{t-3} + v_4 w5_{t-4} + n_t$	

3.3.2 Estimación de los parámetros de los modelos identificados, diagnosis de modelos y predicción.

Identificados los posibles modelos de regresión dinámica para la variable dependiente TIBP180, se inicia con la estimación de los parámetros, para ello, haremos uso de los modelos ARIMA univariantes ajustados a cada una de las variables independientes para definir la estructura que tendrán en su función de transferencia; además, tomaremos el modelo ARIMA de la variable dependiente como el modelo que sigue el proceso de perturbación (n_t), en caso que éste no siga un proceso de ruido blanco. Luego haremos un diagnóstico de cuál es el mejor modelo haciendo uso de los estadísticos BIC y R^2 , también utilizaremos los gráficos de la FAS y FAP residuales para saber que tan bien explica el modelo estimado a la serie en estudio, y ver si sus residuos están incorrelados, y finalmente haremos predicciones con cada modelo estimado para verificar su idoneidad.

3.3.2.1 Ajuste de un modelo de regresión dinámica a la serie TIBP180.

Estimación de los parámetros y diagnosis de los modelos identificados.

Los tres mejores modelos de regresión dinámica cuya estructura en sus funciones de transferencia para las variables independientes son: $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$, $(1 - \beta_0 B)w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + n_t$ y $(1 - \beta_0 B)w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + n_t$.

Modelo de regresión dinámica ajustado a la variable dependiente TIBP180 por LIBOR180.

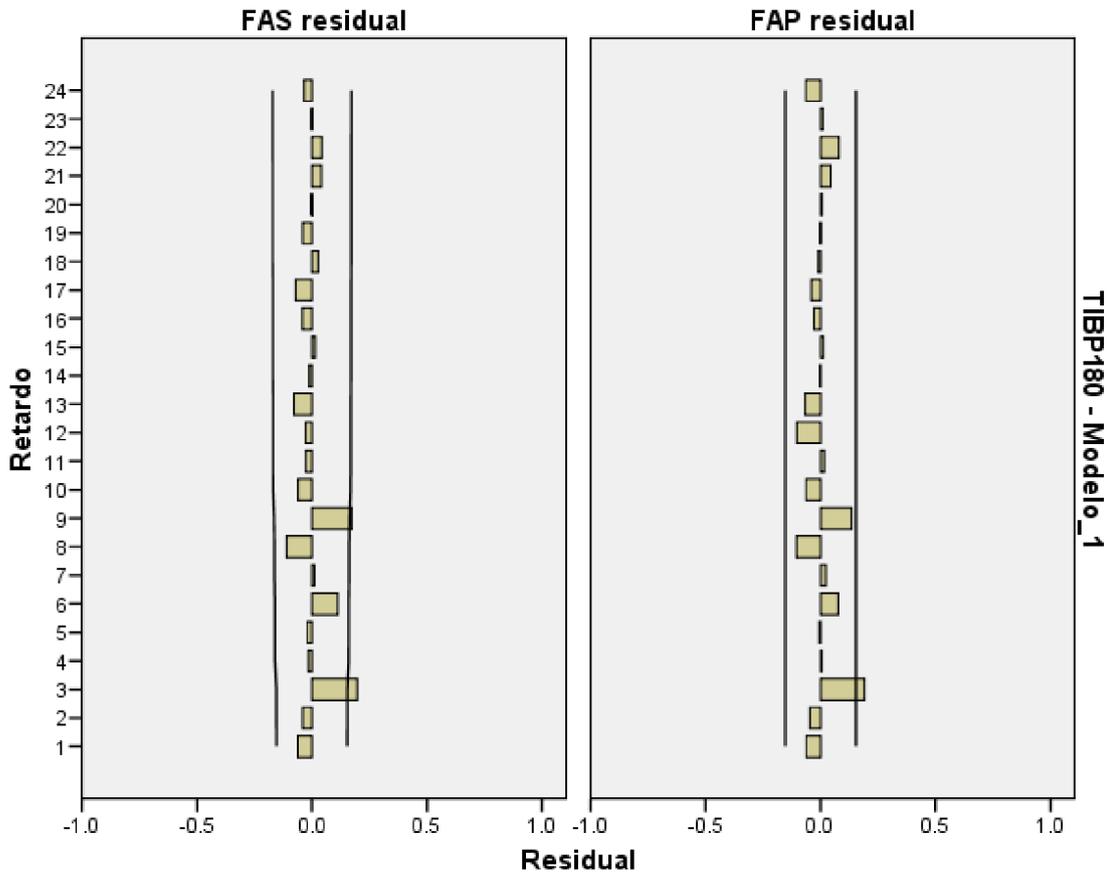
Tabla 9. Parámetros del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$ ajustado a la serie TIBP180.

				Estimación	ET	t	Sig.
TIBP180-Modelo_1	TIBP180	Sin transformación	Constante	-.005	.017	-286	.775
			AR				
			Retardo 1	.312	.077	4.071	.000
			Diferencia				
			1				
	LIBOR180	Sin transformación	Numerador				
			Retardo 0	.082	.072	1.138	.257
			Retardo 1	-.228	.072	-3.174	.002
			Diferencia				
			1				

Tabla 10. Estadísticos del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$ ajustado a la serie TIBP180.

Modelo	Número de predictores	Estadísticos de ajuste del modelo			Ljung-Box Q(18)			Número de valores atípicos
		R-cuadrado estacionaria	R-cuadrado	BIC normalizado	Estadísticos	GL	Sig.	
TIBP180-Modelo_1	1	.169	.993	-3.670	20.467	17	.251	0

Gráfico 11. FAS y FAP residual del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w3_t + v_1 w3_{t-1} + n_t$ ajustado a la serie TIBP180.



La tabla 9 muestra que la constante y el coeficiente del retardo 0 de la variable LIBOR180, no son significativos; mientras que los coeficientes de los primeros retardos de las dos variables son significativos, por otra parte, este modelo explica el 99.3% de la serie TIBP180, y tiene un estadístico BIC = -3.67 que es un valor pequeño y negativo, además su FAS y FAP residual no presenta algún comportamiento determinado, con lo cual los residuos están incorrelados y se aproximan a un ruido blanco, por tanto, podemos pensar que este modelo ajusta bien a la serie dependiente, y la expresión del modelo de regresión dinámica es:

$$(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t$$

Modelo de regresión dinámica ajustado a la variable dependiente TIBP180 por PRIME.

