

Notas de Clase: Series de Tiempo

Benjamín Oliva, Omar Alfaro Rivera & Emiliano Pérez Caullieres

2025-08-21

Contents

Análisis de Series de Tiempo	5
Contexto del documento	5
1 Introducción	7
2 Introducción al Análisis de Series de Tiempo y Elementos de Ecuaciones en Diferencia	9
2.1 La naturaleza de los datos de Series de Tiempo	9
2.2 Ejemplos y aplicaciones de las Series de Tiempo	10
2.3 Ecuaciones en Diferencia para procesos deterministas	18
2.4 Operador de rezago L	36
3 Procesos Estacionarios y Modelos Univariados	41
3.1 Definición de ergodicidad y estacionariedad	41
3.2 Función de autocorrelación	47
3.3 Procesos estacionarios univariados	54
3.4 Procesos Autoregresivos (AR)	55
3.5 Procesos de Medias Móviles (MA)	78
3.6 Procesos ARMA(p, q) y ARIMA(p, d, q)	84
3.7 Función de Autocorrelación Parcial	91
3.8 Selección de las constantes p, q, d en un AR(p), un MA(q), un ARMA(p, q) o un ARIMA(p, d, q)	93
3.9 Pronósticos	97
3.10 Desestacionalización y filtrado de Series	101

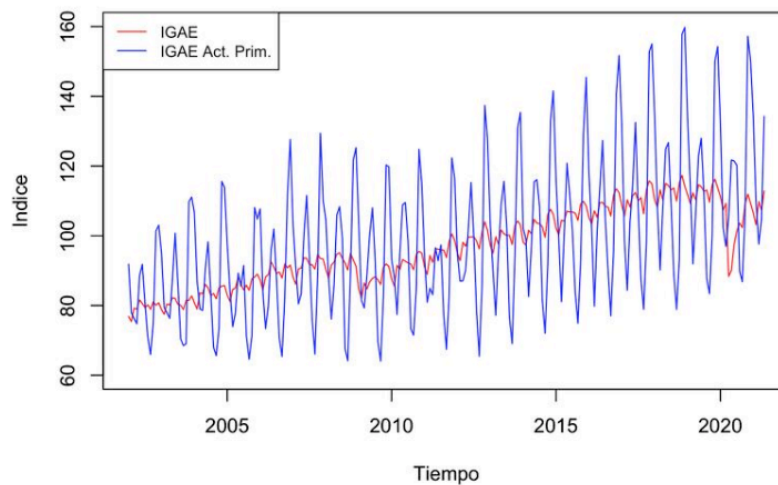
3.11 Motivación	101
3.12 Filtro Hodrick-Prescott	105
4 Procesos No Estacionarios Univariados	111
4.1 Definición y formas de No Estacionariedad	111
4.2 Pruebas de Raíces Unitarias	114
5 Causalidad de Granger, Modelos Multivariados de Vectores Autoregresivos y de Cointegración, Modelos ARDL y Filtro Multivariado	141
5.1 Causalidad de Granger	142
5.2 Definición y representación del Sistema o Modelo de Vectores Autorregresivos (VAR(p))	146
5.3 Análisis de Impulso-Respuesta	162
5.4 Identificación de los Choques Estructurales en los Modelos VAR	168
5.5 Cointegración	180
5.6 Modelos ADRL	193
5.7 Filtro Kalman	200
6 Modelos Univariados y Multivariados de Volatilidad	215
6.1 Motivación	215
6.2 Prueba de normalidad	221
6.3 Modelos ARCH y GARCH Univariados	222
6.4 Ejemplo GARCH(0,1)	232
6.5 Modelos ARCH y GARCH Multivariados	239
6.6 Pruebas para detectar efectos ARCH	241
7 Modelos de Datos Panel, Panel VAR y otros Modelos No Lineales	243
7.1 Modelos de Datos Panel y Panel VAR	243
7.2 Ejemplos: Panel VAR(p)	260
7.3 Cointegración	270
7.4 Modelos No Lineales	271
7.5 Ejemplo: Modelos de Cambio de Régimen (TAR)	274

Análisis de Series de Tiempo

Notas de Clase

Otoño 2024

by Benjamín Oliva, Omar Alfaro Rivera, Emiliano Pérez Caullieres & Luis Gerardo Ruiz



Contexto del documento

Estas son notas de clase de la materia de Análisis de Series de Tiempo de la Facultad de Economía de la Universidad Nacional Autónoma de México.

Este es un trabajo siempre en proceso de mejora. Para cualquier comentario o aclaración escribir a los correos benjov@ciencias.unam.mx o omarxalpha@gmail.com.

La versión PDF de estas notas se encuentra en: <https://github.com/benjov/Series-Tiempo/blob/main/docs/Notas-Series-Tiempo.pdf>

Las tablas de datos empleadas en los diferentes ejemplos se encuentran en: <https://github.com/benjov/Series-Tiempo/tree/main/BD>

Chapter 1

Introducción

Estas notas son un resumen, una síntesis comparativa y, en algunos casos, una interpretación propia de los libros de texto de Brooks (2019), Cowpertwait y Metcalfe (2009), Guerrero-Guzmán (2014), Enders (2015), Franses y van Dijk (2003), Kirchgassner, Wolters, y Hassler (2012), Lutkepohl (2005), Wei (2019), entre otros. En algunos casos se incorpora información adicional para efectos de dar contexto al tema analizado (ver sección de Bibliografía para mayores detalles).

El objetivo de este documento es proporcionar un conjunto de apuntes o notas que sirvan de apoyo para la clase de Series de Tiempo en la Facultad de Economía de la UNAM. Por esta razón, no deben considerarse como notas exhaustivas o como un sustituto de la clase y los laboratorios. Asimismo, es deseable que los alumnos puedan aportar sus observaciones y correcciones a estas notas, las observaciones a estas notas son esperadas y siempre serán bienvenidas y agradecidas.

Este es un trabajo siempre en proceso de mejora, para cualquier comentario o aclaración, contactar al correo benjov@ciencias.unam.mx o omarxalpha@gmail.com.

En estas notas se estudian los temas que típicamente son incluidos como parte de un curso estándar de análisis de series de tiempo y agrega otros tantos, los cuales son:

1. Modelos estacionarios univariados: $AR(p)$, $MA(q)$, $ARMA(p, q)$ y $ARIMA(p, d, q)$, y filtros para eliminar estacionalidad, entre otros;
2. Modelos no estacionarios univariados y Pruebas de raíz unitaria (o pruebas para determinar que una serie es estacionaria);
3. Modelos multivariados, entre lo que se incluye a los Vectores Autoregresivos (VAR) y los procedimientos de Cointegración

4. Modelación de series univariadas con errores con heterocedasticidad y autocorrelación: ARCH(r), GARCH(n), etc.;
5. Modelos multivariados con errores con heterocedasticidad y autocorrelación: M-GARCH y M-GARCH-M;
6. Casos particulares en los que las series incluidas en un modelo multivariado no son del mismo orden de integración, conocidos como modelos ADRL.
7. Modelos de Datos Panel en series de tiempo, y
8. Modelos no lineales como los de cambios de régimen.

Chapter 2

Introducción al Análisis de Series de Tiempo y Elementos de Ecuaciones en Diferencia

2.1 La naturaleza de los datos de Series de Tiempo

El análisis de series de tiempo tiene muchas aplicaciones en diversos campos de la ciencia. Por ejemplo, en la economía continuamente se está expuesto a información recopilada de los mercados financieros, indicadores de empleo, índices o indicadores del nivel de producción, índices de precios, etc. En otros campos de las ciencias sociales se emplea el análisis de series de tiempo para analizar la evolución de la población, los nacimientos, o el número de personas con matrículas escolares. Finalmente, en las ciencias exactas se pueden encontrar casos como los de un epidemiólogo que puede estar interesado en el número de casos de influenza observados en algún periodo de tiempo dado y si a estos se les puede asociar con algún tipo de estacionalidad o si se trata del inicio de un fenómeno atípico.

La primera aproximación que se suele tener a las series de tiempo es mediante el examen de datos puestos en una gráfica, en la cual uno de los ejes es el tiempo y el otro es el valor tomado por la variable. No obstante, en este tipo de exámenes existen dos enfoques. Por un lado, existe el enfoque de la importancia del tiempo, el cual consiste en reconocer cómo lo que sucede hoy es afectado por lo que pasó ayer o, en general, en periodos pasados, o cómo lo que pasa

hoy afectará los eventos futuros. Por otro lado, existe el enfoque del análisis frecuentista o de frecuencia, mediante el cual se busca reconocer la importancia que tiene para los investigadores los ciclos –por ejemplo, ciclos estacionales, momentos de crisis económicas, etc.–

2.2 Ejemplos y aplicaciones de las Series de Tiempo

Un primer ejemplo que puede ilustrar la presencia de los dos tipos de enfoques antes mencionados es la Figura 2.1. En esta figura se muestra la evolución del Indicador Global de la Actividad Económica (IGAE) en su versión global o del total de la economía y en su versión únicamente para las actividades primarias entre enero de 2002 y mayo de 2023.

```
library(readxl)

Base_1 <- read_excel("BD/Base_1_TimeSeries.xlsx")
IGAE_2013 <- ts(Base_1$IGAE_2013, start = 2002, freq = 12)
IGAE_PRIM_2013 <- ts(Base_1$IGAE_PRIM_2013, start = 2002, freq = 12)
ICC <- ts(Base_1$ICC, start = 2002, freq = 12)
ICC_LAG <- ts(Base_1$ICC_LAG, start = 2002, freq = 12)
IPC_BMV <- ts(Base_1$IPC_BMV, start = 2002, freq = 12)
TDC <- ts(Base_1$TDC, start = 2002, freq = 12)

plot(IGAE_2013, type = "l", lwd = 1, col = "red", ylab = "Indice",
      xlab = "Tiempo", ylim = c(60,160))
par(new = T)
# Indicador Global de la Actividad Económica, Act. Prim., base 2008
plot(IGAE_PRIM_2013, type = "l", lwd = 1, col = "blue",
      ylab = "Indice", xlab = "Tiempo", ylim = c(60,160))
# Leyenda
legend("topleft", c("IGAE", "IGAE Act. Prim."), cex = 0.8,
      lty = 1:1, col = c("red", "blue"))

par(new = F)
```

Como se puede observar, el IGAE del total de la economía muestra, principalmente, que el enfoque del tiempo es más relevante. Es decir, que existe cierta persistencia en el indicador, lo que significa que la economía crece en razón del crecimiento reportado en períodos pasados. No obstante, lo que no podemos reconocer es que los eventos futuros tienen un efecto en el desempeño de la economía hoy día. Así, no es común observar cambios abruptos del indicador, salvo por la crisis global de 2008 y la reciente crisis causada por la COVID-19.

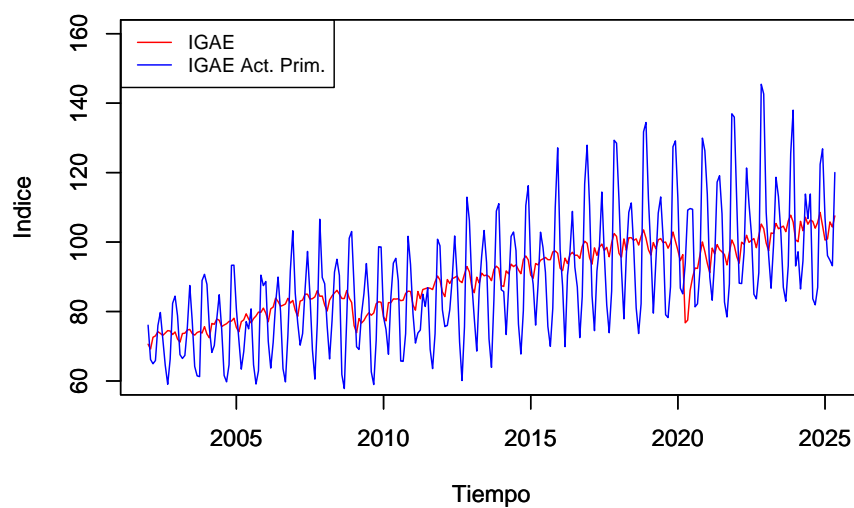


Figure 2.1: Indicador Global de Actividad Económica (IGAE) Global y para las Actividades Primarias (2008=100), Ene.2002 - May.2023

Por el contrario, el IGAE de las actividades primarias muestra una presencia significativa de la importancia de la frecuencia. No pasa desapercibido que existen muchos ciclos en la evolución del indicador. Algo que suena común en las actividades primarias, cuya producción depende de eventos que son cíclicos y que están asociados con el clima u otros factores determinantes de la oferta de productos agrícolas. Otro factor que puede influir en el indicador son elementos de demanda, más que los de oferta. Por ejemplo, el consumo de alimentos típicos de algunas temporadas del año.

Como segundo ejemplo, en la Figura 2.2 se ilustra la evolución reciente del índice de Confianza del Consumidor (ICC) en dos de sus versiones: i) el Índice global y ii) el Índice de confianza de los consumidores cuando estos consideran la situación actual en la economía en relación el año anterior.

```
plot(ICC, type = "l", lwd = 1, col = "red", ylab = "Indice",
     xlab = "Tiempo", ylim = c(29, 50))
# Comando que indica a R que sin borrar la grafica anterior,
# grafique la siguiente.
par(new = T)
# Indice ??Como considera usted la situacion economica del pais
# hoy en dia comparada con la de hace 12 meses?, base enero 2003
plot(ICC_LAG, type = "l", lwd = 1, col = "blue", ylab = "Indice",
     xlab = "Tiempo", ylim = c(29,50))
# Leyenda
legend("bottomleft", c("ICC","ICC lag"), cex = 0.8, lty = 1:1,
     col = c("red", "blue"))
```

```
par(new = F)
```

Destacamos que el ICC mide las expectativas de los consumidores en razón de la información pasada y de la esperada, según dichos consumidores. Así, es probable que las dos series de tiempo exhiban un gran peso para los eventos pasados, pero a la vez, un componente –probablemente menor– del componente de frecuencia. Esto último en razón de que los consumidores suelen considerar en sus expectativas de consumo los períodos cíclicos de la economía: temporadas navideñas, pagos de colegiaturas, etc. Este segundo ejemplo también ilustra que la confianza del consumidor no necesariamente está directamente correlacionada con el desempeño de la economía.

Como tercer ejemplo se muestra la evolución de dos series. La Figura 2.3 ilustra el comportamiento reciente de dos indicadores que son referencia para los inversionistas. Por un lado, se ubica el índice de Precios y Cotizaciones de la BMV (IPC), el cual refleja el valor de las acciones de empresas que cotizan en la BMV y el volumen de acciones comercializadas, en conjunto. En el mundo de las finanzas se ha interpretado que el IPC refleja el rendimiento del capital promedio invertido en las empresas que cotizan en la BMV.

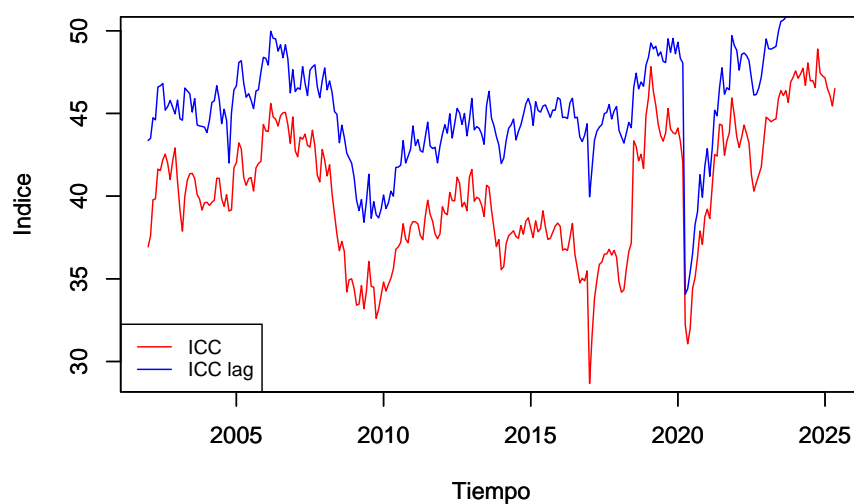


Figure 2.2: Índice de Confianza del Consumidor (ICC): General y resultado de ¿Cómo considera usted la situación económica del país hoy en día comparada con la de hace 12 meses? (puntos), Ene.2002-may.2023

```
par(mfrow=c(1,2))

# Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores
plot(IPC_BMV, type = "l", lwd = 1, col = "red", ylab = "Indice",
      xlab = "Tiempo", main = "Índice de Precios y \nCotizaciones BMV")
# Tipo de Cambio para Solventar Obligaciones en Moneda Extranjera
plot(TDC, type = "l", lwd = 1, col = "blue", ylab = "Pesos X Dolar",
      xlab = "Tiempo", main = "Tipo de Cambio")
```

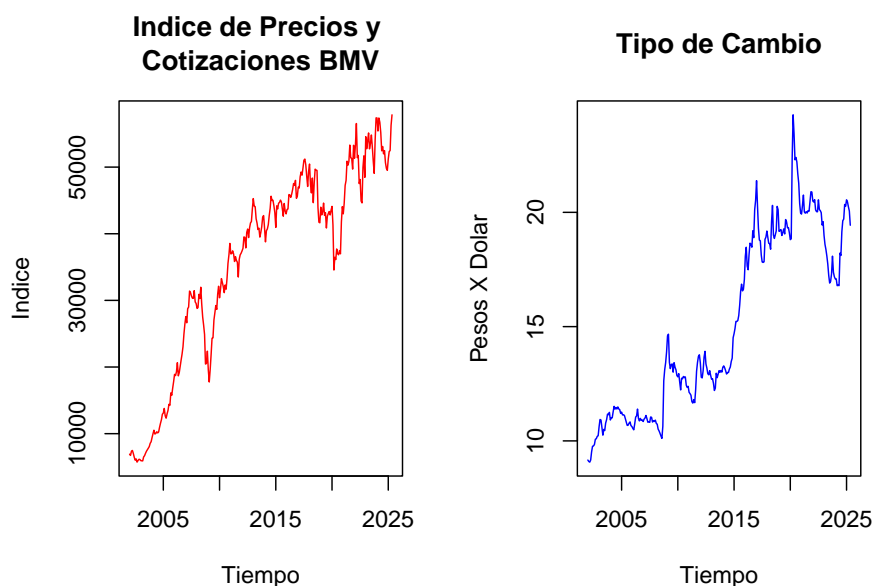


Figure 2.3: índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (Panel Derecho) y Tipo de Cambio para Solventar Obligaciones en Moneda Extranjera, pesos por dólar (Panel izquierdo), Ene.2002-May.2023

```
par(mfrow=c(1,1))
```

Por otro lado, en la Figura 2.3 se presenta la evolución del Tipo de Cambio (TDC) –indicador financiero que se suele utilizar como medio de reserva de valor–. Esto, en razón de que el TDC es conocido como un instrumento que en momentos de crisis toma valores contracíclicos de la economía mexicana. No obstante, ambos indicadores no son comparables. Para hacerlos comparables en la Figura 2.4 se presentan en versión índice con una base en el primer mes de la muestra.

```

IPC_BMV_I <- 100*IPC_BMV/IPC_BMV[1]
TDC_I <- 100*TDC/TDC[1]
# Índice del índice de Precios y Cotizaciones de la Bolsa Mexicana
# de Valores
plot(IPC_BMV_I, type = "l", lwd = 1, col = "red", ylab = "Indice",
      xlab = "Tiempo", ylim = c(80,740))
# Comando que indica a R que sin borrar la grafica anterior,
# grafique la siguiente.
par(new = T)
# Índice del Tipo de Cambio para Solventar Obligaciones en
# Moneda Extranjera
plot(TDC_I, type = "l", lwd = 1, col = "blue", ylab = "Indice",
      xlab = "Tiempo", ylim = c(80,740))
# Leyenda
legend("topleft", c("Índice del IPC", "Índice del TDC"), cex = 0.8,
      lty = 1:1, col = c("red", "blue"))

```

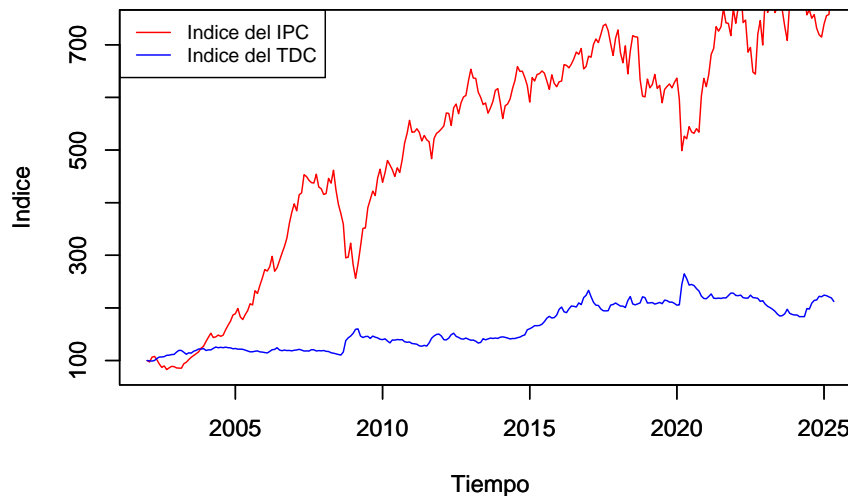


Figure 2.4: Índice del índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores e Índice del Tipo de Cambio para Solventar Obligaciones en Moneda Extranjera (ambos, enero de 2002 = 100), pesos por dólar, Ene.2002-May.2023

```
par(new = F)
```

En la perspectiva de la Figura 2.4 se puede apreciar que el TDC no es tan rentable, ya que una inversión en la BMV mediante el IPC, en el largo plazo, muestra más rendimientos. Asimismo, la Figura 2.4 ilustra que en ambas series se observa un dominio de la condición de tiempo y no uno de la frecuencia. Es decir, tanto el IPC como el TDC no responden a condiciones como ciclos o temporadas que sí son observables en actividades económicas como las primarias.

Finalmente, la Figura 2.5 ilustra una característica que también resulta de gran interés en el análisis de series de tiempo: los datos de alta frecuencia y de comportamiento no regular. En la Figura 2.5 se muestran las diferencias logarítmicas de las series de IGAE de la actividad total, el IPC y el TDC.

```
par(mfrow=c(3,1))
# Indicador Global de la Actividad Económica, base 2008
plot(diff(log(IGAE_2013), lag = 1), type = "l", lwd = 1,
     col = "darkred", ylab = "Var. %", xlab = "Tiempo",
     main = "Indicador Global de la Actividad Economica")
# Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores
plot(diff(log(IPC_BMV), lag = 1), type = "l", lwd = 1,
     col = "darkgreen", ylab = "Var. %", xlab = "Tiempo",
     main = "Índice de Precios y Cotizaciones BMV")
# Tipo de Cambio para Solventar Obligaciones en Moneda Extranjera
plot(diff(log(TDC), lag = 1), type = "l", lwd = 1,
     col = "darkblue", ylab = "Pesos X Dolar", xlab = "Tiempo",
     main = "Tipo de Cambio")
```

```
par(mfrow=c(1,1))
```

Dichas diferencias se pueden interpretar como una tasa de crecimiento de las series por las siguientes razones. Consideremos una serie de tiempo dada por y_t , cuya versión logarítmica es $\ln(y_t)$. De esta forma, la diferencia logarítmica está dada por la ecuación (2.1):

$$\Delta \ln(y_t) = \ln(y_t) - \ln(y_{t-1}) = \ln\left(\frac{y_t}{y_{t-1}}\right) \quad (2.1)$$

Ahora bien, si retomamos la definición de tasa de crecimiento (TC) de una serie de tiempo y_t entre el periodo t y $t - 1$ podemos obtener que:

$$TC = \frac{y_t - y_{t-1}}{y_{t-1}} = \frac{y_t}{y_{t-1}} - 1 \quad (2.2)$$

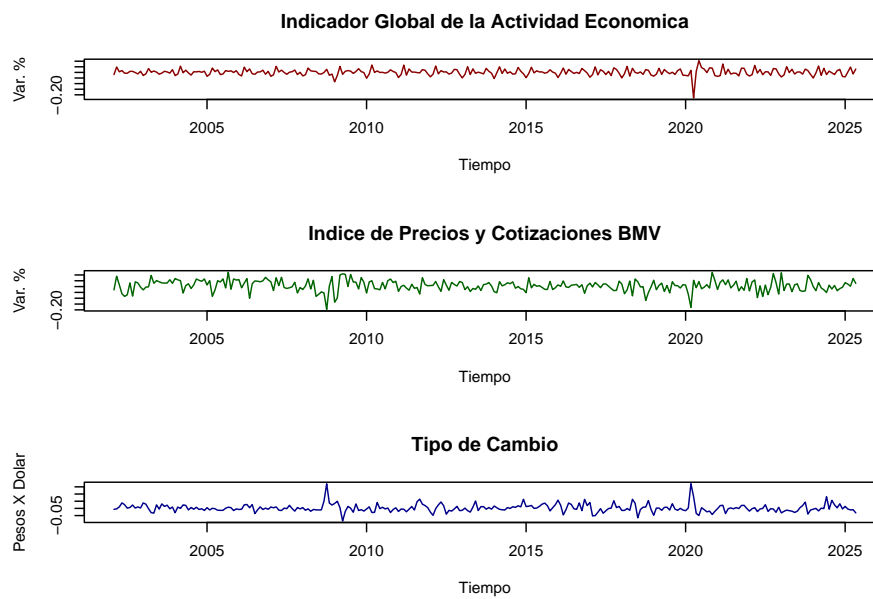


Figure 2.5: Tasas de Crecimiento mensuales (diferencias logarítmicas) de Indicador Global de la Actividad Económica, Índice de Precios y Cotizaciones de la Bolsa Mexicana de Valores (Panel Derecho) y Tipo de Cambio para Solventar Obligaciones en Moneda Extranjera, Ene.2002-May.2023

De esta forma, si tomamos el logaritmo de la expresión de la ecuación (2.2) obtenemos la siguiente aproximación:

$$\frac{y_t}{y_{t-1}} - 1 \approx \ln\left(\frac{y_t}{y_{t-1}}\right) = \ln(y_t) - \ln(y_{t-1}) \quad (2.3)$$

La ecuación (2.3) es cierta cuando los valores de y_t y y_{t-1} son muy parecidos, es decir, cuando las variaciones no son tan abruptas. Otra forma de interpretar la ecuación (2.3) es que para tasas de crecimiento pequeñas, se puede utilizar como una buena aproximación a la diferencia logarítmica mostrada en la ecuación (2.1).

En la Figura 2.5 se reportan las diferencias logarítmicas del IGAE, IPC y TDC, todos, como una media de distintos tipos de rendimientos. Es decir, podemos decir que un capitalista promedio (suponiendo que solo puede invertir en la actividad económica, en la bolsa o en el dólar), puede observar que le es más redituable en función de sus preferencias.

Notése que la dinámica de las variaciones de cada una de las series es significativamente diferente. Destaca que el TDC es una de las variables que, en general, no muestra grandes cambios a lo largo del tiempo. No obstante, se han observado cambios radicales, cuando menos en el año 2008 y durante la pandemia de COVID-19. Esta situación también se ha observado para el IPC. En cambio, el IGAE muestra un comportamiento más estable o estacionario –concepto que abordaremos más adelante–.

2.3 Ecuaciones en Diferencia para procesos deterministas

En las secciones previas se hizo una introducción al concepto de series de tiempo. En esta sección se pretende desarrollar la construcción de los procesos generadores de datos de las series de tiempo. En un sentido más formal, se expondrá que las series de tiempo se pueden considerar como una secuencia de variables aleatorias.

Para tales efectos, se desarrollará una introducción al concepto de ecuaciones en diferencia. Así, las preguntas que se pretende responder son:

1. ¿Cuál es la solución de la ecuación en diferencia que se estudia?
2. ¿Cuáles son las condiciones para que un proceso estocástico, representado mediante una ecuación en diferencia, llegue a alcanzar un punto de equilibrio en el largo plazo?

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 19

El término de *ecuación en diferencia* sirve para denominar un proceso similar o equivalente dentro de las ecuaciones diferenciales, dentro del cual se consideran a un conjunto de variables que están en función del tiempo. Así, si consideramos al tiempo como una variable continua, es decir, consideramos una variable $Z(t)$, podemos expresar las siguientes expresiones para la ecuación diferencial:

$$\frac{dZ(t)}{dt}; \frac{d^2Z(t)}{dt^2}; \dots; \frac{d^kZ(t)}{dt^k} \quad (2.4)$$

Por otro lado, suponiendo el caso del tiempo en forma discreta, es decir, con $t = \dots, -2, -1, 0, 1, 2, \dots$, entonces el comportamiento de la serie de variables dadas por Z_t , la cual se puede expresar como:

$$\Delta Z_t; \Delta^2 Z_t; \dots; \Delta^k Z_t \quad (2.5)$$

Observemos que una forma técnicamente más correcta es escribir las expresiones anteriores como:

$$\frac{\Delta Z_t}{\Delta t}; \frac{\Delta^2 Z_t}{\Delta t^2}; \dots; \frac{\Delta^k Z_t}{\Delta t^k} \quad (2.6)$$

No obstante, no pasa desapercibido que $\Delta t = 1$, por lo que resultan equivalentes ambos conjuntos de expresiones (2.5) y (2.6).

2.3.1 Ecuaciones en Diferencia Lineales de Primer Orden

El primer caso que se suele estudiar en relación a Ecuaciones en Diferencia es el de las Ecuaciones en Diferencia Lineales de Primer Orden. Al respecto, al igual que en el caso continuo, las variaciones de la variable Z_t se pueden expresar como se ilustra en el siguiente ejemplo. Consideremos la siguiente ecuación:

$$Z_t = a_0 + a_1 Z_{t-1} \quad (2.7)$$

Donde, $t = \dots, -2, -1, 0, 1, 2, \dots$, y a_0 y $a_1 \neq 0$ son números reales constantes. De (2.7) podemos despejar la variable Z_{t-1} y obtener una forma de ecuación en diferencia:

$$Z_t - a_1 Z_{t-1} = a_0 \quad (2.8)$$

Ahora denotemos a $LZ_t = Z_{t-1}$, es decir, mediante el operador L se puede rezagar una variable dada. En general, podemos decir que el operador tiene dos propiedades, la primera es que es lineal –en el sentido de que abre sumas y saca escalares– como se muestra en la siguiente expresión para el caso de un (1) rezago:

$$L(\alpha Z_t + \beta) = \alpha Z_{t-1} + \beta \quad (2.9)$$

Donde $\alpha, \beta \in \mathbb{R}$ y $\alpha, \beta \neq 0$. Otro resultado implícito en esta primera propiedad es que el operador rezago aplicado a cualquier escalar dará como resultado el

escalar, puesto que este es una constante sin importar el momento t en el cual se encuentre la variable Z .

La segunda propiedad del operador es que se puede aplicar de forma consecutiva a una misma variable. Es decir, $L(Z_{t-1}) = LLZ_t = L^2Z_t$, por lo que en general tendremos: $L^pZ_t = Z_{t-p}$ (con $p \in \mathbb{Z}$). Así, en el caso de p rezagos la propiedad de linealidad del operador rezago será:

$$L^p(\alpha Z_t + \beta) = \alpha Z_{t-p} + \beta \quad (2.10)$$

Dicho lo anterior, podemos escribir la solución general de (2.8) como:

$$\begin{aligned} Z_t - a_1 LZ_t &= a_0 \\ (1 - a_1 L)Z_t &= a_0 \\ Z_t &= a_0 \frac{1}{1 - a_1 L} + sa_1^t \\ Z_t &= a_0 \frac{1}{1 - a_1} + sa_1^t \end{aligned} \quad (2.11)$$

Donde $a_1 \neq 1$ y $t = \dots, -2, -1, 0, 1, 2, \dots$. Notése que la aplicación del operador rezago L a la constante a_1 dará como resultado el valor de la misma constante, ya que ésta no depende del momento t en el cual observemos a la variable Z_t . En la ecuación (2.11) se adiciona un término sa_1^t que permite ubicar la trayectoria inicial de la solución de la ecuación. El componente no significa un cambio respecto de la ecuación (2.8) original, ya que si buscáramos reconstruir esta ecuación tendríamos:

$$\begin{aligned} (1 - a_1 L)sa_1^t &= sa_1^t - a_1 sLa_1^t \\ &= sa_1^t - a_1 sa_1^{t-1} \\ &= sa_1^t - sa_1^t \\ &= 0 \end{aligned}$$

La ecuación (2.11) se suele interpretar como la solución de largo plazo. Ahora demostraremos por qué es cierta la ecuación y discutiremos algunas condiciones que se deben observar en esta solución para que sea una solución convergente. No obstante, primero discutiremos un método indirecto e incompleto para demostrar el resultado. Dicho método es conocido como el método iterativo. Plantearemos las siguientes ecuaciones particulares donde suponemos la existencia del valor inicial Z_0 del proceso:

$$Z_1 = a_0 + a_1 Z_0$$

$$\begin{aligned} Z_2 &= a_0 + a_1 Z_1 \\ &= a_0 + a_1(a_0 + a_1 Z_0) \\ &= a_0 + a_0 a_1 + a_1^2 Z_0 \\ &= a_0(1 + a_1) + a_1^2 Z_0 \end{aligned}$$

$$\begin{aligned}
Z_3 &= a_0 + a_1 Z_2 \\
&= a_0 + a_1(a_0 + a_0 a_1 + a_1^2 Z_0) \\
&= a_0 + a_0 a_1 + a_0 a_1^2 + a_1^3 Z_0 \\
&= a_0(1 + a_1 + a_1^2) + a_1^3 Z_0
\end{aligned}$$

De lo anterior se puede inferir que el método iterativo convergerá hacia una expresión como la siguiente en el momento t :

$$\begin{aligned}
Z_t &= a_0 + a_1 Z_{t-1} \\
&= a_0(1 + a_1 + a_1^2 + \dots + a_1^{t-1}) + a_1^t Z_0 \\
&= a_0 \sum_{i=0}^{t-1} a_1^i + a_1^t Z_0
\end{aligned} \tag{2.12}$$

Donde, es necesario que en la ecuación (2.12) se cumpla que $|a_1| < 1$ para que la suma sea convergente –más adelante detallaremos esta afirmación–. A este tipo de ecuaciones se les puede denominar como lineales. Esto en razón de que ningún término de la variable Z aparece elevado a ninguna potencia distinta a 1. También, son de primer orden, ya que el rezago de la variable Z es sólo de un período.

En adelante trabajaremos con ecuaciones en las que la variable Z se encuentra rezagada en cualquiera de los siguientes casos:

$$Z_t, Z_{t-1}, Z_{t-2}, Z_{t-3}, \dots, Z_{t-p}, \dots \tag{2.13}$$

Por lo que diremos que en adelante el curso versará sobre ecuaciones en diferencia lineales y de cualquier orden $p \in \mathbb{N}$.

Retomando la ecuación (2.12) y considerando la parte de la suma de los términos de a_1^i , de tal forma que buscaremos dar una expresión más comprensible a dicho término. Definamos la siguiente expresión como:

$$S_{t-1} = \sum_{i=0}^{t-1} a_1^i \tag{2.14}$$

Por lo tanto, S_t estaría dado por la siguiente expresión:

$$\begin{aligned}
S_t &= a_1 \sum_{i=0}^{t-1} a_1^i \\
&= a_1(1 + a_1 + a_1^2 + \dots + a_1^{t-1}) \\
&= a_1 + a_1^2 + a_1^3 + \dots + a_1^t \\
&= a_1 S_{t-1}
\end{aligned} \tag{2.15}$$

Tomando los dos resultados de las ecuaciones (2.14) y (2.15) anteriores, podemos expresar que si a S_{t-1} le restamos S_t , y desarrollando ambos lados de la ecuación anterior podemos obtener:

$$\begin{aligned} S_{t-1} - a_1 S_{t-1} &= S_{t-1} - S_t \\ (1 - a_1) S_{t-1} &= (1 + a_1 + a_1^2 + \dots + a_1^{t-1}) - (a_1 + a_1^2 + a_1^3 + \dots + a_1^t) \\ (1 - a_1) S_{t-1} &= 1 - a_1^t \end{aligned}$$

Así, podemos concluir que:

$$S_{t-1} = \frac{1 - a_1^t}{1 - a_1} \quad (2.16)$$

Conjuntando este último resultado de la ecuación (2.16) con la ecuación (2.12) tenemos la siguiente solución por el método de iteración:

$$Z_t = a_0 \left(\frac{1 - a_1^t}{1 - a_1} \right) + a_1^t Z_0 \quad (2.17)$$

De esta forma la ecuación (2.17) es una solución para la ecuación (2.12), que es una ecuación de un proceso de una Ecuación en Diferencia planteado en la ecuación (2.7). Esta solución aún no es general, en el sentido de que sea válida para cualquier tipo de proceso: convergente o divergente. Dicha convergencia o divergencia estará determinada por el parámetro a_1 . No debe pasar desapercibido que cuando $t \rightarrow \infty$ o cuando la muestra es muy grande (lo que es equivalente), podemos decir que la solución solo puede converger a la siguiente expresión cuando se considera que $|a_1| < 1$:

$$Z_t = a_0 \left(\frac{1}{1 - a_1} \right) \quad (2.18)$$

Retomemos ahora el caso general descrito en la ecuación (2.11) y determinemos una solución general en la cual $a_1 \neq 1$ y $t = \dots, -2, -1, 0, 1, 2, \dots$. Para ello, observemos que el siguiente componente en la ecuación mencionada se puede interpretar como la suma infinita de términos descritos como:

$$\begin{aligned} \frac{1}{1 - a_1} &= 1 + a_1 + a_1^2 + \dots + a_1^t + \dots \\ &= \sum_{i=0}^{\infty} a_1^i \end{aligned} \quad (2.19)$$

Donde claramente es necesario que $|a_1| < 1$. Por lo tanto, sólo faltaría determinar el valor de la constante s en la ecuación (2.11) de la siguiente forma, supongamos que observamos el proceso en el momento inicial, por lo que es posible determinar el valor de la constante conociendo el valor inicial del proceso como sigue:

$$Z_0 = a_0 \frac{1}{1 - a_1} + s \quad (2.20)$$

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 23

De la ecuación (2.20) tenemos que:

$$s = Z_0 - a_0 \frac{1}{1 - a_1} \quad (2.21)$$

Así, juntando la ecuación (2.11) y la ecuación (2.21) tenemos la expresión:

$$Z_t = a_0 \frac{1 - a_1^t}{1 - a_1} + a_1^t Z_0 \quad (2.22)$$

No debe pasar desapercibido que esta solución es la misma que la mostrada en la ecuación (2.17), por lo que en realidad ambas ecuaciones son una solución general indistintamente entre las ecuaciones (2.17) y (2.22). Ambas convergen a la expresión como la ecuación (2.18), con la misma condición de convergencia $|a_1| < 1$. Para ilustrar estas ecuaciones, veamos algunos ejemplos al respecto.

Ejemplo. Consideremos que tenemos un proceso Z_t que es descrito por una ecuación en diferencia lineal de primer orden dada por:

$$Z_t = 2 + 0.9Z_{t-1} \quad (2.23)$$

Siguiendo la expresión mostrada en la ecuación (2.22), obtenemos la expresión:

$$Z_t = 2 \left(\frac{1 - 0.9^t}{1 - 0.9} \right) + 0.9^t Z_0 \quad (2.24)$$

Donde asumiremos que el valor inicial es $Z_0 = 10$ y que la expresión debe converger al valor de 20, cuando t es muy grande o tiende a infinito.

Ejemplo. De forma similar, tomemos otro ejemplo en el cual asumimos la siguiente expresión:

$$Z_t = 2 - 0.5Z_{t-1} \quad (2.25)$$

Siguiendo la expresión mostrada en la ecuación (2.22), obtenemos:

$$Z_t = 2 \left(\frac{1 - (-0.5)^t}{1 + 0.5} \right) + (-0.5)^t Z_0 \quad (2.26)$$

Donde asumiremos que el valor inicial es $Z_0 = 10$ y que la ecuación converge al valor de 1.333333..., cuando t es muy grande o tiende a infinito. Ahora simulemos el comportamiento de ambos procesos y estableceremos los resultados del Cuadro 2.1. Notemos que el segundo proceso converge de una forma más rápida que el primero. El Cuadro 2.1 se ilustra en las siguientes dos Figura 2.6 y Figura 2.7.

```

library(knitr)
library(tidyverse)
library(kableExtra)

Tiempo = c(0:100)
Zt = rep(NA, 101)
Zt2 = rep(NA, 101)
Zt[1] = 10
Zt2[1] = 10

for (i in c(2:101)){
  Zt[i] = 2+0.9*Zt[i-1]
  Zt2[i] = 2 - 0.5*Zt2[i-1]
}
lista = c(1:16, 97:101)
Tiempo1 = Tiempo[lista]
Zt1=Zt[lista]
Zt21=Zt2[lista]

tabla1 = data.frame(Tiempo1, Zt1, Zt21)
colnames(tabla1) <- c("Tiempo", "$Z_t = 2+0.9Z_{t-1}$",
                      "$Z_t = 2-0.5Z_{t-1}$")

kable(tabla1,
      caption = "Dos ejemplos de Ecuaciones Lineales de Primer Orden",
      format = "pandoc") %>%
  kable_styling(font_size = 10)

```

Table 2.1: Dos ejemplos de Ecuaciones Lineales de Primer Orden

Tiempo	$Z_t = 2 + 0.9Z_{t-1}$	$Z_t = 2 - 0.5Z_{t-1}$
0	10.00000	10.000000
1	11.00000	-3.000000
2	11.90000	3.500000
3	12.71000	0.250000
4	13.43900	1.875000
5	14.09510	1.062500
6	14.68559	1.468750
7	15.21703	1.265625
8	15.69533	1.367188
9	16.12580	1.316406
10	16.51322	1.341797
11	16.86189	1.329102
12	17.17570	1.335449

Tiempo	$Z_t = 2 + 0.9Z_{t-1}$	$Z_t = 2 - 0.5Z_{t-1}$
13	17.45813	1.332275
14	17.71232	1.333862
15	17.94109	1.333069
96	19.99960	1.333333
97	19.99964	1.333333
98	19.99967	1.333333
99	19.99970	1.333333
100	19.99973	1.333333

```
tabla <- data.frame( Tiempo, Zt, Zt2 )

ggplot( data = tabla , aes(x = Tiempo, y = Zt) )+
  geom_line(col="blue4") +
  geom_point(col= "blue4") +
  labs(y=expression(Z[t]))
```

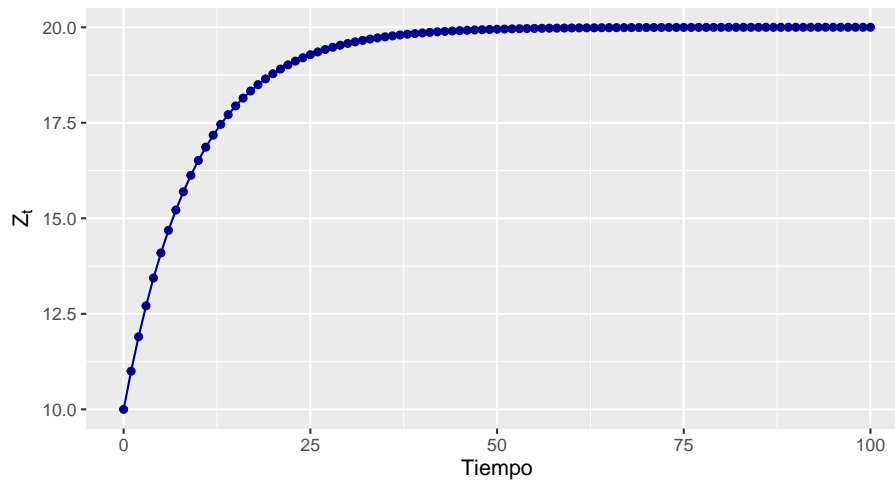


Figure 2.6: Evolución del proceso dado por $Z_t = 2 + 0.9Z_{t-1}$

```
ggplot(data = tabla , aes(x = Tiempo, y=Zt2)) +
  geom_line(col="red4") +
  geom_point(col= "red4") +
  labs(y=expression(Z[t]))
```