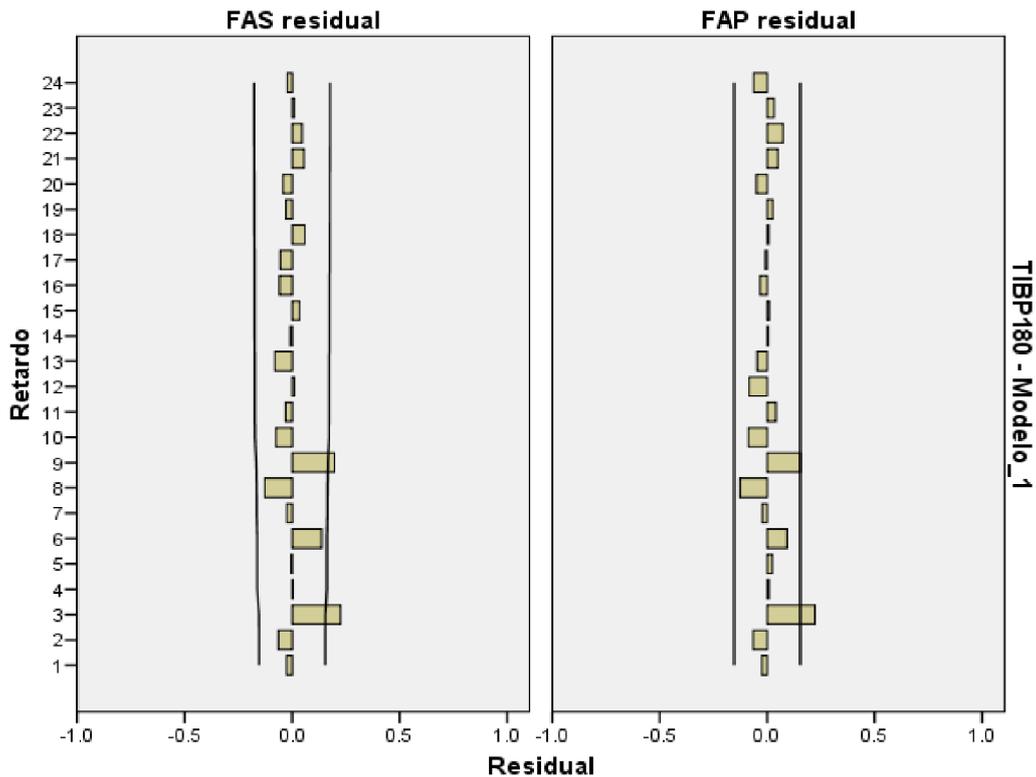
Tabla 11. Parámetros del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + n_t$ ajustado a la serie TIBP180.

				Estimación	ET	t	Sig.
TIBP180-Modelo_1	TIBP180	Sin transformación	Constante	-0.004	.015	-2.48	.804
			AR Retardo 1	.240	.078	3.065	.003
		Diferencia	1				
	PRIME	Sin transformación	Numerador Retardo 0	.211	.076	2.796	.006
			Numerador Retardo 1	-.188	.074	-2.559	.011
		Diferencia	1				

Tabla 12. Estadísticos del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + n_t$ ajustado a la serie TIBP180.

Modelo	Número de predictores	Estadísticos de ajuste del modelo			Ljung-Box Q(18)			Número de valores atípicos
		R-cuadrado estacionaria	R-cuadrado	BIC normalizado	Estadísticos	GL	Sig.	
TIBP180-Modelo_1	1	.204	.994	-3.714	26.590	17	.064	0

Gráfico 12. FAS y FAP residual del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w4_t + v_1 w4_{t-1} + n_t$ ajustado a la serie TIBP180.



La tabla 11 muestra que la constante no es significativa; mientras que los tres coeficientes estimados de los retardos de las dos variables son significativos, por otra parte, este modelo explica el 99.4% de la serie TIBP180, y tiene un estadístico BIC = -3.714 que es un valor pequeño y negativo, además su FAS y FAP residual no presenta algún comportamiento determinado, con lo cual los residuos están incorrelados y se aproximan a un ruido blanco, por tanto, podemos pensar que este modelo ajusta bien a la serie dependiente, y la expresión del modelo de regresión dinámica es:

$$(1 - 0.240B)(1 - B)TIBP180_t = (0.211 - 0.188B)(1 - B)PRIME_t + n_t$$

Modelo de regresión dinámica ajustado a la variable dependiente TIBP180 por FED.

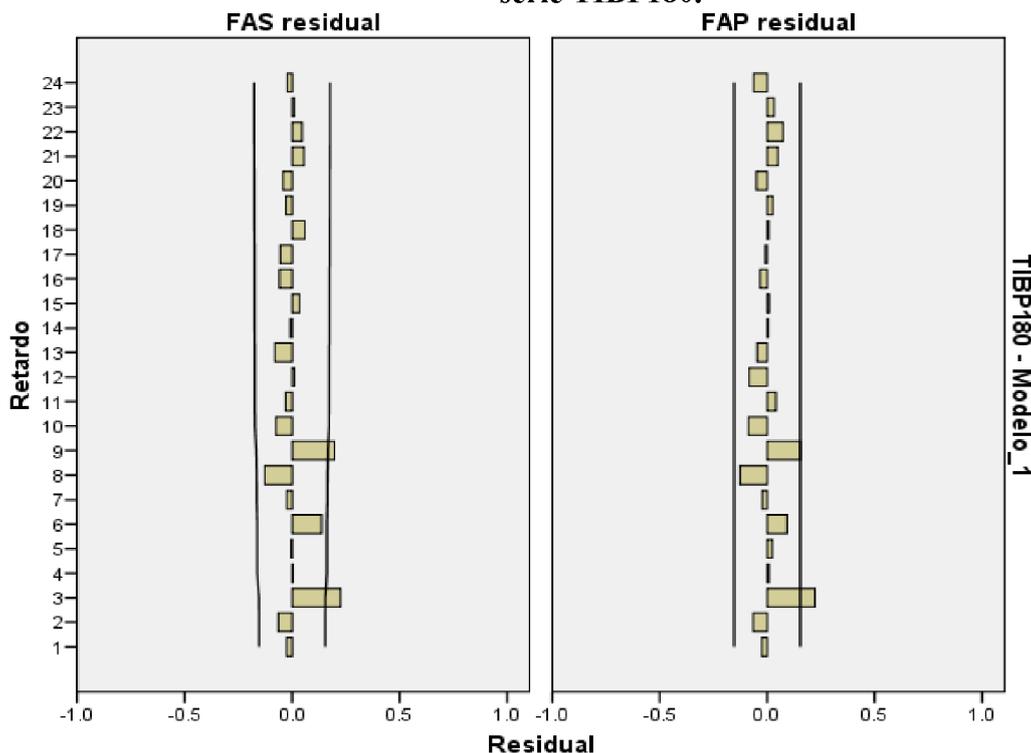
Tabla 13. Parámetros del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + n_t$ ajustado a la serie TIBP180.