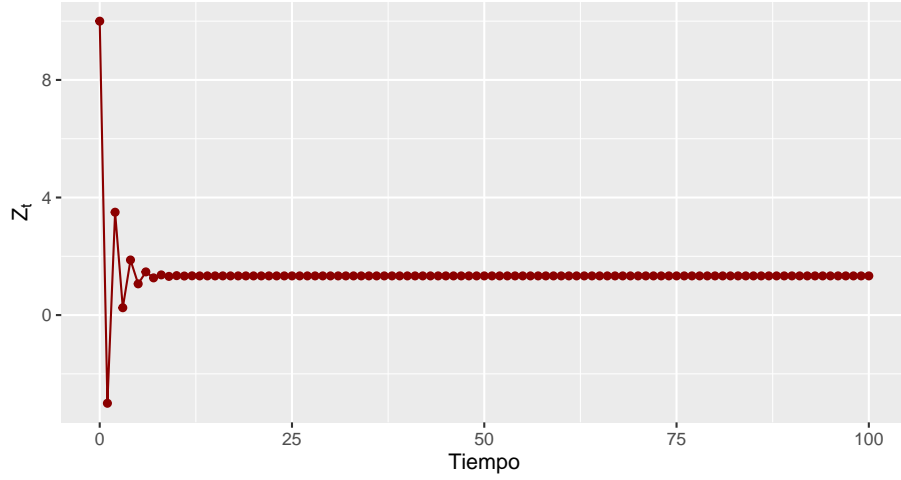


Figure 2.7: Evolución del proceso dado por $Z_t = 2 - 0.5Z_{t-1}$

2.3.2 Ecuaciones en Diferencia Lineales de Segundo Orden y de orden superior

Como un segundo caso a estudiar se ubica el caso de las Ecuaciones en Diferencia Lineales de Segundo Orden y, en su caso, de orden superior. Primero, sea una ecuación como la siguiente, la cual es lineal y de segundo orden, ya que tiene asociado un término de Z_t rezagado dos períodos:

$$Z_t = a_0 + a_1 Z_{t-1} + a_2 Z_{t-2} \quad (2.27)$$

Donde $t = \dots, -2, -1, 0, 1, 2, \dots$ y $a_1, a_2 \neq 0$. Reordenando la ecuación (2.27) podemos escribir:

$$\begin{aligned} Z_t - a_1 Z_{t-1} - a_2 Z_{t-2} &= a_0 \\ Z_t - a_1 L Z_t - a_2 L^2 Z_t &= a_0 \\ (1 - a_1 L - a_2 L^2) Z_t &= a_0 \end{aligned} \quad (2.28)$$

Así, la solución general propuesta para la ecuación (2.28) es la siguiente, la cual es una forma análoga a una Ecuación Lineal en Diferencia de Primer Orden:

$$Z_t = \frac{a_0}{1 - a_1 - a_2} + s_1 g_1^t + s_2 g_2^t \quad (2.29)$$

En donde s_1 y s_2 son constantes que se determinan mediante dos condiciones iniciales –por lo que para resolver este tipo de ecuaciones requerimos conocer

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 27

dos condiciones iniciales—. Los valores de g_1 y g_2 están relacionados con los coeficientes a_1 y a_2 , de esta forma:

$$a_1 = g_1 + g_2 \quad (2.30)$$

$$a_2 = -g_1 g_2 \quad (2.31)$$

Lo anterior surge del siguiente procedimiento y recordando que siempre es posible descomponer una ecuación cuadrática en expresiones como las siguientes:

$$\begin{aligned} (1 - a_1 L - a_2 L^2) &= (1 - g_1 L)(1 - g_2 L) \\ &= 1 - g_1 L - g_2 L + g_1 g_2 L^2 \\ &= 1 - (g_1 + g_2)L + g_1 g_2 L^2 \end{aligned} \quad (2.32)$$

Donde se observa la equivalencia mostrada en las ecuaciones (2.30) y (2.31). Así, considerando la ecuación (2.29) tenemos que:

$$\begin{aligned} (1 - a_1 L - a_2 L^2)Z_t &= (1 - g_1 L)(1 - g_2 L)Z_t \\ &= a_0 + (1 - g_1 L)(1 - g_2 L)s_1 g_1^t \\ &\quad + (1 - g_1 L)(1 - g_2 L)s_2 g_2^t \end{aligned} \quad (2.33)$$

Por lo tanto, buscamos que, para que el proceso sea equivalente y podamos interpretar que la ecuación (2.29) sea una solución general, deberá pasar lo siguiente:

$$(1 - g_1 L)(1 - g_2 L)s_1 g_1^t + (1 - g_1 L)(1 - g_2 L)s_2 g_2^t = 0 \quad (2.34)$$

O, escrito de otra forma:

$$(1 - g_1 L)s_1 g_1^t = (1 - g_2 L)s_2 g_2^t = 0 \quad (2.35)$$

Ahora determinemos cuáles son los valores g_1 y g_2 dados los valores a_1 y a_2 que nos permitan determinar si el proceso será convergente. Para ello debemos resolver la siguiente ecuación que se deriva de la ecuación (2.32):

$$1 - a_1 x - a_2 x^2 = (1 - g_1 x)(1 - g_2 x) = 0 \quad (2.36)$$

Donde, claramente existen dos raíces: $x_1 = g_1^{-1}$ y $x_2 = g_2^{-1}$. Así, la solución estará dada por las raíces de la ecuación característica:

$$\begin{aligned} 1 - a_1 x - a_2 x^2 &= 0 \\ a_2 x^2 + a_1 x - 1 &= 0 \end{aligned} \quad (2.37)$$

Cuya solución es:

$$x = \frac{-a_1 \pm \sqrt{a_1^2 + 4a_2}}{2a_2} \quad (2.38)$$

Es importante distinguir tres diferentes casos en relación con las raíces que surgen como solución de la ecuación (2.37), estos son:

Caso I. Si $a_1^2 + 4a_2 > 0$, la ecuación (2.37) proporcionará dos valores de raíces reales y distintos, eso es $x_1 = g_1^{-1} \neq x_2 = g_2^{-1}$. Si por ahora suponemos que $|g_1| < 1$ y que $|g_2| < 1$, entonces tendremos que:

$$\begin{aligned} (1 - g_1 L)^{-1} (1 - g_2 L)^{-1} a_0 &= \left(\sum_{j=0}^{\infty} g_1^j L^j \right) \left(\sum_{j=0}^{\infty} g_2^j L^j \right) a_0 \\ &= \left(\sum_{j=0}^{\infty} g_1^j \right) \left(\sum_{j=0}^{\infty} g_2^j \right) a_0 \\ &= \frac{a_0}{(1 - g_1)(1 - g_2)} \\ &= \frac{a_0}{1 - a_1 - a_2} \end{aligned} \quad (2.39)$$

Esto último es el punto de equilibrio de la ecuación (2.29); considerando que $|g_1| < 1$ y que $|g_2| < 1$ –notemos que los demás casos son divergentes, ya que la suma anterior no convergería–. De esta forma, la solución de la ecuación estará dada por:

$$\lim_{t \rightarrow \infty} Z_t = \frac{a_0}{1 - a_1 - a_2} \quad (2.40)$$

Caso II. Si $a_1^2 + 4a_2 < 0$ en la ecuación (2.37), entonces las raíces serán números complejos conjugados, es decir:

$$g_i^{-1} = a \pm ib \quad (2.41)$$

$$g_i = u \pm iv \quad (2.42)$$

Dichas raíces las podemos escribir en coordenadas polares como:

$$g_1^{-1} = r e^{i\theta} = r(\cos(\theta) + i \sin(\theta)) \quad (2.43)$$

$$g_2^{-1} = r e^{-i\theta} = r(\cos(\theta) - i \sin(\theta)) \quad (2.44)$$

Donde: $r = \sqrt{u^2 + v^2}$, a esta expresión también se le conoce como módulo. Alternativamente, podemos escribir que $r = \sqrt{g_1 g_2}$. La única condición es que $r < 1$ para que el proceso descrito en la ecuación (2.29) sea convergente.

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 29

Al igual que en el **Caso I**, el punto de equilibrio de la ecuación se debería ubicar alrededor de la ecuación (2.40), siempre que $r < 1$, por lo que el factor que determina la convergencia es el módulo, ya que si el módulo es mayor a 1, el proceso será divergente, pero si es menor a 1 convergerá a (2.40). Para ilustrar, el caso contrario es divergente puesto que representa trayectorias senoidales (oscilatorias) que sólo pueden converger si a medida que pasa el tiempo las ondas son menos amplias.

Analicemos la solución:

$$Z_t = \frac{a_0}{1 - a_1 - a_2} + s_1 g_1^t + s_2 g_2^t$$

Donde s_1 y s_2 las determinamos usando las condiciones iniciales al solucionar:

$$\begin{aligned} Z_0 &= \frac{a_0}{1 - a_1 - a_2} + s_1 g_1^t + s_2 g_2^t \\ Z_1 &= \frac{a_0}{1 - a_1 - a_2} + s_1 g_1^t + s_2 g_2^t \end{aligned}$$

Donde, una vez encontradas las g_1 y g_2 , las incógnitas son s_1 y s_2 . Para determinar g_1 y g_2 realizamos el siguiente procedimiento. Partimos de la ecuación:

$$1 - a_1 x - a_2 x^2 = 0$$

Al factorizar notamos que:

$$(1 - g_1 x)(1 - g_2 x) = 0$$

De lo anterior debe quedar claro que:

$$\begin{aligned} x_1 &= g_1^{-1} \\ x_2 &= g_2^{-1} \end{aligned}$$

Donde g_1 y g_2 se determinan:

$$x = \frac{-a_1 \pm \sqrt{a_1^2 + 4a_2}}{2a_2}$$

Asumiendo que obtienen números complejos:

$$\begin{aligned} x_1 &= g_1^{-1} = u + iv \\ x_2 &= g_2^{-1} = u - iv \end{aligned}$$

Las cuales se pueden reescribir en coordenadas polares:

$$\begin{aligned} g_1^{-1} &= se^{i\theta} = s[\cos(\theta) + i\sen(\theta)] \\ g_2^{-1} &= se^{-i\theta} = s[\cos(\theta) - i\sen(\theta)] \\ s &= \sqrt{u^2 + v^2} = \sqrt{g_1^{-1} \cdot g_2^{-1}} \\ \cos(\theta) &= \frac{u}{s} \\ \sen(\theta) &= \frac{v}{s} \end{aligned}$$

Recuerde que: $r = s^{-1} = \sqrt{g_1 \cdot g_2}$, de esta forma:

$$\begin{aligned} g_1 &= r[\cos(\theta) - i\sen(\theta)] \\ g_2 &= r[\cos(\theta) + i\sen(\theta)] \end{aligned}$$

Retomemos nuestra solución original:

$$\begin{aligned} Z_t &= \frac{a_0}{1 - a_1 - a_2} + s_1 g_1^t + s_2 g_2^t \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t [\cos(\theta) - i\sen(\theta)]^t + s_2 r^t [\cos(\theta) + i\sen(\theta)]^t \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t (e^{-i\theta})^t + s_2 r^t (e^{i\theta})^t \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t (e^{-it\theta}) + s_2 r^t (e^{it\theta}) \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t [\cos(\theta t) - i\sen(\theta t)] + s_2 r^t [\cos(\theta t) + i\sen(\theta t)] \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t [(s_1 + s_2)\cos(\theta t) - i(s_2 - s_1)\sen(\theta t)] \\ &= \frac{a_0}{1 - a_1 - a_2} + s_1 r^t [A\cos(\theta t) - B\sen(\theta t)] \end{aligned}$$

Donde A y B son constantes arbitrarias y la convergencia está definida por r , si $r > 1$ el proceso diverge, y converge si $r < 1$. Finalmente:

$$\begin{aligned} \theta &= \cos^{-1}\cos(\theta) \\ \theta &= \sen^{-1}\sen(\theta) \end{aligned}$$

Caso III. Ahora revisemos el caso en el que $a_1^2 + 4a_2 = 0$, de esta forma las raíces serán idénticas:

$$g = g_1^{-1} = g_2^{-1} = \frac{-a_1}{2a_2} \quad (2.45)$$

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 31

Así, el punto de equilibrio será dado por la solución descrita como:

$$\begin{aligned}(1 - gL)^2 Z_t &= a_0 \\ Z_t &= \frac{a_0}{(1 - gL)^2} + s_1 g^t + s_2 t g^t \\ &= a_0 \sum_{i=0}^{\infty} (1 + i) g^i + s_1 g^t + s_2 t g^t\end{aligned}\quad (2.46)$$

Donde la expresión anterior es resultado de considerar el siguiente procedimiento. Sea:

$$f(g) = \frac{1}{(1 - g)} = \sum_{j=0}^{\infty} g^j$$

Por lo que si hacemos la primer derivada de la expresión anterior tenemos que:

$$\begin{aligned}f'(g) &= \frac{1}{(1 - g)^2} \\ &= \sum_{j=0}^{\infty} j g^{j-1} \\ &= 0 + g^0 + 2g^1 + 3g^2 + \dots \\ &= \sum_{j=0}^{\infty} (1 + j) g^j\end{aligned}$$

Ahora veámos un ejemplo de una Ecuación Lineal en Diferencia de Segundo Orden. Supongamos la ecuación y el desarrollo siguientes:

$$\begin{aligned}Z_t &= 3 + 0.9Z_{t-1} - 0.2Z_{t-2} \\ (1 - 0.9L + 0.2L^2)Z_t &= 3\end{aligned}$$

La solución dada por una ecuación similar a la expresión (2.37), obtendríamos la solución dada por las ecuaciones equivalentes a:

$$\begin{aligned}1 - 0.9x + 0.2x^2 &= 0 \\ -0.2x^2 + 0.9x - 1 &= 0\end{aligned}$$

De donde las raíces del polinomio característico $x_1 = g_1^{-1}$ y $x_2 = g_2^{-1}$ se obtienen de la expresión dada por:

$$\begin{aligned}x &= \frac{-0.9 \pm \sqrt{0.81 + (4)(-0.2)}}{(2)(-0.2)} \\ &= \frac{0.9 \pm 0.1}{0.4}\end{aligned}$$

Dado que el componente $a_1^2 + 4a_2$ es positivo, obtendremos dos raíces reales. Las raíces estarán dadas por $x_1 = 2.5$ y $x_2 = 2.0$, de lo cual podemos determinar que $g_1 = 0.4$ y $g_2 = 0.5$. De esta forma tenemos que $|g_1| < 1$ y $|g_2| < 1$, así la ecuación converge a la expresión dada por las siguientes expresiones:

$$\begin{aligned} Z_t &= \frac{3}{1 - 0.9L + 0.2L^2} + s_1(0.4)^t + s_2(0.5)^t \\ &= \frac{3}{1 - 0.9 + 0.2} + s_1(0.4)^t + s_2(0.5)^t \\ &= \frac{3}{(1 - 0.4)(1 - 0.5)} + s_1(0.4)^t + s_2(0.5)^t \end{aligned}$$

Al final, la ecuación que describe la solución general será:

$$z_t = 10 + s_1(0.4)^t + s_2(0.5)^t \quad (2.47)$$

Para determinar los valores de s_1 y s_2 necesitamos obtener dos valores iniciales de la ecuación. Para lo cual iniciaremos con $t = 0$ y, luego, obtenemos el valor de $t = 1$, y consideremos el valor de $Z_0 = 0$ y $Z_1 = 50$:

$$\begin{aligned} Z_0 &= 10 + s_1(0.4)^0 + s_2(0.5)^0 \\ 0 &= 10 + s_1 + s_2 \\ Z_1 &= 10 + s_1(0.4)^1 + s_2(0.5)^1 \\ 50 &= 10 + 0.4s_1 + 0.5s_2 \end{aligned}$$

Por lo que la solución es: $s_1 = -450$ y $s_2 = 440$, de donde podemos expresar la ecuación como:

$$Z_t = 10 - 450(0.4)^t + 440(0.5)^t \quad (2.48)$$

La ecuación (2.48) anterior convergerá al valor de 10 cuando $t \rightarrow \infty$. Para ilustrar la trayectoria de esta ecuación tomemos un cuadro similar al de los ejemplos anteriores. En el Cuadro 2.2 y la Figura 2.8 mostramos los resultados de la trayectoria para 100 períodos.

```
t = ts(c(0:100))

Zt = 10-450*(0.4^t)+440*(0.5^t)

tabla_2 <- data.frame( t, Zt)

lista = c(1:16, 97:101)
t1 = t[lista]
Zt1=Zt[lista]
```


2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS33

```
tabla1 = data.frame(Tiempo1, Zt1)
colnames(tabla1) <- c("Tiempo", "$Z_t = 10-450(0.4)^t+440(0.5)^t$")

kable(tabla1,
      caption = "Un ejemplo de Ecuación de Segundo Orden",
      format = "pandoc") %>%
kable_styling(font_size = 10)
```

Table 2.2: Un ejemplo de Ecuación de Segundo Orden

Tiempo	$Z_t = 10 - 450(0.4)^t + 440(0.5)^t$
0	0.00000
1	50.00000
2	48.00000
3	36.20000
4	25.98000
5	19.14200
6	15.03180
7	12.70022
8	11.42384
9	10.74141
10	10.38250
11	10.19597
12	10.09987
13	10.05069
14	10.02565
15	10.01294
96	10.00000
97	10.00000
98	10.00000
99	10.00000
100	10.00000

```
ggplot(data = tabla_2, aes(x = t, y=Zt)) +
  geom_line(col="green4") +
  geom_point(col= "green4") +
  labs(y=expression(Z[t]))
```

Finalmente, discutiremos la solución para las Ecuaciones Lineales en Diferencia de Orden p , donde $p \geq 2$. En general, una ecuación de este tipo se puede escribir como:

$$Z_t = a_0 + a_1 Z_{t-1} + a_2 Z_{t-2} + \dots + a_p Z_{t-p} \quad (2.49)$$

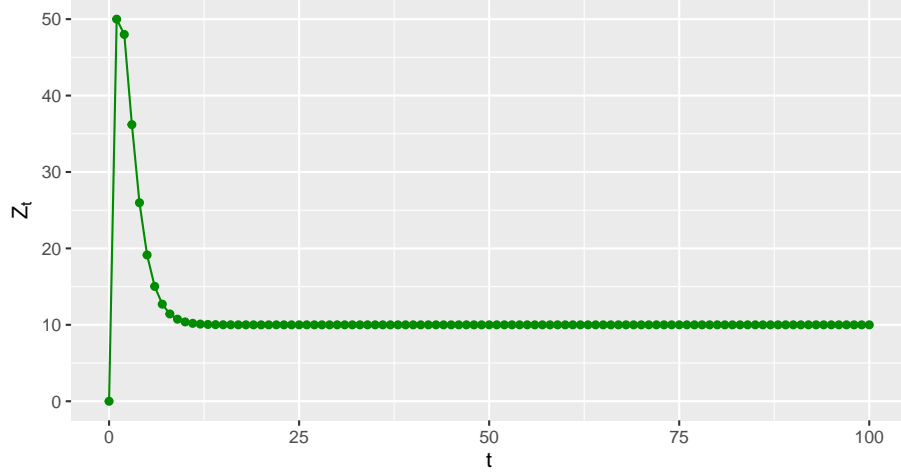


Figure 2.8: Evolución del proceso dado por $Z_t = 3 + 0.9Z_{t-1} - 0.2Z_{t-2}$

Donde $t = \dots, -2, -1, 0, 1, 2, \dots$ y $a_p \neq 0$. La ecuación (2.49) se puede escribir como:

$$\begin{aligned} Z_t - a_1 Z_{t-1} - a_2 Z_{t-2} - \dots - a_p Z_{t-p} &= a_0 \\ Z_t - a_1 L Z_t - a_2 L^2 Z_t - \dots - a_p L^p Z_t &= a_0 \\ (1 - a_1 L - a_2 L^2 - \dots - a_p L^p) Z_t &= a_0 \end{aligned} \quad (2.50)$$

Por el Teorema Fundamental del Álgebra es posible escribir la ecuación (2.50) como:

$$(1 - g_1 L)(1 - g_1 L) \dots (1 - g_p L) Z_t = a_0 \quad (2.51)$$

Utilizando la ecuación (2.50) y la ecuación (2.51) tenemos que la solución general de una ecuación como la descrita en (2.49) se puede escribir como:

$$Z_t = \frac{a_0}{1 - a_1 - a_2 - \dots - a_p} + s_1 g_1^t + s_2 g_2^t + \dots + s_p g_p^t \quad (2.52)$$

$$Z_t = \frac{a_0}{(1 - g_1)(1 - g_1) \dots (1 - g_p)} + s_1 g_1^t + s_2 g_2^t + \dots + s_p g_p^t \quad (2.53)$$

Donde s_1, s_2, \dots, s_p son constantes que se determinan utilizando p valores particulares de Z_t , y la solución general descrita en las ecuaciones (2.52) y

2.3. ECUACIONES EN DIFERENCIA PARA PROCESOS DETERMINISTAS 35

(2.53) implica encontrar p raíces: $x_1 = g_1^{-1}$, $x_2 = g_2^{-1}$, ..., $x_p = g_p^{-1}$ de los siguientes polinomios equivalentes:

$$(1 - g_1)(1 - g_2) \dots (1 - g_p) = 0 \quad (2.54)$$

$$1 - a_1x - a_2x^2 - \dots - a_px^p = 0 \quad (2.55)$$

$$a_px^p + \dots + a_2x^2 + a_1x - 1 = 0 \quad (2.56)$$

Antes de plantear la solución general, analicemos una solución particular cuando un conjunto de las p raíces, digamos un total de m , son iguales, es decir, cuando sucede que $g_1 = g_2 = \dots = g_m = g$ (con $1 < m \leq p$). En este caso la solución general en la ecuación (2.53) se escribe como:

$$Z_t = \frac{a_0}{(1 - g)^m(1 - g_{m+1}) \dots (1 - g_p)} + s_1g^t + s_2tg^t + \dots + s_mt^{m-1}g^t + s_{m+1}g_{m+1}^t + \dots + s_pg_p^t \quad (2.57)$$

Definamos:

$$f(g) = \frac{1}{1 - g} = \sum_{j=0}^{\infty} g^j \quad (2.58)$$

Si retomamos el método descrito párrafos arriba tenemos las siguientes expresiones. Cuando $m = 2$:

$$f'(g) = \frac{1}{(1 - g)^2} = \sum_{j=0}^{\infty} jg^{j-1} = \sum_{j=0}^{\infty} (1 + j)g^j$$

En el otro extremo, cuando $m = p$:

$$f^{(p-1)}(g) = \frac{p-1}{(1 - g)^p} = \sum_{j=0}^{\infty} \frac{(p-1+j)(p-2+j) \dots (2+j)(1+j)}{(p-1)!} g^j \quad (2.59)$$

Así, en el extremo cuando $m = p$ la solución general podría estar dada por:

$$Z_t = a_0 \sum_{j=0}^{\infty} \frac{(p-1+j)(p-2+j) \dots (2+j)(1+j)}{(p-1)!} g^j + g^t \sum_{i=0}^p s_i t^{i-1} \quad (2.60)$$

Donde $|g| < 1$, $t = \dots, -2, -1, 0, 1, 2, \dots$. Para finalizar esta sección, plantearemos la expresión del polinomio característico que nos permitirá hacer el análisis de convergencia de los procesos. Partamos de que la ecuación (2.56) se puede escribir como:

$$(x^{-1})^p - a_1(x^{-1})^{p-1} - a_2(x^{-1})^{p-2} - \dots - a_p = 0 \quad (2.61)$$

La ecuación (2.61) permite interpretar las raíces del polinomio característico de forma directa ya que $x_1^{-1} = g_1$, $x_2^{-1} = g_2$, ..., $x_p^{-1} = g_p$. Así, siempre que $p \geq 1$ en la ecuación (2.49), diremos que el proceso descrito en esa ecuación dará como resultado un proceso convergente si se cumplen las dos condiciones (2.62) y (2.63):

$$|a_p| < 1 \quad (2.62)$$

$$a_1 + a_2 + \dots + a_p < 1 \quad (2.63)$$

Alternativamente, cuando las raíces son reales, lo anterior es equivalente a la expresión (2.64):

$$|g_i| < 1 \quad (2.64)$$

Para $\forall i = 1, 2, \dots, p$. Cuando las raíces son imaginarias, las dos condiciones (2.62) y (2.63) son equivalentes a la expresión (2.65):

$$\sqrt{g_i g_j} = \sqrt{u^2 + v^2} < 1 \quad (2.65)$$

Para $\forall i \neq j$ y $i, j = 1, 2, \dots, p$. Cuando $g_1 = g_2 = \dots = g_p = g$, la condición de la ecuación (2.64) se resume a que $|g| < 1$. En resumen, las condiciones descritas en las ecuaciones (2.64) y (2.65) se pueden ilustrar mediante un círculo unitario como el de la Figura 2.9 en que si las raíces se ubican dentro de éste, podemos decir que el proceso es convergente en el largo plazo.

```
# Crear datos para el círculo
theta <- seq(0, 2*pi, length.out=100)
x <- cos(theta)
y <- sin(theta)
circle_data <- data.frame(x, y)

# Dibujar el círculo
ggplot(circle_data, aes(x=x, y=y)) +
  geom_polygon(fill=NA, color="black") +
  coord_fixed(ratio=1) +
  xlim(-1.5, 1.5) +
  ylim(-1.5, 1.5)
```

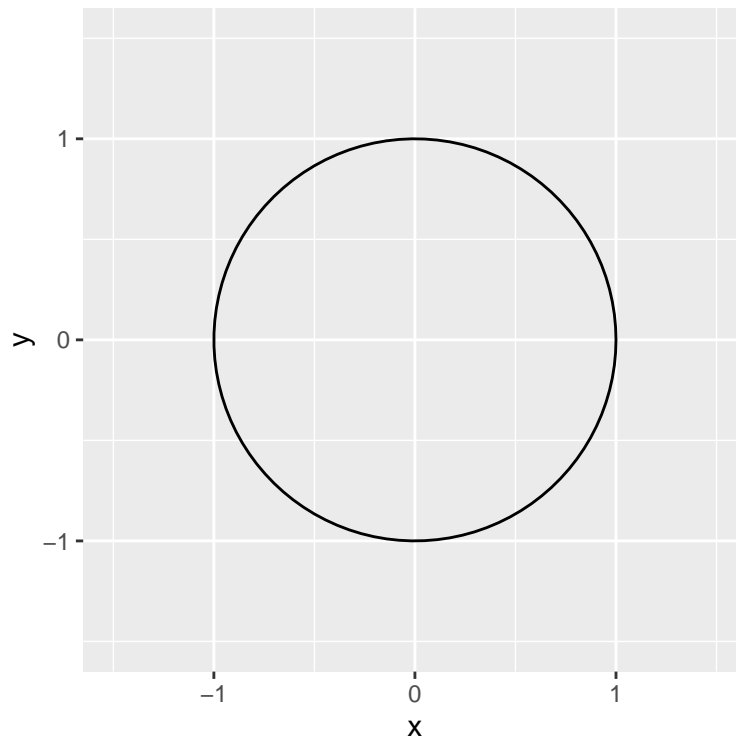


Figure 2.9: Circulo unitario en el que se cumple que $|g_i| < 1$ y $(g_i g_j)^{1/2} = (u^2 + v^2)^{1/2} < 1$

2.4 Operador de rezago L

Denotemos, como se ha mencionado con anterioridad, con L al operador de rezago, el cual nos permitirá construir una relación entre diferencias y medias móviles como se verá más adelante en los procesos univariados $AR(p)$, $MA(q)$ y, en general, $ARIMA(p, d, q)$. Sean X , Y o Z variables con las que denotaremos a una serie de tiempo (note que hasta el momento no hemos definido qué es una serie de tiempo, no obstante, no es necesario definirla para hacer uso del operador).

En esta sección resumiremos algunas propiedades usadas en el capítulo y en capítulos más adelante. Así, si a dicha serie le aplicamos el operador rezago antes definido, el resultado deberá ser que cada uno de los valores de la serie es retardado o regresado un período. Es decir:

$$LZ_t = Z_{t-1} \quad (2.66)$$

De esta forma, si aplicamos el operador rezago L a la nueva serie de tiempo dada por Z_{t-1} podemos obtener Z_{t-2} . Haciendo uso de la ecuación (2.66) podemos obtener:

$$LZ_{t-1} = L(LZ_t) = L^2Z_t = Z_{t-2} \quad (2.67)$$

Mediante una generalización podemos obtener:

$$L^k Z_t = Z_{t-k} \quad (2.68)$$

Para $k = \dots, -2, -1, 0, 1, 2, \dots$. Así, para $k = 0$ obtenemos la identidad dado que $L^0 Z_t = Z_t$, de tal forma que siempre asumiremos que $L^0 = 1$. En otro caso, cuando $k > 0$ a la serie de tiempo a la cual se le aplique el operador rezago L se le deberá aplicar un rezago de k periodos a cada uno de los elementos de la serie. Por el contrario, cuando $k < 0$ el operador rezago significa que se deberá adelantar $|k|$ veces a cada elemento de la serie. Por ejemplo, $L^{-3}Z_t = Z_{t+3}$.

Las reglas descritas en lo subsecuente se mantienen indistintamente cuando aplican para el caso de rezagar como para cuando se adelanta una serie. Como primera propiedad tomemos a la siguiente propiedad:

$$L^m Z_{t-n} = L^m (L^n Z_t) = L^{m+n} Z_t = Z_{t-(n+m)} \quad (2.69)$$

De lo anterior, podemos inferir el siguiente resultado:

$$\Delta Z_t = Z_t - Z_{t-1} = (1 - L)Z_t \quad (2.70)$$

En el caso de la diferencia de orden cuatro o cuarta diferencia se puede expresar como:

$$\Delta_4 Z_t = Z_t - Z_{t-4} = (1 - L^4)Z_t \quad (2.71)$$

Al respecto, vale la pena aclarar que en ocasiones se hará uso de una notación alternativa dada por: Δ^k o Δ_k , donde $k = 1, 2, 3, \dots$, indistintamente, ya que en ambos casos se referirá a una diferencia de orden k . Esta notación resulta de gran utilidad cuando se quiere comparar periodos equivalentes como, por ejemplo, el mismo trimestre pero de un año anterior. De forma similar, para el caso de logaritmos podemos escribir la ecuación (2.71) como:

$$\Delta^4 \ln(Z_t) = \Delta_4 \ln(Z_t) = \ln(Z_t) - \ln(Z_{t-4}) = (1 - L^4) \ln(Z_t) \quad (2.72)$$

Para el caso de una serie de tiempo que se le ha transformado mediante medias móviles, digamos de 4 periodos, podemos escribirla como:

$$Zs_t = \frac{1}{4}(Z_t + Z_{t-1} + Z_{t-2} + Z_{t-3}) = \frac{1}{4}(1 + L + L^2 + L^3)Z_t \quad (2.73)$$

Una generalización del anterior caso puede ser escrita como un polinomio de orden p con el operador rezago L dado como:

$$\begin{aligned} \alpha(L)Z_t &= (1 - \alpha_1 L - \alpha_2 L^2 - \dots - \alpha_p L^p)Z_t \\ &= Z_t - \alpha_1 Z_{t-1} - \alpha_2 Z_{t-2} - \dots - \alpha_p Z_{t-p} \end{aligned} \quad (2.74)$$

Donde α_i puede ser reemplazada por cualquier constante a_i , con $i = 1, 2, 3, \dots$, para escribir ecuaciones como las anteriores. Adicionalmente, podemos decir que la ecuación (2.74) es una generalización del caso de medias móviles, el cual admite una ponderación distinta para cada uno de los elementos rezagados.

Existe la posibilidad de operar más de un polinomio a la vez. Para múltiples polinomios (digamos, los polinomios $\alpha(L)$ y $\beta(L)$) podemos escribir el siguiente resultado:

$$\alpha(L)\beta(L) = \beta(L)\alpha(L) \quad (2.75)$$

Tales polinomios del operador rezago también son llamados *filtros lineales*. A manera de ejemplo, tomemos el siguiente caso de diferencias para una serie de Z_t :

$$\Delta Z_t = (1 - L)Z_t = Z_t - Z_{t-1} \quad (2.76)$$

y un proceso de medias móviles para la misma serie de Z_t :

$$Zs_t = \frac{1}{4}(1 + L + L^2 + L^3)Z_t = \frac{1}{4}(Z_t + Z_{t-1} + Z_{t-2} + Z_{t-3}) \quad (2.77)$$

De tal forma que el producto de ambos procesos se puede escribir como:

$$(1 - L) \times \frac{1}{4}(1 + L + L^2 + L^3)Z_t = \frac{1}{4}(1 - L^4)Z_t \quad (2.78)$$

Es decir, que el producto de dos polinomios, uno de diferencias y otro más de medias móviles, resulta en uno de diferencias pero de mayor grado, en este caso de grado 4.

Chapter 3

Procesos Estacionarios y Modelos Univariados

3.1 Definición de ergodicidad y estacionariedad

A partir de esta sección introduciremos mayor formalidad matemática al análisis de las series de tiempo. Por ello cambiaremos un poco la notación y ocuparemos a X_t en lugar de Z_t como objeto de nuestro análisis. Con X_t denotaremos a una serie de tiempo, ya que con Z_t denotaremos a una variable, sin que ella fuera necesariamente una serie de tiempo en los términos que a continuación discutimos. Asimismo, iniciaremos por establecer una serie de definiciones.

De esta forma, definiremos a una *serie de tiempo* como un vector de variables aleatorias de dimensión T , dado como:

$$X_1, X_2, X_3, \dots, X_T \quad (3.1)$$

Cada una de las X_t ($t = 1, 2, \dots, T$) consideradas como una variable aleatoria. Así, también podemos denotar a la serie de tiempo como:

$$\{X_t\}_{t=1}^T \quad (3.2)$$

Es decir, definiremos a una *serie de tiempo como una realización de un proceso estocástico* –o un Proceso Generador de Datos (PGD). Consideremos una muestra de los múltiples posibles resultados de muestras de tamaño T , la colección dada por:

$$\{X_1^{(1)}, X_2^{(1)}, \dots, X_T^{(1)}\} \quad (3.3)$$

Digamos que la ecuación (3.3) es una de las tantas posibles resultantes del proceso estocástico o PGD. Eventualmente podríamos estar dispuestos a observar

este proceso indefinidamente, de forma tal que estemos interesados en la secuencia dada por $\{X_t^{(1)}\}_{t=1}^{\infty}$, lo cual no dejaría de ser sólo una de las tantas realizaciones o secuencias del proceso estocástico original.

Tan solo por poner un ejemplo, podríamos observar las siguientes realizaciones del mismo PGD:

$$\begin{aligned} &\{X_1^{(2)}, X_2^{(2)}, \dots, X_T^{(2)}\} \\ &\{X_1^{(3)}, X_2^{(3)}, \dots, X_T^{(3)}\} \\ &\{X_1^{(4)}, X_2^{(4)}, \dots, X_T^{(4)}\} \\ &\vdots \\ &\{X_1^{(j)}, X_2^{(j)}, \dots, X_T^{(j)}\} \end{aligned}$$

Donde $j \in \mathbb{Z}$. En lo subsecuente, diremos que una serie de tiempo es una realización del proceso estocástico subyacente. Considerando, en consecuencia, al proceso estocástico con todas sus posibilidades de realización.

Para hacer más sencilla la notación no distinguiremos entre el proceso en sí mismo y una de sus realizaciones, es decir, siempre escribiremos a una serie de tiempo como la secuencia mostrada en la ecuación (3.2), o más precisamente como la siguiente realización:

$$\{X_1, X_2, \dots, X_T\} \quad (3.4)$$

O simplemente:

$$X_1, X_2, \dots, X_T \quad (3.5)$$

El proceso estocástico de dimensión T puede ser completamente descrito por su función de distribución multivariada de dimensión T . No obstante, esto no resulta ser práctico cuando se opere más adelante en el curso. Por ello, en el curso y, en general, en casi todos los textos lo hacen, sólo nos enfocaremos en sus primero y segundo momentos. Es decir, en sus media o valor esperado:

$$\mathbb{E}[X_t]$$

Para $t = 1, 2, \dots, T$; o:

$$\begin{bmatrix} \mathbb{E}[X_1] \\ \mathbb{E}[X_2] \\ \vdots \\ \mathbb{E}[X_T] \end{bmatrix}$$

o,

$$[\mathbb{E}[X_1], \mathbb{E}[X_2], \dots, \mathbb{E}[X_T]]$$

Y en su variación:

$$Var[X_t] = \mathbb{E}[(X_t - \mathbb{E}[X_t])^2]$$

Para $t = 1, 2, \dots, T$, y de sus $T(T-1)/2$ covarianzas:

$$\text{Cov}[X_t, X_s] = \mathbb{E}[(X_t - \mathbb{E}[X_t])(X_s - \mathbb{E}[X_s])]$$

Para $t < s$. Por lo tanto, en la forma matricial podemos escribir lo siguiente:

$$\begin{aligned} & \begin{bmatrix} \text{Var}[X_1] & \text{Cov}[X_1, X_2] & \cdots & \text{Cov}[X_1, X_T] \\ \text{Cov}[X_2, X_1] & \text{Var}[X_2] & \cdots & \text{Cov}[X_2, X_T] \\ \vdots & \vdots & \ddots & \vdots \\ \text{Cov}[X_T, X_1] & \text{Cov}[X_T, X_2] & \cdots & \text{Var}[X_T] \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \rho_{12} & \cdots & \rho_{1T} \\ \rho_{21} & \sigma_2^2 & \cdots & \rho_{2T} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{T1} & \rho_{T2} & \cdots & \sigma_T^2 \end{bmatrix} \end{aligned} \quad (3.6)$$

Donde es claro que en la matriz de la ecuación (3.6) existen $T(T-1)/2$ covarianzas distintas, ya que se cumple que $\text{Cov}[X_t, X_s] = \text{Cov}[X_s, X_t]$, para $t \neq s$.

A menudo, esas covarianzas son denominadas como autocovarianzas puesto que ellas son covarianzas entre variables aleatorias pertenecientes al mismo proceso estocástico pero en un momento t diferente. Si el proceso estocástico tiene una distribución normal multivariada, su función de distribución estará totalmente descrita por sus momentos de primer y segundo orden.

Ahora introduciremos el concepto de ergodicidad, el cual indica que los momentos muestrales, los cuales son calculados en la base de una serie de tiempo con un número finito de observaciones, en la medida que $T \rightarrow \infty$ sus correspondientes momentos muestrales, tienden a los verdaderos valores poblacionales, los cuales definiremos como μ , para la media, y σ_X^2 para la varianza.

Este concepto sólo es cierto si asumimos que, por ejemplo, el valor esperado y la varianza son como se dice a continuación para todo $t = 1, 2, \dots, T$:

$$\mathbb{E}[X_t] = \mu_t = \mu \quad (3.7)$$

$$(3.8)$$

$$\text{Var}[X_t] = \sigma_X^2 \quad (3.9)$$

Mas formalmente, se dice que el PGD o el proceso estocástico es ergódico en la media si:

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\left(\frac{1}{T} \sum_{t=1}^T (X_t - \mu) \right)^2 \right] = 0 \quad (3.10)$$

y ergódico en la varianza si:

$$\lim_{T \rightarrow \infty} \mathbb{E} \left[\left(\frac{1}{T} \sum_{t=1}^T (X_t - \mu)^2 - \sigma_X^2 \right)^2 \right] = 0 \quad (3.11)$$

Estas condiciones se les conoce como *propiedades de consistencia* para las variables aleatorias. Sin embargo, éstas no pueden ser probadas. Por ello se les denomina como un supuesto que pueden cumplir algunas de las series. Más importante aún: **un proceso estocástico que tiende a estar en equilibrio estadístico en un orden ergódico, es estacionario.**

Podemos distinguir dos tipos de estacionariedad. Si asumimos que la función común de distribución del proceso estocástico no cambia a lo largo del tiempo, se dice que el proceso es *estrictamente estacionario*. Como este concepto es difícil de aplicar en la práctica, solo consideraremos a la *estacionariedad débil* o estacionariedad en sus momentos.

Definiremos a la estacionariedad por sus momentos del correspondiente proceso estocástico dado por $\{X_t\}$:

1. *Estacionariedad en media:* Un proceso estocástico es estacionario en media si $E[X_t] = \mu_t = \mu$ es constante para todo t .
2. *Estacionariedad en varianza:* Un proceso estocástico es estacionario en varianza si $Var[X_t] = E[(X_t - \mu_t)^2] = \sigma_X^2 = \gamma(0)$ es constante y finita para todo t .
3. *Estacionariedad en covarianza:* Un proceso estocástico es estacionario en covarianza si $Cov[X_t, X_s] = E[(X_t - \mu_t)(X_s - \mu_s)] = \gamma(|s - t|)$ es sólo una función del tiempo y de la distancia entre las dos variables aleatorias. Por lo que no depende del tiempo denotado por t (no depende de la información contemporánea).
4. *Estacionariedad débil:* Como la estacionariedad en varianza resulta de forma inmediata de la estacionariedad en covarianza cuando se asume que $s = t$, un proceso estocástico es débilmente estacionario cuando es estacionario en media y covarianza.

Puesto que resulta poco factible asumir una estacionariedad diferente a la débil, es adelante siempre que digamos que un proceso es estacionario se referirá al caso débil y sólo diremos que el proceso es estacionario, sin el apelativo de débil.

Ejemplo. Supongamos una serie de tiempo denotada por: $\{U_t\}_{t=0}^T$. Decimos que el proceso estocástico $\{U_t\}$ es un *proceso estocástico puramente aleatorio* o es un *proceso estocástico de ruido blanco o caminata aleatoria*, si este tiene las siguientes propiedades:

1. $E[U_t] = 0, \forall t$;
2. $Var[U_t] = E[(U_t - \mu_t)^2] = E[(U_t - \mu)^2] = E[(U_t)^2] = \sigma^2, \forall t$, y
3. $Cov[U_t, U_s] = E[(U_t - \mu_t)(U_s - \mu_s)] = E[(U_t - \mu)(U_s - \mu)] = E[U_t U_s] = 0, \forall t \neq s$.

En otras palabras, un proceso U_t es un ruido blanco si su valor esperado (promedio) es cero (0), tiene una varianza finita y constante, y además no le importa la historia pasada. Así, su valor presente no se ve influenciado por sus valores pasados, no importando respecto de qué período se tome referencia.

En apariencia, por sus propiedades, este proceso es débilmente estacionario –o simplemente, estacionario–. Todas las variables aleatorias tienen una media de cero, una varianza σ^2 y no existe correlación entre ellas.

Ahora, supongamos que definimos un nuevo proceso estocástico $\{X_t\}$ como:

$$X_t = \begin{cases} U_0 & \text{para } t = 0 \\ X_{t-1} + U_t & \text{para } t = 1, 2, 3, \dots \end{cases} \quad (3.12)$$

Donde $\{U_t\}$ es un proceso puramente aleatorio. Este proceso estocástico, o caminata aleatoria sin tendencia (ajuste - drift), puede ser reescrito como:

$$X_t = \sum_{j=0}^t U_j \quad (3.13)$$

Tratemos de dar más claridad al ejemplo, para ello asumamos que generamos a $\{U_t\}$ por medio del lanzamiento de una moneda. Donde obtenemos una cara (águila) con una probabilidad de 0.5, en cuyo caso decimos que la variable aleatoria U_t tomará el valor de +1, y una cruz (sol) con una probabilidad de 0.5, en cuyo caso decimos que la variable aleatoria U_t toma el valor de -1.

Este planteamiento cumple con las propiedades enunciadas ya que:

1. $\mathbb{E}[U_t] = 0.5 \times -1 + 0.5 \times 1 = 0, \forall t$
2. $Var[U_t] = \mathbb{E}[(U_t - 0)^2] = \frac{1}{2}((-1)^2) + \frac{1}{2}((1)^2) = 1, \forall t$
3. $Cov[U_t, U_s] = \mathbb{E}[(U_t - 0)(U_s - 0)] = \mathbb{E}[U_t \cdot U_s] = 0, \forall t \neq s.$

Retomando a nuestro proceso X_t , diremos que el caso de $X_0 = 0$, para $t = 0$. Si verificamos cuáles son sus primeros y segundos momentos de $\{X_t\}$, tenemos:

$$\mathbb{E}[X_t] = \mathbb{E}\left[\sum_{j=1}^t U_j\right] = \sum_{j=1}^t \mathbb{E}[U_j] = 0 \quad (3.14)$$

En cuanto a la varianza:

$$\begin{aligned}
 Var[X_t] &= Var\left[\sum_{j=1}^t U_j\right] \\
 &= \sum_{j=1}^t Var[U_j] + 2 * \sum_{j \neq k} Cov[U_j, U_k] \\
 &= \sum_{j=1}^t 1 \\
 &= t
 \end{aligned} \tag{3.15}$$

Lo anterior, dado que hemos supuesto que en la caminata aleatoria todas las variables aleatorias son independientes, es decir, $Cov[U_t, U_s] = E[U_t \cdot U_s] = 0$. Por su parte, la covarianza del proceso estocástico se puede ver como:

$$\begin{aligned}
 Cov[X_t, X_s] &= \mathbb{E}\left[\left(\sum_{j=1}^t U_j - 0\right)\left(\sum_{i=1}^s U_i - 0\right)\right] \\
 &= \mathbb{E}[(U_1 + U_2 + \dots + U_t)(U_1 + U_2 + \dots + U_s)] \\
 &= \sum_{j=1}^t \sum_{i=1}^s \mathbb{E}[U_j U_i] \\
 &= \mathbb{E}[U_1^2] + \mathbb{E}[U_2^2] + \dots + \mathbb{E}[U_k^2] \\
 &= \sigma^2 + \sigma^2 + \dots + \sigma^2 \\
 &= 1 + 1 + 1 + 1 \\
 &= \min(t, s)
 \end{aligned}$$

Así, el proceso estocástico dado por la caminata aleatoria sin un término de ajuste es estacionario en media, pero no en varianza o en covarianza, y consecuentemente, en general no estacionario, condición que contraria al caso del proceso simple descrito en U_t .