					Estimación	ET	t	Sig.
TIBP180-Modelo_1	TIBP180	Sin transformación	Constante		-.003	.015	-2.08	.835
			AR	Retardo 1	.255	.078	3.255	.001
			Diferencia	1				
FE	Sin transformación	Numerador	Retardo 0	.122	.076	1.592	.113	
			Retardo 1	-.303	.075	-4.041	.000	
		Diferencia	1					

Tabla 14. Estadísticos del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + n_t$ ajustado a la serie TIBP180.

Modelo	Número de predictores	Estadísticos de ajuste del modelo			Ljung-Box Q(18)			Número de valores atípicos
		R-cuadrado estacionaria	R-cuadrado	BIC normalizado	Estadísticos	GL	Sig.	
TIBP180-Modelo_1	1	.227	.994	-3.742	30.403	17	.024	0

Gráfico 12. FAS y FAP residual del modelo $(1 - \beta_0 B)w1_t = \delta + v_0 w5_t + v_1 w5_{t-1} + n_t$ ajustado a la serie TIBP180.



La tabla 13 muestra que la constante no es significativa; mientras que los dos coeficientes estimados de los primeros retardos de las dos variables son significativos, por otra parte, este modelo explica el 99.4% de la serie TIBP180, y tiene un estadístico BIC = -3.742 que es un valor pequeño y negativo, además su FAS y FAP residual no presentan algún comportamiento determinado, con lo cual los residuos están incorrelados y se aproximan a un ruido blanco, por tanto, podemos pensar que este

modelo ajusta bien a la serie dependiente, y la expresión del modelo de regresión dinámica es:

$$(1 - 0.255B)(1 - B)TIBP180_t = -0.303(1 - B)FED_{t-1} + n_t$$

Predicción de los modelos.

En la etapa de diagnóstico no pudimos establecer cuál era el mejor modelo de los 3 ajustados, por tanto, se hicieron predicciones con cada uno, a continuación se presenta el gráfico y tabla de las mejores predicciones:

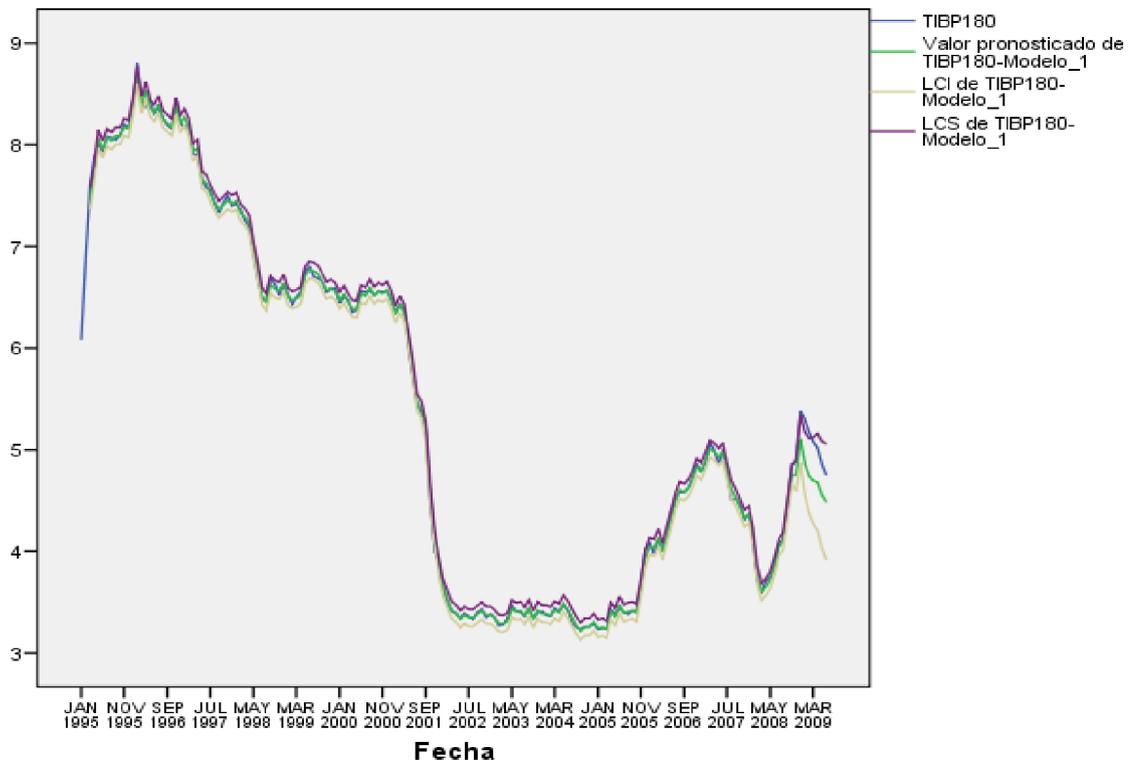
Tabla 15. Predicciones del modelo $(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t$ ajustado a la serie TIBP180.