Es fácil ver que muchas de las posibilidades de realización de este proceso estocástico (series de tiempo) pueden tomar cualquiera de las rutas consideradas en la Figura 3.1.

```

set.seed(1234)
# Utilizaremos una función guardada en un archivo a parte
# Llamamos a la función:
source("Caminata.R")

# Definimos argumentos de la función
Opciones <- c(-1, 1)
#

```

```

Soporte <- 10000

# Vamos a réplicar el proceso con estos parámetros
Rango <- 200
#
Caminos <- 10

#
for(i in 1:Caminos){
  TT <- data.matrix(data.frame(Caminata(Opciones, Soporte)[1]))
  #
  G_t <- data.matrix(data.frame(Caminata(Opciones, Soporte)[2]))
  #
  plot(TT, G_t, col = "blue", type = "l", ylab = "Ganancias",
        xlab = "Tiempo", ylim = c(-Rango,Rango))
  #
  par(new = TRUE)
  #
  i <- i +1
}
#
par(new = FALSE)

```

3.2 Función de autocorrelación

Para ampliar la discusión, es posible calcular la fuerza o intensidad de la dependencia de las variables aleatorias dentro de un proceso estocástico, ello mediante el uso de las autocovarianzas. Cuando las covarianzas son normalizadas respecto de la varianza, el resultado es un término que es independiente de las unidades de medida aplicadas, y se conoce como la *función de autocorrelación*.

Para procesos estacionarios, dicha función de autocorrelación esta dada por:

$$\rho(\tau) = \frac{\mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)]}{\mathbb{E}[(X_t - \mu)^2]} = \frac{\gamma(\tau)}{\gamma(0)} \quad (3.16)$$

Donde $\tau = \dots, -2, -1, 0, 1, 2, \dots$. Dicha función tiene las siguientes propiedades:

1. $\rho(0) = 1$. Es fácil demostrar que la función $\rho(0)$ es:

$$\rho(0) = \frac{\mathbb{E}[(X_t - \mu)(X_{t+0} - \mu)]}{\mathbb{E}[(X_t - \mu)^2]} = \frac{\mathbb{E}[(X_t - \mu)^2]}{\mathbb{E}[(X_t - \mu)^2]} = 1 \quad (3.17)$$

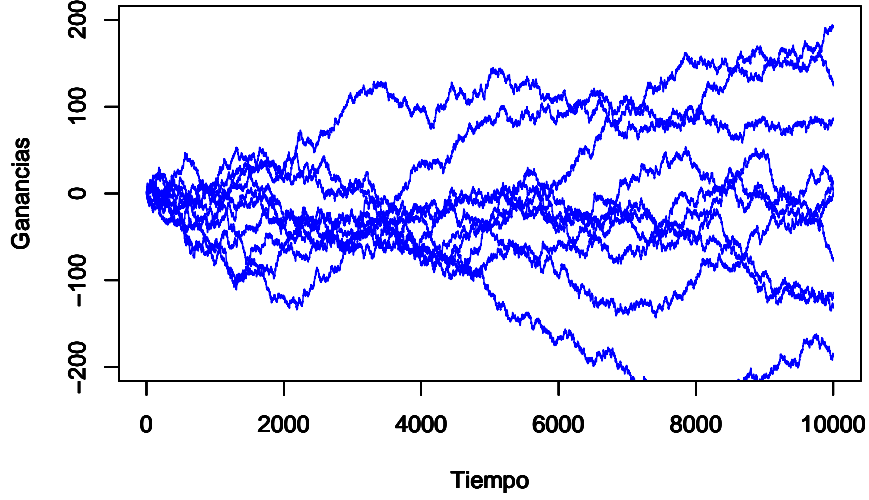


Figure 3.1: Ejemplo de 10 trayectorias de la caminata aleatoria, cuando sólo es posible cambios de +1 y -1

2. $\rho(\tau) = \rho(-\tau)$. Partiendo de la definición de $\rho(\tau)$ podemos ver que la distancia que existe entre t y $t+\tau$ es τ , de esta forma la autocorrelación de la variable X entre los periodos antes señalados debería ser la misma para el caso en que $\rho(-\tau)$. Partamos de la ecuación para ver más claramente:

$$\rho(\tau) = \frac{\mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)]}{\mathbb{E}[(X_t - \mu)^2]} = \frac{\mathbb{E}[(X_t - \mu)(X_{t-\tau} - \mu)]}{\mathbb{E}[(X_t - \mu)^2]} = \rho(-\tau) \quad (3.18)$$

3. $|\rho(\tau)| \leq 1$, para todo τ .

Derivado de las propiedades 1 y 2 antes descritas se puede concluir que sólo es necesario conocer la función de autocorrelación para el caso de $\tau = 1, 2, 3, \dots$, ya que de estos casos podemos derivar los valores de la función de autocorrelación complementarios de $\tau = \dots, -3, -2, -1$.

Partiendo de los supuestos de ergodicidad en relación a la media, varianza y covarianzas de un proceso estacionario, podemos estimar dichos parámetros con las siguientes formulaciones o propuestas de estimadores puntuales:

$$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T X_t \quad (3.19)$$

$$\hat{\gamma}(0) = \frac{1}{T} \sum_{t=1}^T (X_t - \hat{\mu})^2 = \hat{\sigma}^2 \quad (3.20)$$

$$\hat{\gamma}(\tau) = \frac{1}{T} \sum_{t=1}^{T-\tau} (X_t - \hat{\mu})(X_{t+\tau} - \hat{\mu}), \text{ para } \tau = 1, 2, \dots, T-1 \quad (3.21)$$

No hacemos la demostración en estas notas –sería deseable que el alumno revisara la siguiente afirmación–, pero estos últimos son estimadores consistentes de μ , $\gamma(0)$ y $\gamma(\tau)$. Por su parte, un estimador consistente de la función de autocorrelación estará dado por:

$$\hat{\rho}(\tau) = \frac{\sum_{t=1}^{T-\tau} (X_t - \hat{\mu})(X_{t+\tau} - \hat{\mu})}{\sum_{t=1}^T (X_t - \hat{\mu})^2} = \frac{\hat{\gamma}(\tau)}{\hat{\gamma}(0)} \quad (3.22)$$

El estimador de la ecuación (3.22) es asintóticamente insesgado. Por ejemplo, para el caso de un proceso de ruido blanco o caminata aleatoria, su varianza puede ser aproximada por el valor dado $1/T$. Ésta tiene, asintóticamente, una distribución normal. Dado esto, el intervalo de confianza al 95% será el dado por $\pm 2/\sqrt{T}$, en el cual se encuentra la mayoría de los coeficientes de autocorrelación estimados.

Ahora discutamos algunos ejemplos o aplicaciones. Cuando se realiza la evaluación de la estimación de un modelo de series de tiempo es importante saber si los residuales del modelo realmente tienen propiedades de un proceso puramente aleatorio, en particular, si ellos no están correlacionados entre sí. Así, la hipótesis a probar será:

$$H_0 : \rho(\tau) = 0, \text{ para todo } \tau = 1, 2, \dots, m \text{ y } m < T \quad (3.23)$$

Esta expresión se puede interpretar como una prueba respecto de si la correlación entre la información de periodos atrás es cero con la información contemporánea. Para hacer una prueba global de la hipótesis de si un número m de coeficientes de autocovarianzas son cero Box y Pierce (1970) desarrollarán la siguiente estadística:

$$Q^* = T \sum_{j=1}^m \hat{\rho}(j)^2 \quad (3.24)$$

Bajo la hipótesis nula esta estadística se distribuye asintóticamente como una chi cuadrado (χ^2) con $m-k$ grados de libertad y con k que representa al número de parámetros estimados.

Haciendo una aplicación estricta de la distribución de esta estadística, sabemos que esta se mantiene asintóticamente. Greta, Ljung y Box (1978) propusieron

la siguiente modificación de la estadística para muestras pequeñas:

$$Q = T(T+2) \sum_{j=1}^m \frac{\hat{\rho}(j)^2}{T-j} \quad (3.25)$$

La cual también se distribuye asintóticamente como χ^2 con $m - k$ grados de libertad.

También es intuitivamente claro que la hipótesis nula de no autocorrelación de residuales debería ser rechazada si alguno de los valores $\hat{\rho}(j)$ es muy grande, es decir, si Q o Q^* es muy grande. O más precisamente, si estas estadísticas son más grandes que los correspondientes valores críticos de la distribución χ^2 con $m - k$ grados de libertad a algún grado dado de significancia.

Una alternativa para esta prueba es una del tipo Multiplicadores de Lagrange (o LM) desarrollada por Breusch (1978) y Godfrey (1978). La cual, al igual que las estadísticas Q y Q^* , la hipótesis nula está dada por:

H_0 : Los residuales no están autocorrelacionados.

H_a : Los residuales muestran alguna autocorrelación de forma autoregresiva o de medias móviles.

La prueba consiste en realizar una regresión auxiliar en la cual los residuales se estiman en función de las variables explicativas del modelo original y en los residuales mismos pero rezagados hasta el término m (regresión auxiliar). La prueba resulta en una estadística con una distribución χ^2 con m grados de libertad la cual está dada por la expresión:

$$LM = T \times R^2 \quad (3.26)$$

Donde R^2 es el resultante de la regresión auxiliar y T es el número de observaciones totales.

En comparación con una prueba Durbin - Watson que es comúnmente usada en la econometría tradicional, para probar autocorrelación de los residuales, las estadísticas Q , Q^* y LM tienen las siguientes ventajas:

1. Permiten corroborar la existencia de autocorrelación para cualquier orden, y no solo para un primer orden (es decir, para cualquier valor de $\tau = 1, 2, 3, \dots$);
2. Los resultados se mantienen aún y cuando exista una probable variable endógena en forma rezagada, y
3. No depende del orden o la forma en que se acomoden las observaciones, algo que es muy probable que ocurra en la econometría tradicional.

El hecho de que los residuales no estén autocorrelacionados no implica que estos sean independientes y normalmente distribuidos. La ausencia de autocorrelación no implica una independencia estocástica si las variables son normalmente distribuidas.

A menudo se asume que estos residuales están distribuidos normalmente, ya que la mayoría de las pruebas estadísticas tienen este supuesto detrás. No obstante, ello también depende de los otros momentos de la distribución, específicamente del tercer y cuarto momento. Los cuales se expresan como:

$$\mathbb{E}[(X_t - \mathbb{E}[X_t])^i], i = 3, 4$$

El tercer momento es necesario para determinar el sesgo, el cual está dado como:

$$\hat{S} = \frac{1}{T} \frac{\sum_{t=1}^T (X_t - \hat{\mu})^3}{\sqrt{\hat{\gamma}(0)^3}} \quad (3.27)$$

Para distribuciones simétricas (como en el caso de la distribución normal) el valor teórico para el sesgo es cero.

La curtosis, la cual está dada en función del cuarto momento, se puede expresar como:

$$\hat{K} = \frac{1}{T} \frac{\sum_{t=1}^T (X_t - \hat{\mu})^4}{\hat{\gamma}(0)^2} \quad (3.28)$$

Para el caso de una distribución normal, esta estadística toma el valor de 3. Valores más grandes que 3 indican que la distribución tiene colas anchas. En tales casos se ubican los datos financieros.

Usando el valor de las estadísticas para medir el sesgo y la curtosis, S y K , respectivamente, Jarque y Bera (1980) propusieron una prueba de normalidad, la cual puede ser aplicada a series de tiempo en niveles o en diferencias indistintamente. Dicha prueba se expresa como:

$$JB = \frac{T}{6} \left(\hat{S} + \frac{1}{4}(\hat{K} - 3)^2 \right) \quad (3.29)$$

La cual tiene una distribución χ^2 con 2 grados de libertad y donde T es el tamaño de la muestra. La hipótesis de que las observaciones están distribuidas de forma normal se rechaza si el valor de la estadística de prueba es más grande que los correspondientes valores críticos en tablas.

Ejemplo. Veamos un ejemplo para ilustrar el uso de la función de autocorrelación. Tomemos como variable al número de pasajeros transportados por el sistema de transporte del metro de la CDMX.¹ Los datos empleados fueron

¹Los datos y algoritmo está disponible en el repositorio de GitHub y corresponde a la Clase 3.

tomados del INEGI y son una serie de tiempo en el período que va de enero de 2000 a mayo de 2023, es decir, 281 observaciones. Como se puede apreciar en la Figura 3.2, el número de pasajeros por mes ha oscilado significativamente a lo largo del tiempo. Incluso podemos observar un cambio estructural de la serie entre 2011 y 2012. Asimismo, podemos ubicar una caída atípica que ocurrió en septiembre de 2017. Pero lo más relevante es la caída asociada a la pandemia de COVID-19 de 2020.

```
library(ggplot2)
library(dplyr)
library(readxl)

Datos <- read_excel("BD/Base_Transporte.xlsx",
                    sheet = "Datos", col_names = TRUE)

ggplot(data = Datos, aes(x = Periodo, y = Pax_Metro)) +
  geom_line(linewidth = 0.5, color = "darkblue") +
  #geom_point(size = 1.0, color = "darkblue") +
  #theme_bw() +
  xlab("Tiempo") +
  ylab("Millones de pasajeros") +
  theme(plot.title = element_text(size = 11, face = "bold", hjust = 0)) +
  theme(plot.subtitle = element_text(size = 10, hjust = 0)) +
  theme(plot.caption = element_text(size = 10, hjust = 0)) +
  theme(plot.margin = unit(c(1,1,1,1), "cm")) +
  labs(
    title = "Pasajeros Transportados en el Metro de la CDMX",
    subtitle = "(Ene-2000 a Jul-2021)",
    caption = "Fuente: Elaboración propia con información del INEGI, \nhttps://www.inegi.org.mx/contenidos/resultados/indicadores/transporte/indicador-1000000-pasajeros-transportados-en-el-metro-de-la-cdmx.aspx"
  )

#
ggsave("Pax_Metro.png", width = 20, height = 15, units = "cm")
```

A esta serie de tiempo le calculamos los principales estadísticos hasta ahora estudiados y obtenemos el Cuadro 3.1. En dicho cuadro se destaca que se muestra la función de autocorrelación para los tres primeros rezagos. Para mayor detalle, en la Figura 3.3 se muestra la función de autocorrelación, en donde las bandas descritas por las líneas azules son el intervalo de confianza dentro de las cuales no se puede rechazar la hipótesis nula de que $H_0 : \hat{\rho}(p) = 0$, para todo $\tau = 1, 2, \dots, T - 1$.

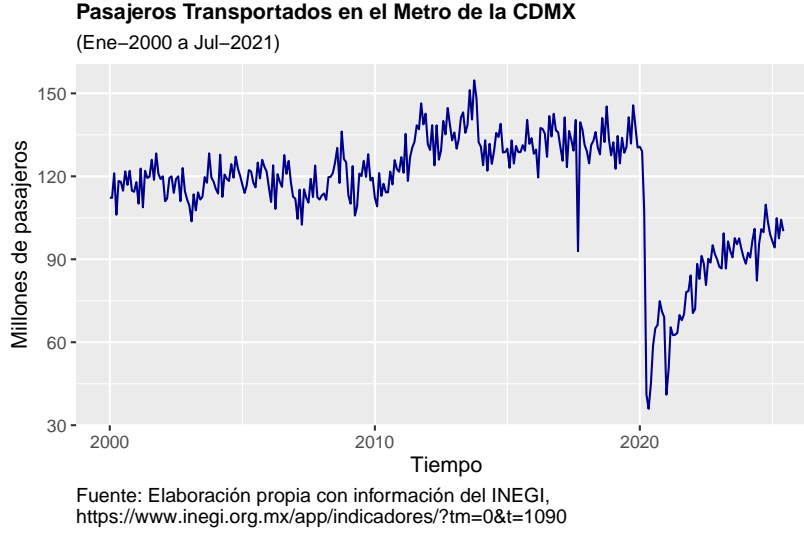


Figure 3.2: Evolución del número de pasajeros en el Metro de la CDMX, enero 2000 a mayo 2023

Table 3.1: Estadísticas descriptivas del número de pasajeros en el Metro de la CDMX, enero de 2000 a junio de 2019

Estadística	Valor
$\hat{\mu} = \frac{1}{T} \sum_{t=1}^T X_t$	124.3000
$\hat{\gamma}(0) = \frac{1}{T} \sum_{t=1}^T (X_t - \hat{\mu})^2$	103.6400
$\hat{\gamma}(1) = \frac{1}{T} \sum_{t=1}^{T-1} (X_t - \hat{\mu})(X_{t+1} - \hat{\mu})$	63.1100
$\hat{\gamma}(2) = \frac{1}{T} \sum_{t=1}^{T-2} (X_t - \hat{\mu})(X_{t+2} - \hat{\mu})$	72.9100
$\hat{\gamma}(3) = \frac{1}{T} \sum_{t=1}^{T-3} (X_t - \hat{\mu})(X_{t+3} - \hat{\mu})$	63.6900
$\hat{\rho}(1) = \frac{\sum_{t=1}^{T-1} (X_t - \hat{\mu})(X_{t+1} - \hat{\mu})}{\sum_{t=1}^T (X_t - \hat{\mu})^2} = \frac{\hat{\gamma}(1)}{\hat{\gamma}(0)}$	0.6089
$\hat{\rho}(2) = \frac{\sum_{t=1}^{T-2} (X_t - \hat{\mu})(X_{t+2} - \hat{\mu})}{\sum_{t=1}^T (X_t - \hat{\mu})^2} = \frac{\hat{\gamma}(2)}{\hat{\gamma}(0)}$	0.7035
$\hat{\rho}(3) = \frac{\sum_{t=1}^{T-3} (X_t - \hat{\mu})(X_{t+3} - \hat{\mu})}{\sum_{t=1}^T (X_t - \hat{\mu})^2} = \frac{\hat{\gamma}(3)}{\hat{\gamma}(0)}$	0.6145
$Q^* = T \sum_{j=1}^1 \hat{\rho}(j)^2$	86.7577
$Q^* = T \sum_{j=1}^2 \hat{\rho}(j)^2$	290.9279

```
Pax_Metro <- ts(Datos$Pax_Metro,
               start = 2000,
               freq = 12)

acf(Pax_Metro,
    lag.max = 150,
    xlab = 'Resagos k en meses',
    main = "Funcion de Autocorrelación del número de pasajeros del metro")
```

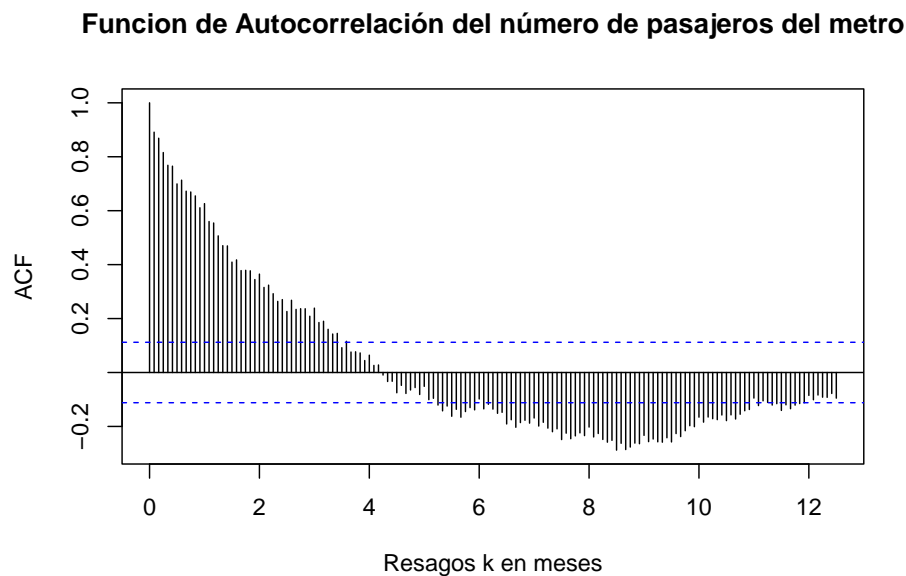


Figure 3.3: Función de Autocorrelación: 150 rezagos del número de pasajeros en el Metro de la CDMX, enero de 2000 a mayo de 2023

3.3 Procesos estacionarios univariados

En este capítulo analizaremos el método o metodología de análisis de series de tiempo propuesto por Box y Jenkins (1970). Los modelos propuestos dentro de esta metodología o conjunto de métodos se han vuelto indispensables para efectos de realizar pronósticos de corto plazo.

En este sentido, se analizarán los métodos más importantes en series de tiempo: procesos autoregresivos (AR, por sus siglas en inglés) y procesos de medias móviles (MA, por sus siglas en inglés). Asimismo, se realizará un análisis de los

procesos que resultan de la combinación de ambos, conocida como ARMA, los cuales son más comúnmente usados para realizar pronósticos.

3.4 Procesos Autoregresivos (AR)

Los procesos autoregresivos tienen su origen en el trabajo de Cochrane y Orcutt de 1949, mediante el cual analizaron los residuales de una regresión clásica como un proceso autoregresivo. Puede consultarse el apéndice para la discusión del modelo de regresión clásica.