		Previsión									
Modelo		Oct 2008	Nov 2008	Dic 2008	Ene 2009	Feb 2009	Mar 2009	Abr 2009	May 2009	Jun 2009	
	Valor Observado	4.86	4.88	5.38	5.3	5.18	5.08	5.02	4.86	4.75	
TIBP180-Modelo_1	Previsión	4.73	4.76	5.13	4.95	4.89	4.90	4.95	4.88	4.87	
	LCS	4.89	4.99	5.49	5.43	5.50	5.64	5.80	5.85	5.96	
	LCI	4.62	4.53	4.78	4.46	4.28	4.17	4.09	3.91	3.79	

Para cada modelo, las predicciones comienzan después del último valor no perdido del rango del período de estimación solicitado y finalizan en el último período para el que hay disponibles

Gráfico 13. Predicciones del modelo

$(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t$ ajustado a la serie TIBP180.



Con los resultados obtenidos anteriormente podemos concluir que el mejor modelo ajustado a la serie TIBP180 es: $(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t$, donde n_t se comporta aproximadamente como un ruido blanco.

3.4. Análisis de Modelos Multivariantes de Series Temporales.

El análisis de los modelos multivariantes se desarrollará con el fin de establecer un modelo que describa de mejor manera la tasa de interés bancaria nacional TIBP180 con la tasa de interés bancaria internacional LIBOR180 explicadas entre ellas. Las variables se han tomado de esa forma ya que el mejor modelo de regresión dinámica que explicaba las tasas de interés bancario nacional TIBP180 es la variable LIBOR180. El estudio de las series conjuntas de series temporales establece inicialmente la estacionariedad de las mismas así como también contempla las tres fases que cualquier modelo estimado establece las cuáles son la **identificación, estimación y diagnosis**, las cuáles se describen a continuación.

3.4.1 Identificación del modelo para las series temporales TIBP180 y LIBOR 180.

Para empezar a trabajar se debe tener principalmente la certeza que estamos con series independientes, es decir que ninguna de ellas puede escribirse como combinación lineal de otra o de otras, si ese fuese el caso nos quedaríamos con una serie menos.

Independencia de variables

Con la prueba de Granger determinaremos el grado de exogeneidad o de independencia de las variables. El grado de independencia es importante ya que de esa forma se introducirán los datos en el programa EVIEWS6.

Tabla 16. Prueba de Exogeneidad de Granger.

Pairwise Granger Causality Tests

Date: 11/23/09 Time: 20:04

Sample: 1 174

Lags: 2

Null Hypothesis:	Obs	F-Statistic	Prob.
DIFLIBOR180 does not Granger Cause DIFTIBP180	171	4.17117	0.0171
DIFTIBP180 does not Granger Cause DIFLIBOR180		1.73246	0.1800

En la tabla 16 se presenta la prueba de independencia Granger; en esta prueba podemos observar que la variable TIBP180 tiene relación de causalidad con la LIBOR180 lo cuál es ilógico pensarlo ya que en la realidad no tienen nada que ver las tasas de interés bancario nacional sobre las tasas de interés bancario internacional. Esta relación se debe a situaciones del azar, por tanto, no debe considerarse en el análisis.

En el análisis univariante se estableció que las series LIBOR180 y TIBP180 no son estacionarias en media y en consecuencia era necesario hacer una diferencia. Por lo tanto en el análisis bivariado se verificará si existen vectores de cointegración, para ello, se aplica el test de los Rangos:

Tabla 17. Prueba de Existencia de Vectores de Cointegración.

Date: 11/23/09 Time: 20:13
 Sample (adjusted): 4 174
 Included observations: 171 after adjustments
 Trend assumption: Linear deterministic trend
 Series: DIFTIBP180 DIFLIBOR180
 Lags interval (in first differences): 1 to 1

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.327018	102.3454	15.49471	0.0001
At most 1 *	0.183293	34.62321	3.841466	0.0000

Trace test indicates 2 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.327018	67.72218	14.26460	0.0000
At most 1 *	0.183293	34.62321	3.841466	0.0000

Max-eigenvalue test indicates 2 cointegrating eqn(s) at the 0.05 level

* denotes rejection of the hypothesis at the 0.05 level

**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegrating Coefficients (normalized by b'S11*b=l):

DIFTIBP180	DIFLIBOR180
-6.686010	2.799977
-2.473488	-4.362052

Unrestricted Adjustment Coefficients (alpha):

D(DIFTIBP180)	0.095228	0.022783
D(DIFLIBOR180)	-0.021719	0.096929

1 Cointegrating Equation(s): Log likelihood 100.2017

Normalized cointegrating coefficients (standard error in parentheses)

DIFTIBP180	DIFLIBOR180
1.000000	-0.418781 (0.08405)

Adjustment coefficients (standard error in parentheses)

D(DIFTIBP180)	-0.636693 (0.07585)
D(DIFLIBOR180)	0.145215 (0.11824)

La prueba indica que hay 2 vectores de cointegración con un nivel de 0.05. La hipótesis nula es no existen vectores de cointegración y la alterna es existen vectores de cointegración. En este caso se observa que existen dos vectores de cointegración según los dos primeros test que son: el de máximo eigenvalor (Maximum eigenvalue) y traza (trace), con un p-valor menor que 0.05, por lo tanto, rechazamos la hipótesis nula, en consecuencia, hay dos vectores de cointegración.

Después de haber verificado la independencia, estacionariedad y la existencia de cointegración procedemos a la identificación del modelo multivariante.

En la parte univariante se verificó que la variable TIBP180 la predice bastante bien un ARI(1,1), así como también la variable LIBOR180. Siguiendo la misma lógica es posible que un modelo VAR prediga de buena manera una serie temporal conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$. La identificación del modelo puede obtenerse a partir de los modelos univariantes elaborados a partir de y_t , ya que cada variable se explicaba así misma de buena manera. Así como también nos sirven como herramienta los criterios de información AIC y BIC para determinar el orden del modelo entre varios modelos estimados con órdenes distintos y con la misma muestra efectiva, tomando como criterio el menor valor de AIC y BIC de entre los modelos estimados con diferente orden.

Tabla 18. Orden del criterio de selección de los retardos.

VAR Lag Order Selection Criteria
 Endogenous variables: DIFTIBP180 DIFLIBOR180
 Exogenous variables: C
 Date: 11/23/09 Time: 20:17
 Sample: 1 174
 Included observations: 165

Lag	LogL	LR	FPE	AIC	SC	HQ
0	79.38027	NA	0.001342	-0.937943	-0.900295	-0.922660
1	107.3414	54.90540	0.001004	-1.228380	-1.115437*	-1.182532*
2	110.0534	5.259751	0.001019	-1.212769	-1.024530	-1.136356
3	118.8072	16.76480*	0.000962*	-1.270390*	-1.006855	-1.163412
4	122.5934	7.159329	0.000965	-1.267798	-0.928968	-1.130255
5	123.2634	1.250671	0.001005	-1.227435	-0.813309	-1.059327
6	123.9303	1.228838	0.001047	-1.187034	-0.697613	-0.988361
7	126.1044	3.952751	0.001070	-1.164901	-0.600184	-0.935663
8	128.4518	4.211171	0.001093	-1.144870	-0.504857	-0.885067

* indicates lag order selected by the criterion
 LR: sequential modified LR test statistic (each test at 5% level)
 FPE: Final prediction error
 AIC: Akaike information criterion
 SC: Schwarz information criterion
 HQ: Hannan-Quinn information criterion

En la tabla18 se presentan los valores de los estadísticos de criterio de información que tienen como finalidad seleccionar el mejor modelo tomando el menor valor que presenten los mismos. Se puede observar que para el retardo 1 los valores de los estadísticos del BIC= -1.115437, HQ= -1.182532 son los más pequeños aunque también se puede optar por pronosticar un modelo con retardo 3 ya que tiene los estadísticos de comparación entre los modelos más pequeños, sin embargo, en el análisis de regresión dinámica se determino que existe relación entre las variables hasta retardos de orden 1, por lo cual trabajaremos con el primer modelo.

3.4.2 Estimación y diagnosis de un modelo VAR para TIBP180 y LIBOR 180.

Después de haber verificado los errores para un VAR(1) y revisado algunos estadísticos para determinar el mejor modelo que se utilizará para la estimación de los datos que corresponden a la serie conjunta Y_t , determinamos que el mejor modelo es un VAR(1).

A continuación se presenta la estimación del modelo VAR(1).

Tabla 19. Modelo VAR(1).

Vector Autoregression Estimates		
Date: 11/23/09 Time: 19:52		
Sample (adjusted): 3 174		
Included observations: 172 after adjustments		
Standard errors in () & t-statistics in []		
	DIFTIBP1	DIFLIBOR
DIFTIBP1(-1)	0.355552 (0.06665) [5.33473]	-0.118030 (0.09157) [-1.28900]
DIFLIBOR(-1)	0.095111 (0.05056) [1.88118]	0.464560 (0.06946) [6.68797]
C	-0.006790 (0.01175) [-0.57792]	-0.017338 (0.01614) [-1.07411]
R-squared	0.179212	0.209296
Adj. R-squared	0.169498	0.199939
Sum sq. resids	3.939513	7.435973
S.E. equation	0.152679	0.209761
F-statistic	18.44979	22.36679
Log likelihood	80.71618	26.08276
Akaike AIC	-0.903676	-0.268404
Schwarz SC	-0.848778	-0.213506
Mean dependent	-0.012326	-0.031221
S.D. dependent	0.167536	0.234511
Determinant resid covariance (dof adj.)		0.000982
Determinant resid covariance		0.000948
Log likelihood		110.5312
Akaike information criterion		-1.215479
Schwarz criterion		-1.105683

En este despliegue se incluyen las ecuaciones que integran el modelo VAR(1), apareciendo en primer lugar los resultados individuales para cada uno de los coeficientes estimados (primera línea), acompañados de su desviación típica (segunda línea) y el estadístico t asociado (tercera línea), ambos incluidos entre paréntesis.

En la parte inferior se recogen los estadísticos conjuntos de cada una de las ecuaciones, para finalizar con cuatro estadísticos globales del modelo, incluyendo el determinante de la matriz de covarianzas de los residuos el cual nos muestra un valor muy cercano a cero, por lo que podemos decir que el modelo es una buena estimación del vector conjunto Y_t , el logaritmo de verosimilitud conjunto y los criterios informativos de Akaike y Schwarz son los más pequeños de todos los modelos estimados.

En nuestro caso hemos seleccionado un VAR(1), ya que el segundo valor mínimo de los criterios informativos (que señala el mejor modelo) se alcanza, precisamente con 1 retardos, al menos en el criterio de Schwartz, por lo que nos quedaremos con dicha especificación. La evaluación de los modelos VAR es bastante compleja, ya que la significatividad individual de los parámetros está muy condicionada por la necesaria correlación entre los regresores, precisamente por ese motivo el análisis individual de los coeficientes no es demasiado útil en proceso de selección de tamaño del modelo, y las posibilidades de validación a priori se limitan, casi exclusivamente, a los coeficientes de determinación, y el análisis de los residuos del modelo.

Tabla 20. Test de autocorrelaciones de Portmanteau.

Este test nos sirve para determinar la autocorrelación residual entre los retardos posteriores al retardo del modelo estimado.

VAR Residual Portmanteau Tests for Autocorrelations
 Null Hypothesis: no residual autocorrelations up to lag h
 Date: 11/23/09 Time: 20:21
 Sample: 1 174
 Included observations: 172

Lags	Q-Stat	Prob.	Adj Q-Stat	Prob.	df
1	3.513901	NA*	3.534451	NA*	NA*
2	6.901664	0.1412	6.962069	0.1379	4
3	19.47349	0.0125	19.75706	0.0113	8
4	27.10207	0.0075	27.56728	0.0064	12
5	28.57529	0.0270	29.08460	0.0234	16
6	34.68715	0.0218	35.41738	0.0180	20
7	36.43871	0.0497	37.24325	0.0414	24
8	40.31389	0.0620	41.30745	0.0503	28
9	47.15538	0.0411	48.52670	0.0307	32
10	59.35015	0.0085	61.47423	0.0051	36
11	65.85711	0.0062	68.42577	0.0034	40
12	67.66032	0.0125	70.36421	0.0070	44

*The test is valid only for lags larger than the VAR lag order.
 df is degrees of freedom for (approximate) chi-square distribution

La prueba de Portmanteau nos indica que entre los residuos existe autocorrelación en los distintos periodos de tiempo a partir del retardo 2 ya que todos los valores posteriores a ese retardo son significativos, lo que nos indica que los residuos no son buenos en la estimación del modelo.