3.4.1 AR(1)

Como primer caso analizaremos al proceso autoregresivo de primer orden, $AR(1)$, el cual podemos definir como una Ecuación Lineal en Diferencia de Primer Orden Estocástica. Diremos que una Ecuación Lineal en Diferencia de Primer Orden es estocástica si en su representación analítica considera un componente estocástico como en la ecuación (3.30) descrita a continuación:

$$X_t = a_0 + a_1 X_{t-1} + U_t \quad (3.30)$$

Donde a_0 es un término constante, U_t es un proceso estacionario, con media cero (0), una varianza finita y constante (σ^2) y una covarianza que depende de la distancia entre t y cualquier $t-s$ (γ_s)—que no depende de los valores pasados o futuros de la variable—, X_0 es el valor inicial del proceso X_t . No obstante, en ocasiones vamos a asumir que la covarianza será cero (0), por lo que en esos casos tendremos un proceso puramente aleatorio. Considerando la ecuación (3.30) y un proceso de sustitución sucesivo podemos establecer lo siguiente, empezando con X_1 :

$$X_1 = a_0 + a_1 X_0 + U_1$$

Para X_2 :

$$\begin{aligned} X_2 &= a_0 + a_1 X_1 + U_2 \\ &= a_0 + a_1(a_0 + a_1 X_0 + U_1) + U_2 \\ &= a_0 + a_1 a_0 + a_1^2 X_0 + a_1 U_1 + U_2 \end{aligned}$$

Para X_3 :

$$\begin{aligned} X_3 &= a_0 + a_1 X_2 + U_3 \\ &= a_0 + a_1(a_0 + a_1 a_0 + a_1^2 X_0 + a_1 U_1 + U_2) + U_3 \\ &= a_0 + a_1 a_0 + a_1^2 a_0 + a_1^3 X_0 + a_1^2 U_1 + a_1 U_2 + U_3 \end{aligned}$$

Así, para cualquier X_t , $t = 1, 2, 3, \dots$, obtendríamos:

$$\begin{aligned}
X_t &= a_0 + a_1 X_{t-1} + U_t \\
&= a_0 + a_1(a_0 + a_1 a_0 + a_1^2 a_0 + \dots + a_1^{t-2} a_0 + a_1^{t-1} X_0 \\
&\quad + a_1^{t-2} U_1 + \dots + a_1 U_{t-2} + U_{t-1}) + U_t \\
&= a_0 + a_1 a_0 + a_1^2 a_0 + a_1^3 a_0 + \dots + a_1^{t-1} a_0 + a_1^t X_0 \\
&\quad + a_1^{t-1} U_1 + \dots + a_1^2 U_{t-2} + a_1 U_{t-1} + U_t \\
&= (1 + a_1 + a_1^2 + a_1^3 + \dots + a_1^{t-1}) a_0 + a_1^t X_0 \\
&\quad + a_1^{t-1} U_1 + \dots + a_1^2 U_{t-2} + a_1 U_{t-1} + U_t \\
&= \frac{1 - a_1^t}{1 - a_1} a_0 + a_1^t X_0 + \sum_{j=0}^{t-1} a_1^j U_{t-j}
\end{aligned} \tag{3.31}$$

De esta forma en la ecuación (3.31) observamos un proceso que es explicado por dos partes: una que depende del tiempo y otra que depende de un proceso estocástico. Asimismo, debe notarse que la condición de convergencia es idéntica al caso de ecuaciones en diferencia estudiadas al inicio del curso: $|a_1| < 1$, por lo que cuando $t \rightarrow \infty$, la expresión (3.31) será la siguiente:

$$X_t = a_0 \frac{1}{1 - a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-j} \tag{3.32}$$

Así, desaparece la parte dependiente del tiempo y únicamente prevalece la parte que es dependiente del proceso estocástico. Esta es la solución de largo plazo del proceso $AR(1)$, la cual depende del proceso estocástico. Notemos, además, que esta solución implica que la variable o la serie de tiempo X_t es también un proceso estocástico que hereda las propiedades de U_t . Así, X_t es también un proceso estocástico estacionario, como demostraremos más adelante.

Observemos que la ecuación (3.32) se puede reescribir si consideramos la formulación que en la literatura se denomina como la descomposición de Wold, en la cual se define que es posible asumir que $\psi_j = a_1^j$ y se considera el caso en el cual $|a_1| < 1$. De esta forma tendremos que, por ejemplo, cuando:

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{\infty} a_1^{2j} = \frac{1}{1 - a_1^2}$$

Alternativamente y de forma similar a las ecuaciones en diferencia estudiadas previamente, podemos escribir el proceso $AR(1)$ mediante el uso del operador rezago como:

$$\begin{aligned}
X_t &= a_0 + a_1 L X_t + U_t \\
X_t - a_1 L X_t &= a_0 + U_t \\
(1 - a_1 L) X_t &= a_0 + U_t \\
X_t &= \frac{a_0}{1 - a_1 L} + \frac{1}{1 - a_1 L} U_t
\end{aligned} \tag{3.33}$$

En esta última ecuación retomamos el siguiente término para reescribirlo como:

$$\frac{1}{1 - a_1 L} = 1 + a_1 L + a_1^2 L^2 + a_1^3 L^3 + \dots \quad (3.34)$$

Tomando este resultado para sustituirlo en la ecuación (3.33), obtenemos la siguiente expresión:

$$\begin{aligned} X_t &= (1 + a_1 L + a_1^2 L^2 + a_1^3 L^3 + \dots) a_0 + (1 + a_1 L + a_1^2 L^2 + a_1^3 L^3 + \dots) U_t \\ &= (1 + a_1 + a_1^2 + a_1^3 + \dots) a_0 + U_t + a_1 U_{t-1} + a_1^2 U_{t-2} + a_1^3 U_{t-3} + \dots \\ &= \frac{a_0}{1 - a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-j} \end{aligned} \quad (3.35)$$

Donde la condición de convergencia y estabilidad del proceso descrito en esta ecuación es que $|a_1| < 1$. Por lo que hemos demostrado que, mediante el uso del operador de rezago, es posible llegar al mismo resultado que obtuvimos mediante el procedimiento de sustituciones iterativas.

La ecuación (3.35) se puede interpretar como sigue. La solución o trayectoria de equilibrio de un AR(1) se divide en dos partes. La primera es una constante que depende de los valores de a_0 y a_1 . La segunda parte es la suma ponderada de las desviaciones o errores observados y acumulados en el tiempo hasta el momento t .

Ahora obtendremos los momentos que describen a la serie de tiempo cuando se trata de un proceso AR(1). Para ello debemos obtener la media, la varianza y las covarianzas de X_t . Para los siguientes resultados debemos recordar y tener en mente que si U_t es un proceso puramente aleatorio, entonces:

1. $\mathbb{E}[U_t] = 0$ para todo t
2. $Var[U_t] = \sigma^2$ para todo t
3. $Cov[U_t, U_s] = 0$ para todo $t \neq s$

Dicho lo anterior y partiendo de la ecuación (3.35), el primer momento o valor esperado de la serie de tiempo será el siguiente:

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E} \left[\frac{a_0}{1 - a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-j} \right] \\ &= \frac{a_0}{1 - a_1} + \sum_{j=0}^{\infty} a_1^j \mathbb{E}[U_{t-j}] \\ &= \frac{a_0}{1 - a_1} = \mu \end{aligned} \quad (3.36)$$

Respecto de la varianza podemos escribir la siguiente expresión a partir de la ecuación (3.35):

$$\begin{aligned}
Var[X_t] &= \mathbb{E}[(X_t - \mu)^2] \\
&= \mathbb{E}\left[\left(\frac{a_0}{1-a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-j} - \frac{a_0}{1-a_1}\right)^2\right] \\
&= \mathbb{E}[(U_t + a_1 U_{t-1} + a_1^2 U_{t-2} + a_1^3 U_{t-3} + \dots)^2] \\
&= \mathbb{E}[U_t^2 + a_1^2 U_{t-1}^2 + a_1^4 U_{t-2}^2 + a_1^6 U_{t-3}^2 + \dots \\
&\quad + 2a_1 U_t U_{t-1} + 2a_1^2 U_t U_{t-2} + \dots] \\
&= \mathbb{E}[U_t^2] + a_1^2 \mathbb{E}[U_{t-1}^2] + a_1^4 \mathbb{E}[U_{t-2}^2] + a_1^6 \mathbb{E}[U_{t-3}^2] + \dots \\
&= \sigma^2 + a_1^2 \sigma^2 + a_1^4 \sigma^2 + a_1^6 \sigma^2 + \dots \\
&= \sigma^2 (1 + a_1^2 + a_1^4 + a_1^6 + \dots) \\
&= \sigma^2 \frac{1}{1-a_1^2} = \gamma(0)
\end{aligned} \tag{3.37}$$

Previo a analizar la covarianza de la serie de tiempo, recordemos que para el proceso puramente aleatorio U_t su varianza y covarianza puede verse como $\mathbb{E}[U_t, U_s] = \sigma^2$, para $t = s$, y $\mathbb{E}[U_t, U_s] = 0$, para cualquier otro caso, respectivamente.

Dicho lo anterior, partiendo de la ecuación (3.35) la covarianza de la serie estará dada por:

$$\begin{aligned}
Cov(X_t, X_{t-\tau}) &= \mathbb{E}[(X_t - \mu)(X_{t-\tau} - \mu)] \\
&= \mathbb{E}\left[\left(\frac{a_0}{1-a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-j} - \frac{a_0}{1-a_1}\right) \right. \\
&\quad \left. \times \left(\frac{a_0}{1-a_1} + \sum_{j=0}^{\infty} a_1^j U_{t-\tau-j} - \frac{a_0}{1-a_1}\right)\right] \\
&= a_1^\tau \mathbb{E}[U_{t-\tau}^2 + a_1 U_{t-\tau-1}^2 + a_1^2 U_{t-\tau-2}^2 + a_1^3 U_{t-\tau-3}^2 + \dots] \\
&= a_1^\tau \sigma^2 \frac{1}{1-a_1^2} = \gamma(\tau)
\end{aligned} \tag{3.38}$$

Notése que con estos resultados en las ecuaciones (3.37) y (3.38) podemos construir la función de autocorrelación teórica como sigue:

$$\begin{aligned}
\rho(\tau) &= \frac{\gamma(\tau)}{\gamma(0)} \\
&= a_1^\tau
\end{aligned} \tag{3.39}$$

Donde $\tau = 1, 2, 3, \dots$ y $|a_1| < 1$. Este último resultado significa que cuando el proceso autoregresivo es de orden 1 (es decir, AR(1)) la función de autocorrelación teóricamente es igual al parámetro a_1 elevado al número de rezagos

considerados. No obstante, note que esto no significa que la autocorrelación observada sea como lo expresa en el planteamiento anterior. Por el contrario, una observación sencilla mostraría que la autocorrelación observada sería ligeramente distinta a la autocorrelación teórica. Ahora veamos algunos ejemplos.

Ejemplo. En el primer ejemplo simularemos una serie y mostraremos el análisis de un proceso construido considerando un proceso puramente aleatorio como componente U_t . Por su parte, en un segundo ejemplo aplicaremos el análisis a una serie de tiempo de una variable económica observada.

Para el primer ejemplo, consideremos un proceso dado por la forma de un $AR(1)$ como en la ecuación (3.33) cuya solución está dada por la ecuación (3.35). En específico, supongamos que el término o componente estocástico U_t es una serie generada a partir de números aleatorios de una función normal con media 0 y desviación estándar 4. Los detalles del proceso simulado se muestran en las siguientes gráficas.

La Figura 3.4 ilustra el comportamiento que se debería observar en una serie considerando el procedimiento iterativo de construcción. Por su parte, la Figura 3.5 ilustra el proceso o trayectoria de la solución de la serie de tiempo. Finalmente, las Figuras 3.6 y 3.7 muestran el correlograma calculado considerando una función de autocorrelación aplicada al proceso real y una función de autocorrelación aplicada al proceso teórico, respectivamente.

```
library(ggplot2)
library(dplyr)
library(latex2exp)

a0 <- 5; a1 <- 0.9; X_0 <- (a0/(1 - a1)); T <- 1000

X_t <- data.frame(Tiempo = c(0:T))

set.seed(12345)

# Agregamos un término estocástico al data frame

X_t$U_t <- rnorm(T+1, mean = 0, sd = 4)

# Agregamos columnas con NA's para un proceso teorico y uno real
X_t$X_t <- NA
X_t$XR_t <- NA

# La serie teórica inicia en un valor inicial X_0
X_t$X_t[1] <- X_0

# La serie real inicia en un valor inicial X_0
```

```

X_t$XR_t[1] <- X_0

# Agregamos una columna para la función de Autocorrelación teórica:
X_t$rho <- NA

for (i in 2:(T + 1)) {
  # Real:
  X_t$XR_t[i] = a0 + a1*X_t$XR_t[i-1] + X_t$U_t[i-1]

  # Teórico:
  X_t$X_t[i] = X_t$X_t[i-1] + (a1^(i-1))*X_t$U_t[i-1]

  # Autocorrelación:
  X_t$rho[i-1] = a1^(i-1)
}

ggplot(data = X_t, aes(x = Tiempo, y = XR_t)) +
  geom_line(size = 0.5, color = "darkred") +
  #theme_bw() +
  xlab("Tiempo") +
  ylab(TeX("$X_t$")) +
  theme(plot.title = element_text(size = 11, face = "bold",
                                   hjust = 0)) +
  theme(plot.subtitle = element_text(size = 10, hjust = 0)) +
  theme(plot.caption = element_text(size = 10, hjust = 0)) +
  theme(plot.margin = unit(c(1,1,1,1), "cm")) +
  labs(
    title = "Comportamiento del Proceso Real ('Estimado')",
    subtitle = "Con un error con Distribución Normal (media = 0, desviación estándar = 1)",
    caption = "Fuente: Elaboración propia."
  )
)

ggsave("G_AR_1_Real.png", width = 20, height = 10, units = "cm")

```

```

ggplot(data = X_t, aes(x = Tiempo, y = X_t)) +
  geom_line(size = 0.5, color = "darkblue") +
  #theme_bw() +
  xlab("Tiempo") +
  ylab(TeX("$X_t$")) +
  theme(plot.title = element_text(size = 11, face = "bold",
                                   hjust = 0)) +
  theme(plot.subtitle = element_text(size = 10, hjust = 0)) +
  theme(plot.caption = element_text(size = 10, hjust = 0)) +
  theme(plot.margin = unit(c(1,1,1,1), "cm")) +

```

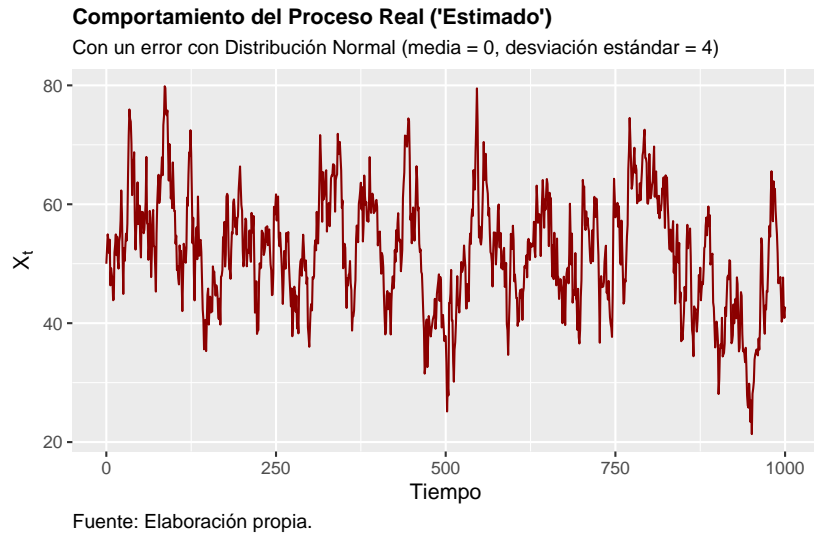


Figure 3.4: Comportamiento del Proceso Real ('Estimado')

```
labs(
  title = "Comportamiento del Proceso Teórico",
  subtitle = "Con un error con Distribución Normal (media = 0, desviación estándar = 4)",
  caption = "Fuente: Elaboración propia."
)

ggsave("G_AR_1_Teo.png", width = 20, height = 10, units = "cm")

acf(X_t$XR_t, lag.max = 30, col = "blue",
  ylab = "Autocorrelacion",
  xlab="Rezagos",
  main="Funcion de Autocorrelacion Real")

barplot(X_t$rho[1:30], names.arg = c(1:30), col = "blue",
  border="blue", density = c(10,20),
  ylab = "Autocorrelacion",
  xlab="Rezagos",
  main="Funcion de Autocorrelacion Teórica")
```

Recordemos que una trayectoria de equilibrio o solución de un $AR(1)$ es como se muestra en la ecuación (3.35). Así, nuestra serie simulada cumple con la

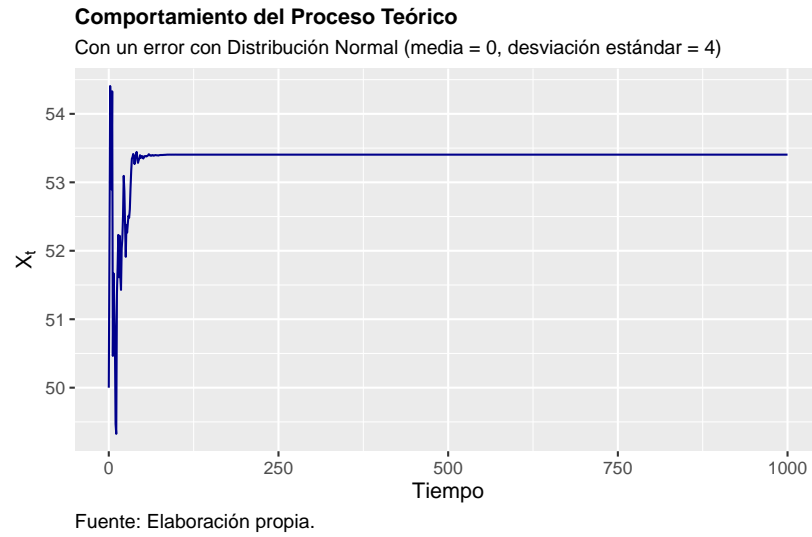


Figure 3.5: Comportamiento del Proceso Teórico

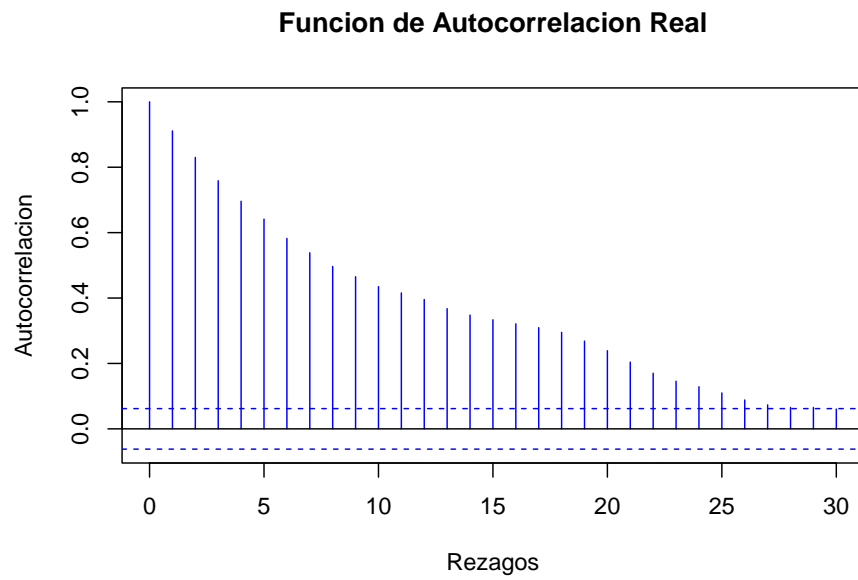


Figure 3.6: Funcion de Autocorrelacion Real

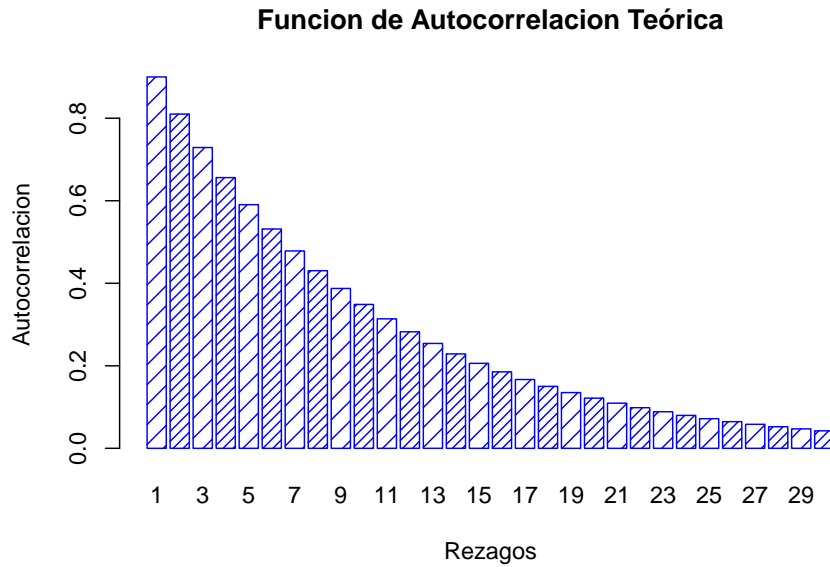


Figure 3.7: Funcion de Autocorrelacion Teórica

característica de que los errores son más relevantes cuando la serie es corta. Por el contrario, los errores son menos relevantes, cuando la serie es muy larga. La Figura 3.8 ilustra esta observación de la trayectoria de equilibrio.

```
ggplot(data = X_t, aes(x = Tiempo)) +
  geom_line(aes(y = XR_t), size = 0.5, color = "darkred") +
  geom_line(aes(y = X_t), size = 0.5, color = "darkblue") +
  #theme_bw() +
  xlab("Tiempo") +
  ylab(TeX("$X_t$")) +
  theme(plot.title = element_text(size = 11, face = "bold",
                                   hjust = 0)) +
  theme(plot.subtitle = element_text(size = 10, hjust = 0)) +
  theme(plot.caption = element_text(size = 10, hjust = 0)) +
  theme(plot.margin = unit(c(1,1,1,1), "cm")) +
  labs(
    title = "Comportamiento de los Procesos Real y Teórico",
    subtitle = "Con un error con Distribución Normal (media = 0, desviación estándar = 4)",
    caption = "Fuente: Elaboración propia."
  )
```

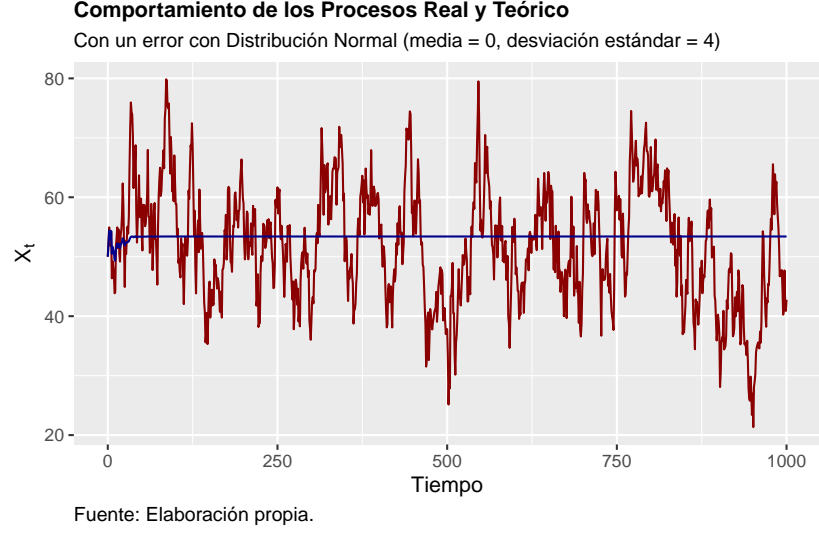


Figure 3.8: Comportamiento de los Procesos Real y Teórico

```
ggsave("G_AR_1_Comb.png", width = 20, height = 10, units = "cm")
```

Ejemplo. Para el segundo ejemplo consideremos una aplicación a una serie de tiempo en específico: Pasajeros transportados mensualmente en el Sistema de Transporte Colectivo Metro (pasajeros medidos en millones).²

A la serie se le aplicará una metodología de estimación dada por el método de Máxima Verosimilitud (ML, por sus siglas en inglés). Antes de realizar el proceso de estimación, consideremos una transformación de diferencias logarítmicas con el objeto de obtener una serie de tiempo expresada en tasas de crecimiento³ y con un comportamiento parecido a un proceso estacionario.