Tabla 21. Prueba de normalidad de los residuos.

Con esta prueba determinamos la existencia de normalidad entre los residuales de cada serie, así como los residuales de las series de forma conjunta.

VAR Residual Normality Tests
 Orthogonalization: Cholesky (Lutkepohl)
 Null Hypothesis: residuals are multivariate normal
 Date: 11/23/09 Time: 20:24
 Sample: 1 174
 Included observations: 172

Component	Skewness	Chi-sq	df	Prob.
1	0.012136	0.004222	1	0.9482
2	-2.512525	180.9664	1	0.0000
Joint		180.9706	2	0.0000

Component	Kurtosis	Chi-sq	df	Prob.
1	6.089521	68.40683	1	0.0000
2	18.65924	1757.351	1	0.0000
Joint		1825.758	2	0.0000

Component	Jarque-Bera	df	Prob.
1	68.41105	2	0.0000
2	1938.318	2	0.0000
Joint	2006.729	4	0.0000

Para determinar la normalidad de los residuos utilizamos el estadístico de Jarque-Bera; se observa que los valores del p-valor son menores a 0.05 los cuáles se asume que no hay normalidad de los residuos de forma conjunta, así como también de forma individual para las series TIBP180 y LIBOR180.

Tabla 22. Funciones de respuesta a impulsos.

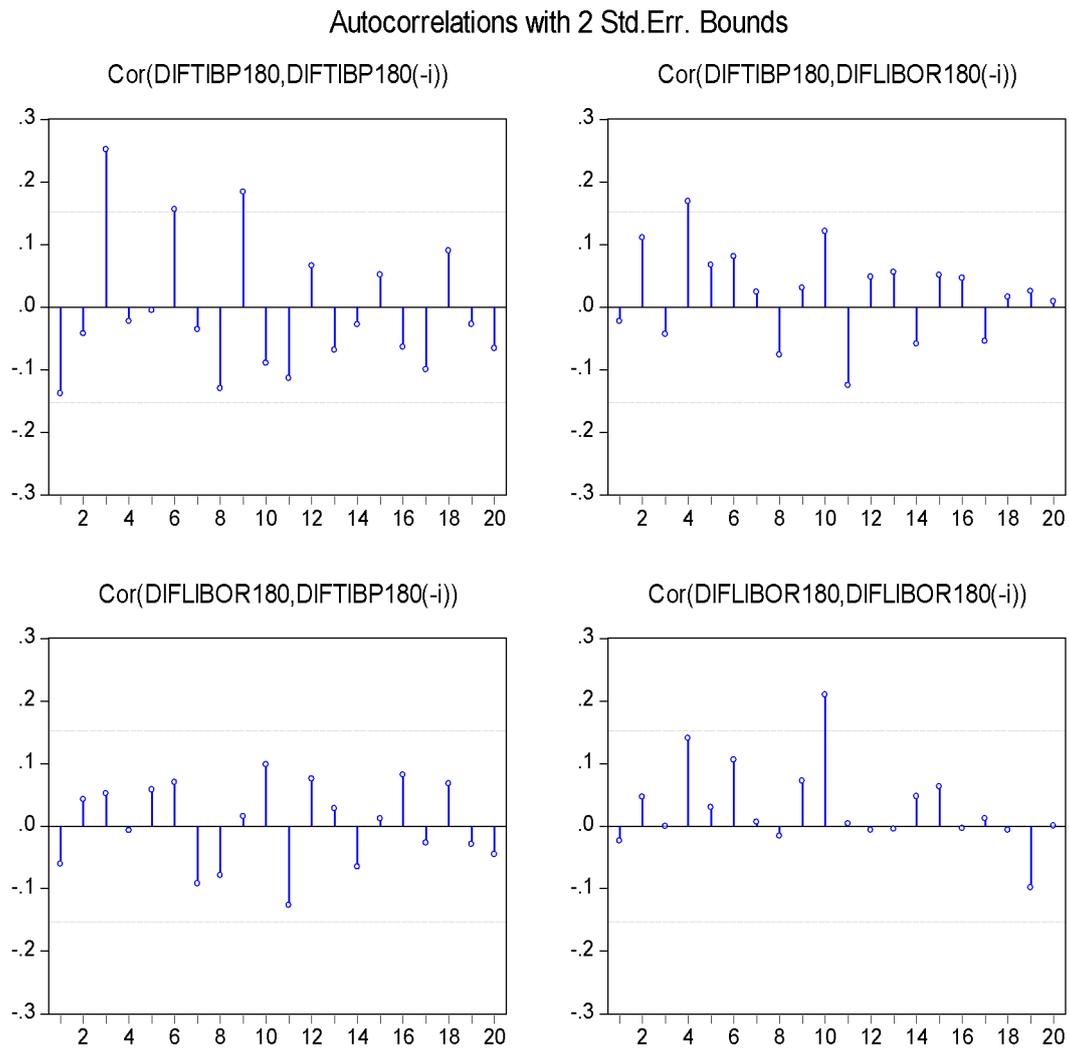
Response of DIFTIBP180:		
Period	DIFTIBP180	DIFLIBOR180
1	0.152679	0.000000
2	0.058397	0.019522
3	0.020959	0.016010
4	0.006888	0.009687
5	0.001951	0.005120
6	0.000385	0.002490
7	-2.82E-05	0.001139
8	-9.11E-05	0.000495
9	-6.97E-05	0.000205
10	-4.11E-05	8.08E-05

Response of DIFLIBOR180:		
Period	DIFTIBP180	DIFLIBOR180
1	0.043228	0.205259
2	0.002062	0.095355
3	-0.005935	0.041994
4	-0.005231	0.017619
5	-0.003243	0.007042
6	-0.001737	0.002667
7	-0.000852	0.000945
8	-0.000393	0.000305
9	-0.000172	8.31E-05
10	-7.15E-05	1.44E-05

Cholesky Ordering: DIFTIBP180 DIFLIBOR180

En la función respuesta a impulsos se observa un comportamiento de efectos positivos y negativos en las dos variables; es decir que un incremento en cualquiera de las variables puede causar un efecto negativo o positivo según sea el caso, aunque con crecimientos y decrecimientos pequeños en diferentes períodos. Esto se debe a que las tasas se mantienen oscilando entre ciertos valores los cuales a medida que pasa el tiempo surgen cambios pero dichos cambios son mínimos ya que no pueden ser susceptibles; si así fuese hubiesen comportamientos explosivos en los períodos, los cuales no son notables en nuestro caso.

Gráfico 14. Errores de Autocorrelación.



Observamos que la mayoría de los errores se encuentran correctamente dentro de la banda de confianza, sólo hay un error significativo en las correlaciones, por lo que esperaríamos que el modelo haga buenas predicciones para los valores de la serie conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$.

3.4.3 Predicciones para el modelo VAR de la serie TIBP180 y LIBOR 180.

En el siguiente gráfico se presenta la predicción del modelo VAR(1) que se ha establecido para la serie conjunta Y_t .

Gráfico 15. Predicciones del modelo.

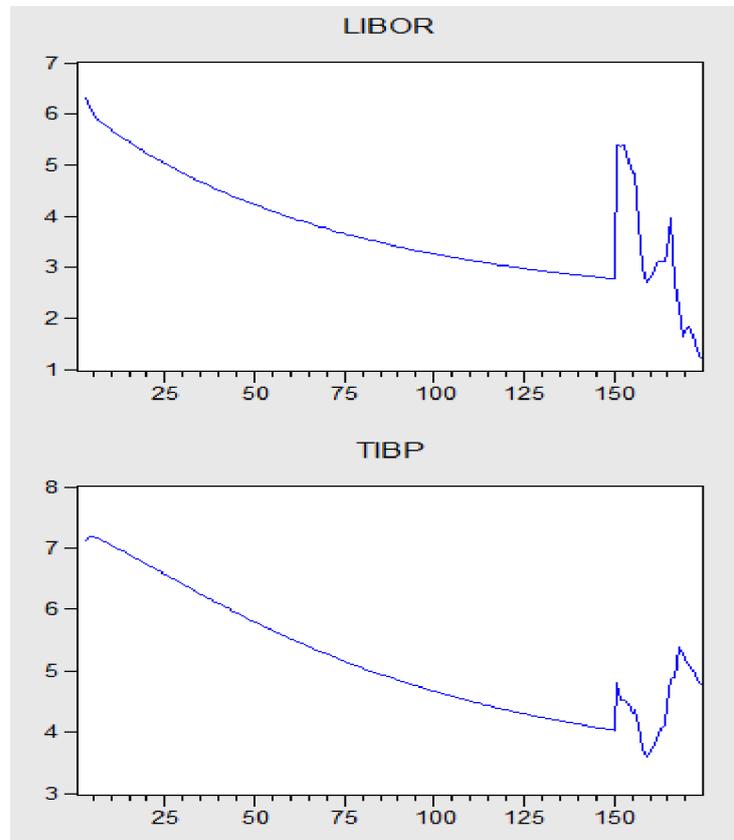


Tabla 23. Predicciones del modelo VAR(1).

Valores Originales			Valores Predichos		
obs	TIBP	LIBOR	obs	LIBOR_0	TIBP_0
4	7.88	6.40	4	6.141256	7.195129
5	8.00	6.18	5	6.019157	7.193792
6	7.93	5.91	6	5.930334	7.170799
7	8.08	5.86	7	5.860090	7.141568
8	8.07	5.93	8	5.799754	7.110869
9	8.05	5.88	9	5.744681	7.079943
10	8.10	5.89	10	5.692532	7.048985
11	8.20	5.74	11	5.642169	7.017944
12	8.17	5.62	12	5.593053	6.986756
13	8.38	5.43	13	5.544932	6.955383
14	8.80	5.17	14	5.497684	6.923815
15	8.38	5.39	15	5.451246	6.892061
16	8.37	5.54	16	5.405584	6.860136
17	8.43	5.60	17	5.360673	6.828061
18	8.31	5.75	18	5.316495	6.795860
19	8.37	5.84	19	5.273034	6.763552
20	8.26	5.70	20	5.230273	6.731161
21	8.19	5.82	21	5.188199	6.698706

En la tabla anterior observamos que los valores predichos están cerca de los valores reales, aunque deben de tomarse con cautela, pues los residuos no son exactamente un ruido blanco y en el modelo para la serie conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$, aparecen dos tipos de errores que podrían estar correlacionados.

El modelo vectorial que describe la serie conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$ es un VAR(1), el cual viene dado de la siguiente manera:

$$\begin{bmatrix} 1-B & 0 \\ 0 & 1-B \end{bmatrix} \begin{bmatrix} TIBP180_t \\ LIBOR180_t \end{bmatrix} = \begin{bmatrix} -0.0068 \\ -0.0173 \end{bmatrix} + \begin{bmatrix} 0.3556 & 0.0951 \\ -0.1180 & 0.4645 \end{bmatrix} \begin{bmatrix} 1-B & 0 \\ 0 & 1-B \end{bmatrix} \begin{bmatrix} TIBP180_{t-1} \\ LIBOR180_{t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

En los resultados anteriores se esperaba que el modelo VAR(1) que describe la serie conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$ tuviese relación con el modelo de regresión dinámica ajustado por las series TIBP180 y LIBOR180 lo cuál no siempre se cumple, ya que, si por ejemplo, en el modelo vectorial anterior tomamos $\phi_{21} = 0$ y despreciando las constantes, tenemos:

$$(1-0.3556B)(1-B)TIBP180_t = 0.0951(1-B)LIBOR180_{t-1} + a_{1t} \quad (3.1)$$

$$(1-0.4645B)(1-B)LIBOR180_t = a_{2t} \quad (3.2)$$

Las ecuaciones (3.1) y (3.2) nos muestran que la variable $LIBOR180_t$ no depende del pasado de la serie $TIBP180_t$, en consecuencia tenemos una relación causal unidireccional de LIBOR180 hacia TIBP180, que es la relación establecida en los modelos de regresión dinámica, pero, podemos observar que los coeficientes estimados de la ecuación (3.1) difieren con los obtenidos en el modelo de la sección anterior que es: $(1-0.312B)(1-B)TIBP180_t = -0.228(1-B)LIBOR180_{t-1} + a_{1t}$ (3.3), esto se debe a que, en general las ecuaciones (3.1) y (3.2) no representan modelos de función de transferencia ya que sus residuos a_{1t} y a_{2t} podrían estar correlados, como vimos en los resultados obtenidos en el ajuste del modelo VAR(1). Para obtener una aproximación a la función de transferencia del modelo (3.3) se necesita ortogonalizar los residuos del modelo VAR(1) para que estos sean incorrelados.