Así, para cada una de las series que analicemos en diferencias logarítmicas respecto del momento k las expresaremos bajo la siguiente transformación:

$$DLX_t = \log(X_t) - \log(X_{t-k})$$

Donde $k = 1, 2, 3, \dots$ y $\log(\cdot)$ es la función logaritmo natural. Esta expresión se puede interpretar como una tasa de crecimiento, puesto que asumimos variaciones pequeñas para las cuales se cumple que: $\log(X_t) - \log(X_{t-k}) \approx \frac{X_t - X_{t-k}}{X_t}$.

²Fuente: INEGI, <https://www.inegi.org.mx/app/indicadores/?tm=0&t=1090>.

³Estas tasas no son porcentuales, para hacerlas porcentuales faltaría multiplicar por 100 cada calor de la serie.

Primero, para realizar el análisis de una serie de tiempo deberemos decidir si éste se realizará para la serie en niveles o en diferencias. Por convención, decimos que la serie está en niveles si ésta se analiza sin hacerle ninguna transformación o si se analiza aplicando solo logaritmos. Cuando la serie se analiza en diferencias significa que la diferencia se hace sin aplicar logaritmos o aplicando logaritmos. Sin embargo, la convención es hacer un análisis en diferencias logarítmicas.

Para decidir cómo analizar la serie de pasajeros en el metro de la CDMX en la Figura 3.9 se muestra la gráfica de la serie en niveles (sin transformación logarítmica y con transformación logarítmica) y en diferencias logarítmicas mensuales (es decir, con $k = 1$).

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)

Datos <- read_excel("BD/Base_Transporte.xlsx",
                    sheet = "Datos", col_names = TRUE)

# En Niveles
Pax_Metro <- ts(Datos$Pax_Metro, start = c(2000, 1),
               freq = 12)

# En Logaritmos:
Pax_LMetro <- ts(log(Datos$Pax_Metro), start = c(2000, 1),
               freq = 12)

# Diferencias mensuales:
Pax_DLMetro <- ts( log(Datos$Pax_Metro) -
                  lag( log(Datos$Pax_Metro), 1 ),
                  start = c(2000, 1), freq = 12)

#
par(mfrow = c(3,1))

plot(Pax_Metro, xlab = "Tiempo",
     main = "Pasajeros transportados (Millones) en el SCM",
     col = "darkgreen")

plot(Pax_LMetro, xlab = "Tiempo",
     main = "LN Pasajeros transportados (Millones) en el SCM",
     col = "darkblue")

plot(Pax_DLMetro, xlab = "Tiempo",
     main = "Diff LN Pasajeros transportados (Millones) en el SCM",
```

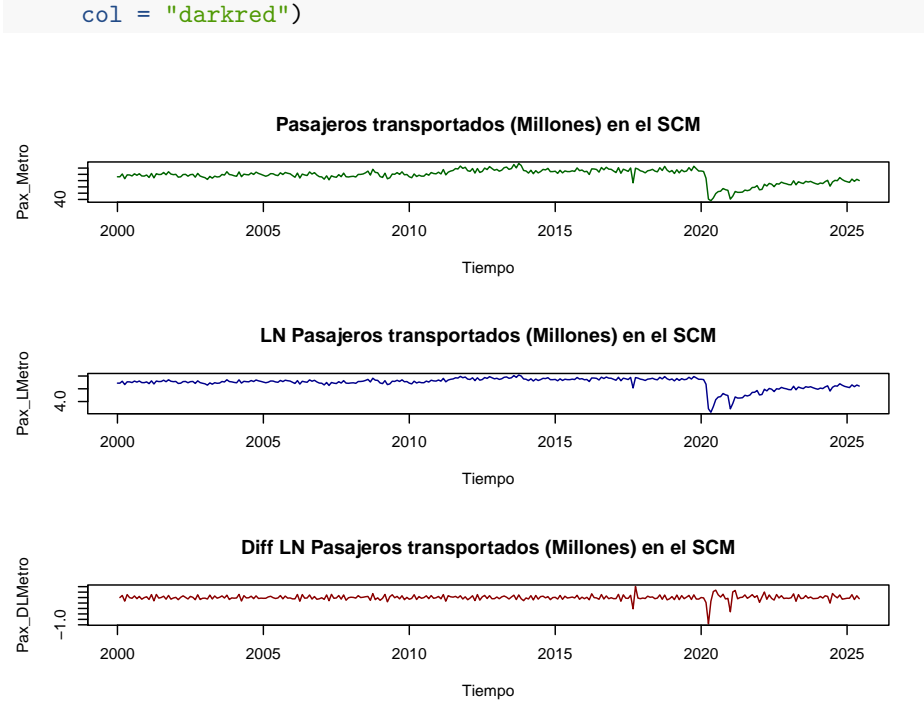


Figure 3.9: Pasajeros transportados (Millones) en el metro de la CDM en niveles y en diferencias logarítmicas

```
par(mfrow=c(1,1))
```

A continuación, estimaremos una $AR(1)$ para la serie en niveles bajo la transformación logarítmica ($PaxLMetro_t$) y en diferencias logarítmicas ($PaxDLMetro_t$). Para el primer caso, obtenemos el siguiente resultado:

Table 3.2: $AR(1)$ para la variable $PaxLMetro_t$.

$PaxLMetro_t$	=	4.7419	+	0.5916	$PaxLMetro_{t-1}$
		(0.0105)		(0.0526)	
$\hat{\sigma}^2$	=	0.004335	AIC	=	-602.73

Para el segundo caso, obtenemos el siguiente resultado:

Table 3.3: AR(1) para la variable $PaxDLMetro_t$.

$PaxDLMetro_t$	=	0.0007 (0.0023)	—	0.6194 (0.0511)	$PaxDLMetro_{t-1}$
$\hat{\sigma}^2$	=	0.003344	AIC	=	−660.53

En ambos casos observamos que el parámetro asociado al componente AR es significativo y cumple con la restricción de ser en valor absoluto menor a 1, por lo que la solución asociada al proceso será convergente. También en ambos casos se reporta la estadística o Criterio de Información de Akaike (AIC, por sus siglas en inglés), misma que más adelante discutiremos su importancia y aplicación.

3.4.2 AR(2)

Una vez analizado el caso de $AR(1)$ analizaremos el caso del $AR(2)$. La ecuación generalizada del proceso autoregresivo de orden 2 (denotado como $AR(2)$) puede ser escrita como:

$$X_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + U_t \quad (3.40)$$

Donde U_t denota un proceso puramente aleatorio con media cero (0), varianza constante (σ^2) y autocovarianza cero ($Cov(U_t, U_s) = 0$, con $t \neq s$), y un parámetro $a_2 \neq 0$. Así, utilizando el operador rezago podemos reescribir la ecuación (3.40) como:

$$\begin{aligned} X_t - a_1 X_{t-1} - a_2 X_{t-2} &= a_0 + U_t \\ (1 - a_1 L^1 - a_2 L^2) X_t &= a_0 + U_t \end{aligned}$$

Donde, vamos a denotar a $\alpha(L) = (1 - a_1 L^1 - a_2 L^2)$, y lo denotaremos como un polinomio que depende del operador rezago y que es distinto de cero. De esta forma podemos reescribir la ecuación (3.40) como:

$$\alpha(L) X_t = a_0 + U_t \quad (3.41)$$

Ahora, supongamos que existe el inverso multiplicativo del polinomio $\alpha(L)$, el cual será denotado como: $\alpha^{-1}(L)$ y cumple con que:

$$\alpha^{-1}(L) \alpha(L) = 1 \quad (3.42)$$

Así, podemos escribir la solución a la ecuación (3.40) como:

$$X_t = \alpha^{-1}(L) a_0 + \alpha^{-1}(L) U_t$$

Si utilizamos el hecho que $\alpha^{-1}(L)$ se puede descomponer a través del procedimiento de Wold en un polinomio de forma similar al caso de $AR(1)$, tenemos que:

$$\alpha^{-1}(L) = \psi_0 + \psi_1 L + \psi_2 L^2 + \dots \quad (3.43)$$

Por lo tanto, el inverso multiplicativo $\alpha^{-1}(L)$ se puede ver como:

$$1 = (1 - a_1 L - a_2 L^2)(\psi_0 + \psi_1 L + \psi_2 L^2 + \dots) \quad (3.44)$$

Desarrollando la ecuación (3.44) tenemos la siguiente expresión:

$$\begin{aligned} 1 &= \psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots \\ &\quad - a_1 \psi_0 L - a_1 \psi_1 L^2 - a_1 \psi_2 L^3 - \dots \\ &\quad - a_2 \psi_0 L^2 - a_2 \psi_1 L^3 - \dots \end{aligned}$$

Ahora, podemos agrupar todos los términos en función del exponente asociado al operador rezago L . La siguiente es una solución particular y es una de las múltiples que podrían existir que cumpla con la ecuación (3.44). Sin embargo, para efectos del análisis, sólo necesitamos una de esas soluciones. Utilizaremos las siguientes condiciones que deben cumplirse en una de las posibles soluciones:

$$\begin{aligned} L^0 &: \Rightarrow \psi_0 = 1 \\ L^1 &: \psi_1 - a_1 \psi_0 = 0 \Rightarrow \psi_1 = a_1 \\ L^2 &: \psi_2 - a_1 \psi_1 - a_2 \psi_0 = 0 \Rightarrow \psi_2 = a_1^2 + a_2 \\ L^3 &: \psi_3 - a_1 \psi_2 - a_2 \psi_1 = 0 \Rightarrow \psi_3 = a_1^3 + 2a_1 a_2 \\ &\vdots \end{aligned}$$

De esta forma podemos observar que en el límite siempre obtendremos una ecuación del tipo $\psi_j - a_1 \psi_{j-1} - a_2 \psi_{j-2} = 0$ asociada a cada uno de los casos en que exista un L^j , donde $j \neq 0, 1$, y la cual siempre podremos resolver conociendo que las condiciones iniciales son: $\psi_0 = 1$ y $\psi_1 = a_1$.

Así, de las relaciones antes mencionadas y considerando que $\alpha^{-1}(L)$ aplicada a una constante como a_0 , tendrá como resultado otra constante. De esta forma podemos escribir que la solución del proceso $AR(2)$ en la ecuación (3.40) será dada por una expresión como sigue:

$$X_t = \frac{a_0}{1 - a_1 - a_2} + \sum_{j=0}^{\infty} \psi_{t-j} U_{t-j} \quad (3.45)$$

Donde todos los parámetros ψ_i están determinados por los parámetros a_0 , a_1 y a_2 . En particular, $\psi_0 = 1$ y $\psi_1 = a_1$, como describimos anteriormente. Al igual

que en el caso del $AR(1)$, en la ecuación (3.45) las condiciones de estabilidad estarán dadas por las soluciones del siguiente polinomio característico:⁴

$$\lambda^2 - \lambda a_1 - a_2 = 0 \quad (3.46)$$

Así, la condición de estabilidad de la trayectoria es que $|\lambda_i| < 1$, para $i = 1, 2$. Es decir, es necesario que cada una de las raíces sea, en valor absoluto, siempre menor que la unidad. Estas son las condiciones de estabilidad para el proceso $AR(2)$.

Finalmente, al igual que en un $AR(1)$, a continuación determinamos los momentos de una serie que sigue un proceso $AR(2)$. Iniciamos con la determinación de la media de la serie:

$$\mathbb{E}[X_t] = \mu = \frac{a_0}{1 - a_1 - a_2} \quad (3.47)$$

Lo anterior es cierto puesto que $\mathbb{E}[U_{t-i}] = 0$, para todo $i = 0, 1, 2, \dots$. Para determinar la varianza utilizaremos las siguientes relaciones basadas en el uso del valor esperado, varianza y covarianza de la serie. Adicionalmente, para simplificar el trabajo asumamos que $a_0 = 0$, lo cual implica que $\mu = 0$. Dicho lo anterior, partamos de:

$$\begin{aligned} \mathbb{E}[X_t X_{t-\tau}] &= \mathbb{E}[(a_1 X_{t-1} + a_2 X_{t-2} + U_t) X_{t-\tau}] \\ &= a_1 \mathbb{E}[X_{t-1} X_{t-\tau}] + a_2 \mathbb{E}[X_{t-2} X_{t-\tau}] + \mathbb{E}[U_t X_{t-\tau}] \end{aligned}$$

Donde $\tau = 0, 1, 2, 3, \dots$ y $\mathbb{E}[U_t X_{t-\tau}] = 0$ para todo $\tau \neq 0$.⁵ Dicho esto, podemos derivar el valor del valor esperado para diferentes valores de τ :

$$\begin{aligned} \tau = 0 & : \quad \gamma(0) = \alpha_1 \gamma(1) + \alpha_2 \gamma(2) + \sigma^2 \\ \tau = 1 & : \quad \gamma(1) = \alpha_1 \gamma(0) + \alpha_2 \gamma(1) \\ \tau = 2 & : \quad \gamma(2) = \alpha_1 \gamma(1) + \alpha_2 \gamma(0) \\ & \vdots \end{aligned}$$

Donde debe ser claro que $\mathbb{E}[(X_t - \mu)(X_{t-\tau} - \mu)] = \mathbb{E}[X_t X_{t-\tau}] = \gamma(\tau)$. Así, en general cuando $\tau \neq 0$:

$$\gamma(\tau) = a_1 \gamma(\tau - 1) + a_2 \gamma(\tau - 2) \quad (3.48)$$

Realizando la sustitución recursiva y solucionando el sistema respectivo obtenemos que la varianza y las covarianzas estarán determinadas por:

$$\text{Var}[X_t] = \gamma(0) = \frac{1 - a_2}{(1 + a_2)[(1 - a_2)^2 - a_1^2]} \sigma^2 \quad (3.49)$$

⁴Note que raíces son equivalentes al inverso de las del polinomio dado por $\lambda^2 a_2 - \lambda a_1 - 1 = 0$.

⁵Es fácil demostrar esta afirmación, sólo requiere de desarrollar la expresión y utilizar el hecho de que U_t es un proceso pueramente aleatorio, por lo que la covarianza es cero (0).

$$\gamma(1) = \frac{a_1}{(1+a_2)[(1-a_2)^2 - a_1^2]} \sigma^2 \quad (3.50)$$

$$\gamma(2) = \frac{a_1^2 + a_2 - a_2^2}{(1 + a_2)[(1 - a_2)^2 - a_1^2]} \sigma^2 \quad (3.51)$$

Recordemos que las funciones de autocorrelación se obtienen de la división de cada una de las funciones de covarianza ($\gamma(\tau)$) por la varianza ($\gamma(0)$). Así, podemos construir la siguiente expresión:

$$\rho(\tau) - a_1\rho(\tau - 1) - a_2\rho(\tau - 2) = 0 \quad (3.52)$$

Ejemplo. Utilizaremos la serie de Pasajeros en vuelos nacionales (en vuelos de salidas) para estimar un $AR(2)$ mediante el método de máxima verosimilitud (ML, por sus siglas en inglés). Antes de realizar el proceso de estimación, consideremos una transformación de la serie en logaritmos y una más en diferencias logarítmicas; lo anterior con el objeto de obtener una serie de tiempo suavizada y expresada en tasas de crecimiento, con un comportamiento parecido a un proceso estacionario.

Así, para cada una de las series que analicemos en diferencias logarítmicas, las expresaremos bajo la siguiente transformación:

$$DLX_t = \log(X_t) - \log(X_{t-k})$$

Donde $k = 1, 2, 3, \dots$ y $\log(\cdot)$ es la función logaritmo natural. Por convención, decimos que la serie está en niveles si esta se analiza sin hacerle ninguna transformación o se analiza en logaritmos. Cuando la serie se analiza en diferencias significa que la diferencia se hace sin aplicar logaritmos. Y cuando la serie analizada está en diferencias logarítmicas también diremos que está en diferencias. Sin embargo, lo común es hacer un análisis en logaritmos y en diferencias logarítmicas.

Primero, para decidir si se realizará un AR(2) para la serie en niveles o en diferencias, analizaremos su gráfica. La serie en niveles, en niveles bajo una transformación logarítmica y en diferencias logarítmicas mensuales de los pasajeros en vuelos nacionales se muestra en la Figura 3.10.

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)

Datos <- read_excel("BD/Base_Transporte.xlsx",
                    sheet = "Datos", col_names = TRUE)
```

```

# En Niveles
Pax_Nal <- ts(Datos$Pax_Nal,
              start = c(2000, 1),
              freq = 12)

# Logaritmos:
LPax_Nal <- ts(log(Datos$Pax_Nal),
               start = c(2000, 1),
               freq = 12)

# Diferencias mensuales:
DLPax_Nal <- ts(log(Datos$Pax_Nal) -
                lag(log(Datos$Pax_Nal), 1),
                start = c(2000, 1), freq = 12)

#
par(mfrow=c(3,1))

plot(Pax_Nal, xlab = "Tiempo", ylab = "Pasajeros",
      main = "Pasajeros en vuelos nacionales de salida",
      col = "darkgreen")

plot(LPax_Nal, xlab = "Tiempo", ylab = "LN Pasajeros",
      main = "LN Pasajeros en vuelos nacionales de salida",
      col = "darkblue")

plot(DLPax_Nal, xlab = "Tiempo", ylab = "DLN Pasajeros",
      main = "Diff LN Pasajeros en vuelos nacionales de salida",
      col = "darkred")

par(mfrow=c(1,1))

```

A continuación, estimaremos un $AR(2)$ para la serie en niveles bajo una transformación logarítmica ($LPaxNal_t$) y en diferencias logarítmicas ($DLPaxNal_t$). Para el primer caso obtenemos el siguiente resultado:

Table 3.4: $AR(2)$ para la variable $LPaxNal_t$.

$LPaxNal_t$	=	14.6267 (0.1816)	+	0.7637 (0.0637)	$LPaxNal_{t-1}$
		0.2025 (0.0646)	$LPaxNal_{t-2}$		
$\hat{\sigma}^2$	=	0.01138	AIC	=	-372.64

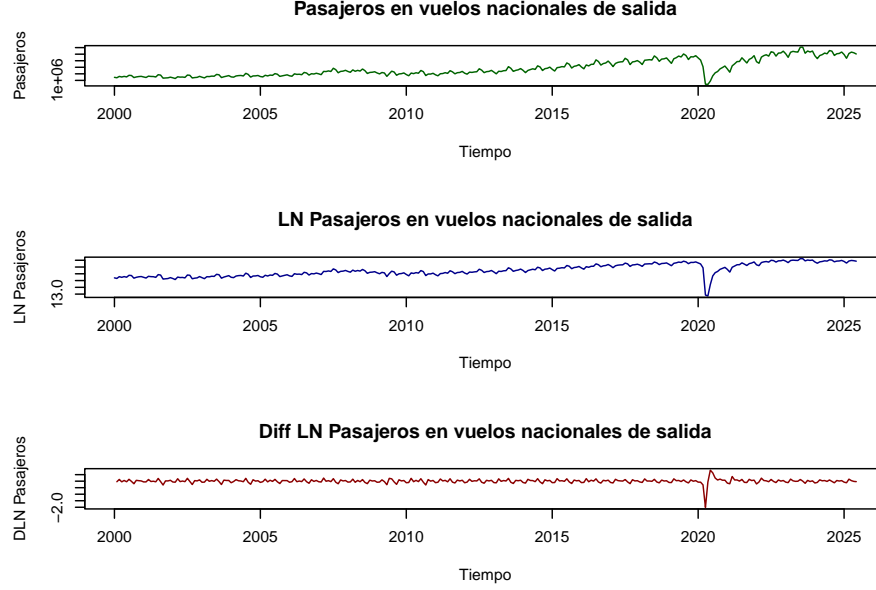


Figure 3.10: Pasajeros en vuelos de salidas nacionales en niveles y en diferencias logaritmicas

Para el segundo caso obtenemos el siguiente resultado:

Table 3.5: AR(2) para la variable $DLPaxNal_t$.