4. Conclusiones.

- En el estudio de los modelos univariantes de series temporales, se establece que el modelo ARIMA(1,1,0) es el que mejor ajusta a las series: TIBP180, LIBOR180, PRIME, FED, e IVAE, mientras que el ARIMA(0,1,0) y ARIMA(1,0,0) son los que mejor ajustan a las series IGInflacion y THP1 respectivamente.
- Al analizar los modelos de regresión dinámica, mediante la función de correlación cruzada, obtuvimos que la tasa de interés bancarias nacional TPH1 así como los dos índices IGInflacion e IVAE no poseen ninguna relación causal con el resto de series en estudio, por lo tanto, se descartó ajustar un modelo de regresión dinámica o incluir en alguno de los estimados, mientras que la serie TIBP180 posee relación instantánea con las 3 tasas internacionales en estudio (LIBOR180, PRIME y FED). La serie que mejor explica a la tasa de interés que los bancos nacionales ofrecen a sus depositantes (TIBP180) es la tasa promedio de interés que impone el banco interbancario de Londres día a día (LIBOR180), con lo que confirmamos que las tasas que se imponen a nivel nacional tienen relación con las internacionales, y la expresión de este modelo está dada por:

$$(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t.$$
- El análisis multivariante de series temporales para la variable conjunta $Y_t = [TIBP180_t, LIBOR180_t]'$ establece que puede ser modelado con un VAR(1). Cabe mencionar que el mejor modelo establecido para Y_t no cumple con las condiciones mínimas para considerarlo como el idóneo para predecir la serie conjunta Y_t , ya que la LIBOR180 de ninguna forma puede ser explicada por la TIBP180, únicamente se cumple en el caso inverso. El modelo VAR(1) viene dado por:

$$\begin{bmatrix} 1-B & 0 \\ 0 & 1-B \end{bmatrix} \begin{bmatrix} TIBP180_t \\ LIBOR180_t \end{bmatrix} = \begin{bmatrix} -0.0068 \\ -0.0173 \end{bmatrix} + \begin{bmatrix} 0.3556 & 0.0951 \\ -0.1180 & 0.4645 \end{bmatrix} \begin{bmatrix} 1-B & 0 \\ 0 & 1-B \end{bmatrix} \begin{bmatrix} TIBP180_{t-1} \\ LIBOR180_{t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

- Para efectos de hacer predicciones de la serie TIBP180, se recomienda el modelo de regresión dinámica dado por:

$$(1 - 0.312B)(1 - B)TIBP180_t = -0.228(1 - B)LIBOR180_{t-1} + n_t$$

5. Recomendaciones.

- Se recomienda el estudio de los modelos de regresión dinámica y modelos multivariantes de series temporales, cuando se desee estudiar las relaciones que podrían existir entre dos o más series de tiempo, ya que permiten establecer modelos bastante fiables y determinan el nivel de relación existente entre las series.
- En futuros trabajos de investigación sobre modelos de regresión dinámica y modelos multivariantes de series temporales, se sugiere que no se tomen variables estrictamente económicas o que posean mucha colinealidad, ya que ésta impide que se estimen modelos no tan simples y se haga un mejor estudio entre variables.

6. Bibliografía.

Referencias.

- Aznar, Antonio y Trivez, Francisco Javier. “Métodos de predicción en economía análisis de series temporales”, Tomo II, 1993, Editorial Ariel.
- Correa, Juan Byron. “Análisis comparativo en la estimación de modelos de regresión”, 2002, Universidad del Valle.
- Escudero García, José Ramón y Vallejo Seco, Guillermo. “Aplicación del diseño de series temporales múltiples a un caso de intervención en dos clases de Enseñanza General Básica”, 2000, Universidad de Oviedo.
- Galeano San Miguel, Pedro. Tesis Doctoral “Atípicos, cambios estructurales y discriminación en Series Temporales Multivariantes”, 2004, Universidad Carlos III de Madrid.
- Jaimes Berrios, Alma Azucena y Quintanilla Aparicio, María Isabel. Tesis de Licenciatura “Análisis de Series Temporales con Outliers e Intervenciones y sus Aplicaciones”, 2008, Universidad de El Salvador.
- Mauricio, José Alberto. “Introducción al Análisis de Series Temporales”, 2007, Universidad Complutense Madrid.
- Pankratz, Alan. “Forecasting whit Dynamic Regression Models”, 1991, Wiley-Interscience Publication.
- Peña, Daniel. “Análisis de Series temporales”, 2005, Alianza Editorial.

- Ramos Alberto, Pedro Armando. Tesis de Maestría “Ajuste de un modelo ARIMA a la emanación de gas CO_2 del volcán de San Vicente”, 2007, Universidad de El Salvador.
- Renderos, Deysi Maribel y Molina Masferrer, Mario Giovanni. Tesis de Maestría “Ajuste de un modelo ARIMA para la precipitación diaria en la zona oriental de El Salvador, 2008, Universidad de El Salvador.
- Ribeiro, Marcio Bruno y Teixeira, Joaquinillo Rodolpho. “Análisis econométrico de la inversión privada en Brasil”, 2001, Universidad de Brasilia.
- S. Wei, William. “Times Series Analysis Univariate and Multivariate Methods”, 1994, Addison-Wesley Publishing Company, Inc.
- Trigueros Arguello, Álvaro. “Impacto del incremento del desempleo en Estados Unidos sobre las remesas familiares en El Salvador”, 2009, FUSADES.
- Vicente Martínez, Eva María. Tesis Doctoral “Análisis de Series Temporales Macroeconómicas de Estados Unidos relacionadas con la Inflación”, 2004, Universidad Complutense de Madrid.