$DLPaxNal_t$	=	0.0050 (0.0036)	+	0.3205 (0.0592)	$DLPaxNal_{t-1}$
		0.4242 (0.0591)			$DLPaxNal_{t-2}$
$\hat{\sigma}^2$	=	0.009378	AIC	=	-418.3

Para ambos casos, entre paréntesis indicamos los errores estándar y reportamos el estadístico de Akaike, AIC. Finalmente, podemos determinar si las soluciones serán convergentes para ello en la Figura 3.11 mostramos las raíces asociadas a cada uno de los polinomios. De la inspección visual, podemos concluir que ambas propuestas de AR(2) representan una solución convergente y estable.

```
source("arroots.R")
source("plot.armacroots.R")
```



```
AR_LPax_Nal <- arima(LPax_Nal, order = c(2, 0, 0),
                     method = "ML")

AR_DLPax_Nal <- arima(DLPax_Nal, order = c(2, 0, 0),
                      method = "ML")

par(mfrow=c(1,2))

plot.armacoots(aroots(AR_LPax_Nal),
               main="Inverse AR roots of \nAR(2): LN Pax Nal")

#
plot.armacoots(aroots(AR_DLPax_Nal),
               main="Inverse AR roots of \nAR(2): Diff LN Pax Nal")
```

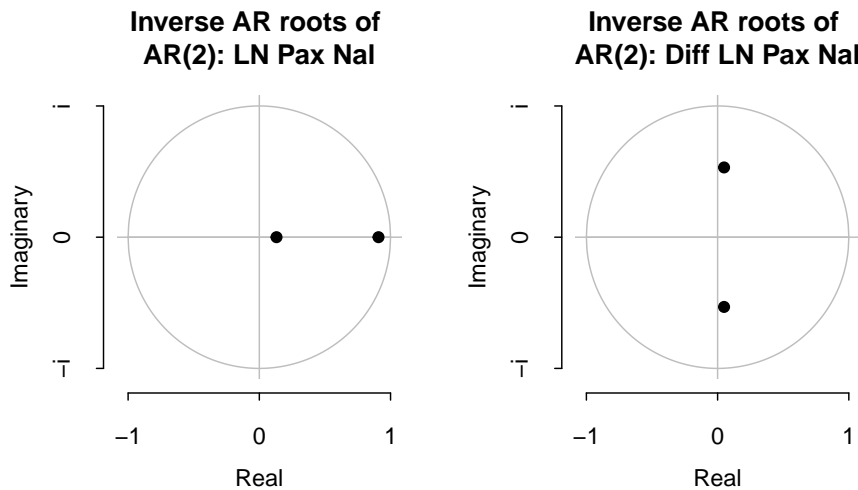


Figure 3.11: Inverso de las Raíces del polinomio característico

```
par(mfrow=c(1,1))
```

3.4.3 AR(p)

Veremos ahora una generalización de los procesos autoregresivos (AR). Esta generalización es conocida como un proceso $AR(p)$ y que puede ser descrito por

la siguiente ecuación en diferencia estocástica:

$$X_t = a_0 + a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + \dots + a_p X_{t-p} + U_t \quad (3.53)$$

Donde $a_p \neq 0$, y U_t es un proceso puramente aleatorio con media cero (0), varianza constante (σ^2) y covarianza cero (0). Usando el operador rezago, L^k , para $k = 0, 1, 2, \dots, p$, obtenemos la siguiente expresión de la ecuación (3.53):

$$(1 - a_1 L - a_2 L^2 - a_3 L^3 - \dots - a_p L^p) X_t = a_0 + U_t \quad (3.54)$$

Definamos el polinomio $\alpha(L)$ como:

$$\alpha(L) = 1 - a_1 L - a_2 L^2 - a_3 L^3 - \dots - a_p L^p \quad (3.55)$$

De forma similar que en los procesos $AR(1)$ y $AR(2)$, las condiciones de estabilidad del proceso $AR(p)$ estarán dadas por la solución de la ecuación característica:

$$\lambda^p - a_1 \lambda^{p-1} - a_2 \lambda^{p-2} - a_3 \lambda^{p-3} - \dots - a_p = 0 \quad (3.56)$$

Así, solo si el polinomio anterior tiene raíces cuyo valor absoluto sea menor a uno ($|\lambda_i| < 1$) y si $1 - a_1 L - a_2 L^2 - a_3 L^3 - \dots - a_p L^p < 1$ podremos decir que el proceso es convergente y estable. Lo anterior significa que la ecuación (3.55) puede expresarse en términos de la descomposición de Wold o como la suma infinita de términos como:

$$\frac{1}{1 - a_1 L - a_2 L^2 - a_3 L^3 - \dots - a_p L^p} = \psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots \quad (3.57)$$

Donde, por construcción de $\alpha(L)\alpha^{-1}(L) = 1$ implica que $\psi_0 = 1$. De forma similar a los procesos $AR(1)$ y $AR(2)$, es posible determinar el valor de los coeficientes ψ_j en términos de los coeficientes a_i . Así, la solución del proceso $AR(p)$ estará dada por:

$$X_t = \frac{a_0}{1 - a_1 - a_2 - a_3 - \dots - a_p} + \sum_{j=0}^{\infty} \psi_j U_{t-j} \quad (3.58)$$

Considerando la solución de la ecuación (3.53) expresada en la ecuación (3.58) podemos determinar los momentos del proceso y que estarán dados por una media como:

$$\mathbb{E}[X_t] = \mu = \frac{a_0}{1 - a_1 - a_2 - a_3 - \dots - a_p} \quad (3.59)$$

Lo anterior, considerando que $\mathbb{E}[U_t] = 0$, para todo t . Para determinar la varianza del proceso, sin pérdida de generalidad, podemos definir una ecuación: $\gamma(\tau) = \mathbb{E}[X_{t-\tau} X_t]$, la cual (omitiendo la constante, ya que la correlación de una

constante con cualquier variable aleatoria que depende del tiempo es cero (0)) puede ser escrita como:

$$\gamma(\tau) = \mathbb{E}[(X_{t-\tau}) \cdot (a_1 X_{t-1} + a_2 X_{t-2} + a_3 X_{t-3} + \dots + a_p X_{t-p} + U_t)] \quad (3.60)$$

Donde $\tau = 0, 1, 2, \dots, p$ y $a_0 = 0$, lo que implica que $\mu = 0$. De lo anterior obtenemos el siguiente conjunto de ecuaciones mediante sustituciones de los valores de τ :

$$\begin{aligned} \gamma(0) &= a_1 \gamma(1) + a_2 \gamma(2) + \dots + a_p \gamma(p) + \sigma^2 \\ \gamma(1) &= a_1 \gamma(0) + a_2 \gamma(1) + \dots + a_p \gamma(p-1) \\ &\vdots \\ \gamma(p) &= a_1 \gamma(p-1) + a_2 \gamma(p-2) + \dots + a_p \gamma(0) \end{aligned}$$

De esta forma, es fácil observar que la ecuación general para $p > 0$ estará dada por:

$$\gamma(p) - a_1 \gamma(p-1) - a_2 \gamma(p-2) - \dots - a_p \gamma(0) = 0 \quad (3.61)$$

Dividiendo la ecuación (3.61) por $\gamma(0)$, se obtiene la siguiente ecuación:

$$\rho(p) - a_1 \rho(p-1) + a_2 \rho(p-2) + \dots + a_p \rho(0) = 0 \quad (3.62)$$

Así, podemos escribir el siguiente sistema de ecuaciones:

$$\begin{aligned} \rho(1) &= a_1 + a_2 \rho(1) + a_3 \rho(2) + \dots + a_p \rho(p-1) \\ \rho(2) &= a_1 \rho(1) + a_2 + a_3 \rho(1) + \dots + a_p \rho(p-2) \\ &\vdots \\ \rho(p) &= a_1 \rho(p-1) + a_2 \rho(p-2) + \dots + a_p \end{aligned}$$

Lo anterior se puede expresar como un conjunto de vectores y matrices de la siguiente forma:

$$\begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(p) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & 1 & \dots & \rho(p-2) \\ \rho(2) & \rho(1) & \dots & \rho(p-3) \\ \vdots & \vdots & \dots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_p \end{bmatrix} \quad (3.63)$$

De lo anterior podemos escribir la siguiente ecuación que es una forma alternativa para expresar los valores de los coeficientes a_i de la solución del proceso $AR(p)$:

$$\rho = \mathbf{R} \mathbf{a} \quad (3.64)$$

Es decir, podemos obtener la siguiente expresión:

$$\mathbf{a} = \mathbf{R}^{-1} \rho \quad (3.65)$$

Ejemplo. Utilizaremos la serie de Pasajeros en vuelos internacionales de salida para estimar un $AR(p)$ mediante el método de máxima verosimilitud (ML). Antes de realizar el proceso de estimación, consideremos una transformación de la serie en logaritmos y una más en diferencias logarítmicas; lo anterior con el objeto de obtener una serie de tiempo suavizada y expresada en tasas de crecimiento, con un comportamiento parecido a un proceso estacionario.

Primero, para decidir si se realizará un $AR(p)$ para la serie en niveles o en diferencias, analizaremos su gráfica. La serie de Pasajeros en vuelos internacionales de salidas se muestra en la Figura 3.12. En esta se muestra la gráfica de la serie en niveles (sin transformación logarítmica y con transformación logarítmica) y en diferencias logarítmicas mensuales (es decir, con diferencia respecto del mes inmediato anterior).

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)

Datos <- read_excel("BD/Base_Transporte.xlsx",
                    sheet = "Datos", col_names = TRUE)

# En Niveles
Pax_Int <- ts(Datos$Pax_Int,
              start = c(2000, 1),
              freq = 12)

# Logaritmos:
LPax_Int <- ts(log(Datos$Pax_Int),
               start = c(2000, 1),
               freq = 12)

# Diferencias mensuales:
DLPax_Int <- ts(log(Datos$Pax_Int) - lag(log(Datos$Pax_Int), 1),
                start = c(2000, 1),
                freq = 12)

#
par(mfrow=c(3,1))

plot(Pax_Int, xlab = "Tiempo", ylab = "Pasajeros",
     main = "Pasajeros en vuelos internacionales de salida",
     col = "darkgreen")

plot(LPax_Int, xlab = "Tiempo", ylab = "LN Pasajeros",
     main = "LN Pasajeros en vuelos internacionales de salida",
```

```
col = "darkblue")

plot(DLPax_Int, xlab = "Tiempo", ylab = "DLN Pasajeros",
     main = "Diff LN Pasajeros en vuelos internacionales de salia",
     col = "darkred")
```

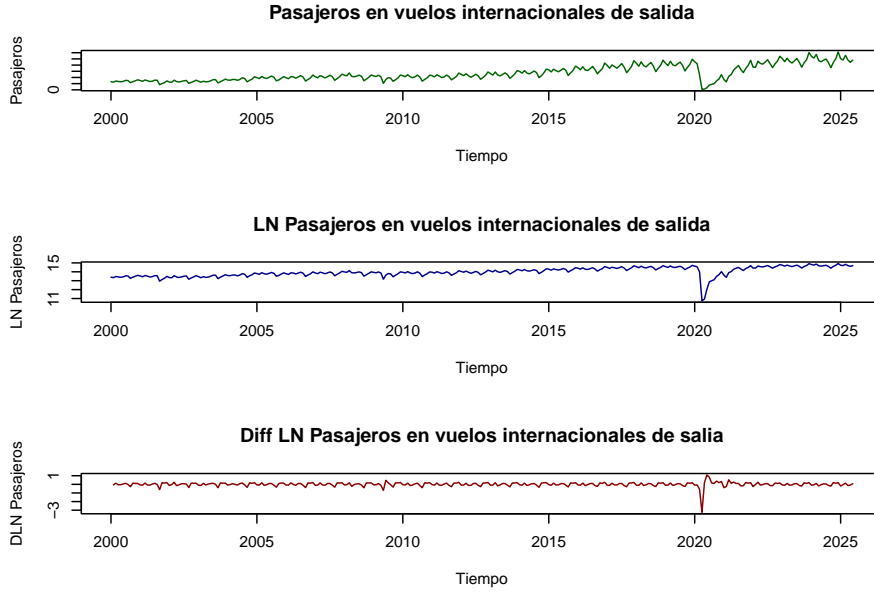


Figure 3.12: Pasajeros en vuelos internacionales de salida en niveles y en diferencias logarítmicas

```
par(mfrow=c(1,1))
```

De la gráfica en la Figura 3.12 observamos que quizá la mejor forma de estimar un $AR(p)$ es mediante la serie en diferencias, ya que ésta es la que parece ser una serie estacionaria. A continuación, estimaremos una $AR(4)$ para la serie en diferencias logarítmicas ($DLPaxInt_t$):

Table 3.6: $AR(4)$ para la variable $DLPaxInt_t$.

$DLPaxInt_t$	=	0.0050 (0.0052)	+	0.2701 (0.0655)	$DLPaxInt_{t-1}$
		-0.4326 (0.0664)			$DLPaxInt_{t-2}$

		-0.1956	$DLPaxNal_{t-3}$		
		(0.0664)			
		-0.0316	$DLPaxNal_{t-4}$		
		(0.0653)			
$\hat{\sigma}^2$	=	0.02371	AIC	=	-198.16

Entre paréntesis indicamos los errores estándar y reportamos el estadístico de Akaike, AIC. Finalmente, podemos determinar si las soluciones serán convergentes, para ello en la Figura 3.13 mostramos las raíces asociadas a cada uno de los polinomios. De la inspección visual podemos concluir que el AR(4) representan una solución convergente y estable.

```
source("arroots.R")

source("plot.armaroots.R")

AR_LPax_Int <- arima(LPax_Int, order = c(4, 0, 0),
                    method = "ML")

AR_DLPax_Int <- arima(DLPax_Int, order = c(4, 0, 0),
                    method = "ML")

par(mfrow=c(1,2))

plot.armaroots(arroots(AR_LPax_Int),
              main="Inverse AR roots of \nAR(2): LN Pax Int")

#
plot.armaroots(arroots(AR_DLPax_Int),
              main="Inverse AR roots of \nAR(2): Diff LN Pax Int")

par(mfrow=c(1,1))
```

3.5 Procesos de Medias Móviles (MA)

3.5.1 MA(1)

Una vez planteado el proceso generalizado de $AR(p)$, iniciamos el planteamiento de los procesos de medias móviles, denotados como $MA(q)$. Iniciemos con el planteamiento del proceso $MA(1)$, que se puede escribir como una ecuación como la siguiente:

$$X_t = \mu + U_t - b_1 U_{t-1} \quad (3.66)$$

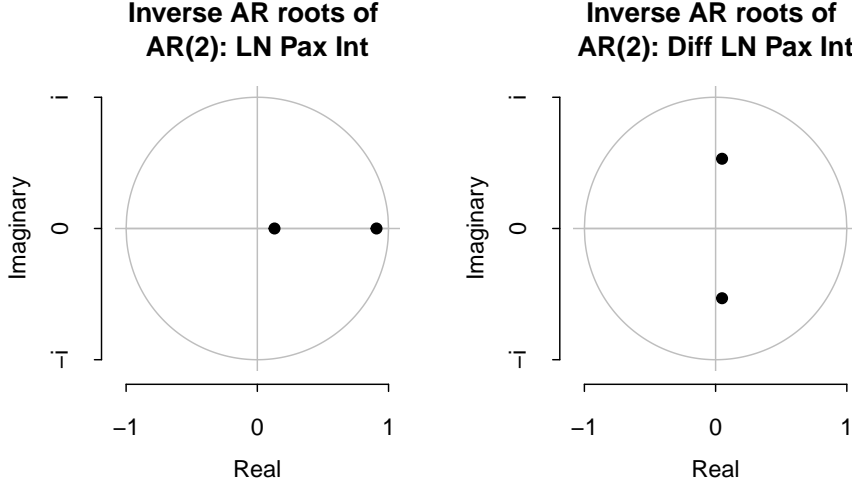


Figure 3.13: Inverso de las Raíces del polinomio característico

O como:

$$X_t - \mu = (1 - b_1 L)U_t \quad (3.67)$$

Donde U_t es un proceso puramente aleatorio, es decir, con $\mathbb{E}[U_t] = 0$, $\text{Var}[U_t] = \sigma^2$, y $\text{Cov}[U_t, U_s] = 0$. Así, un proceso $MA(1)$ puede verse como un proceso AR con una descomposición de Wold en la que $\psi_0 = 1$, $\psi_1 = -b_1$ y $\psi_j = 0$ para todo $j > 1$.

Al igual que los procesos autoregresivos, determinaremos los momentos de un proceso $MA(1)$. En el caso de la media, observamos que será:

$$\begin{aligned} \mathbb{E}[X_t] &= \mu + \mathbb{E}[U_t] - b_1 \mathbb{E}[U_{t-1}] \\ &= \mu \end{aligned} \quad (3.68)$$

Por su parte, la varianza estará dada por:

$$\begin{aligned} \text{Var}[X_t] &= \mathbb{E}[(X_t - \mu)^2] \\ &= \mathbb{E}[(U_t - b_1 U_{t-1})^2] \\ &= \mathbb{E}[U_t^2 - 2b_1 U_t U_{t-1} + b_1^2 U_{t-1}^2] \\ &= \mathbb{E}[U_t^2] - 2b_1 \mathbb{E}[U_t U_{t-1}] + b_1^2 \mathbb{E}[U_{t-1}^2] \\ &= \sigma^2 + b_1^2 \sigma^2 \\ &= (1 + b_1^2) \sigma^2 = \gamma(0) \end{aligned} \quad (3.69)$$

De esta forma, la varianza del proceso es constante en cualquier periodo t . Para determinar la covarianza utilizaremos la siguiente ecuación:

$$\begin{aligned}
 \mathbb{E}[(x_t - \mu)(x_{t+\tau} - \mu)] &= \mathbb{E}[(U_t - b_1 U_{t-1})(U_{t+\tau} - b_1 U_{t+\tau-1})] \\
 &= \mathbb{E}[U_t U_{t+\tau} - b_1 U_t U_{t+\tau-1} - b_1 U_{t-1} U_{t+\tau} \\
 &\quad + b_1^2 U_{t-1} U_{t+\tau-1}] \\
 &= \mathbb{E}[U_t U_{t+\tau}] - b_1 \mathbb{E}[U_t U_{t+\tau-1}] \\
 &\quad - b_1 \mathbb{E}[U_{t-1} U_{t+\tau}] + b_1^2 \mathbb{E}[U_{t-1} U_{t+\tau-1}] \quad (3.70)
 \end{aligned}$$

Si hacemos sustituciones de diferentes valores de τ en la ecuación (3.70) notaremos que la covarianza será distinta de cero únicamente para el caso de $\tau = 1, -1$. En ambos casos tendremos como resultado:

$$\begin{aligned}
 \mathbb{E}[(x_t - \mu)(x_{t+1} - \mu)] &= \mathbb{E}[(x_t - \mu)(x_{t-1} - \mu)] \\
 &= -b_1 \mathbb{E}[U_t U_t] \\
 &= -b_1 \mathbb{E}[U_{t-1} U_{t-1}] \\
 &= -b_1^2 \sigma^2 = \gamma(1) \quad (3.71)
 \end{aligned}$$

De esta forma tendremos que las funciones de autocorrelación estarán dadas por los siguientes casos:

$$\begin{aligned}
 \rho(0) &= 1 \\
 \rho(1) &= \frac{-b_1}{1 + b_1^2} \\
 \rho(\tau) &= 0 \text{ para todo } \tau > 1
 \end{aligned}$$

Ahora regresando a la ecuación (3.66), su solución la podemos expresar como:

$$\begin{aligned}
 U_t &= -\frac{\mu}{1 - b_1} + \frac{1}{1 - b_1 L} X_t \\
 &= -\frac{\mu}{1 - b_1} + X_t + b_1 X_{t-1} + b_1^2 X_{t-2} + \dots
 \end{aligned}$$

Donde la condición para que se cumpla esta ecuación es que $|b_1| < 1$. La manera de interpretar esta condición es como una condición de estabilidad de la solución y cómo una condición de invertibilidad. Notemos que un $MA(1)$ (y en general un $MA(q)$) es equivalente a un $AR(\infty)$, es decir, cuando se invierte un MA se genera un AR con infinitos rezagos.

En esta sección no desarrollaremos un ejemplo, primero explicaremos en qué consiste una modelación del tipo $MA(q)$ y después plantearemos un ejemplo en concreto.

3.5.2 MA(q)

En general, el proceso de medias móviles de orden q , $MA(q)$, puede ser escrito como:

$$X_t = \mu + U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q} \quad (3.72)$$

Podemos reescribir la ecuación (3.72) utilizando el operador rezago. Así, tendremos el proceso de $MA(q)$ como:

$$\begin{aligned} X_t - \mu &= (1 - b_1 L - b_2 L^2 - \dots - b_q L^q) U_t \\ X_t - \mu &= \beta(L) U_t \end{aligned} \quad (3.73)$$

Donde U_t es un proceso puramente aleatorio con $\mathbb{E}[U_t] = 0$, $Var[U_t] = \mathbb{E}[U_t^2] = \sigma^2$ y $Cov[U_t, U_s] = \mathbb{E}[U_t U_s] = 0$, y $\beta(L) = 1 - b_1 L - b_2 L^2 - \dots - b_q L^q$ es un polinomio del operador rezago L . La ecuación (3.73) puede ser interpretada como un proceso $AR(q)$ sobre la serie U_t .

Ahora determinemos los momentos de un proceso $MA(q)$:

$$\begin{aligned} \mathbb{E}[X_t] &= \mathbb{E}[\mu + U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q}] \\ &= \mu + \mathbb{E}[U_t] - b_1 \mathbb{E}[U_{t-1}] - b_2 \mathbb{E}[U_{t-2}] - \dots - b_q \mathbb{E}[U_{t-q}] \\ &= \mu \end{aligned} \quad (3.74)$$

En el caso de la varianza tenemos que se puede expresar como:

$$\begin{aligned} Var[X_t] &= \mathbb{E}[(X_t - \mu)^2] \\ &= \mathbb{E}[(U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q})^2] \\ &= \mathbb{E}[U_t^2 + b_1^2 U_{t-1}^2 + b_2^2 U_{t-2}^2 + \dots + b_q^2 U_{t-q}^2 \\ &\quad - 2b_1 U_t U_{t-1} - \dots - 2b_{q-1} b_q U_{t-q+1} U_{t-q}] \\ &= \mathbb{E}[U_t^2] + b_1^2 \mathbb{E}[U_{t-1}^2] + b_2^2 \mathbb{E}[U_{t-2}^2] + \dots + b_q^2 \mathbb{E}[U_{t-q}^2] \\ &\quad - 2b_1 \mathbb{E}[U_t U_{t-1}] - \dots - 2b_{q-1} b_q \mathbb{E}[U_{t-q+1} U_{t-q}] \\ &= \sigma^2 + b_1^2 \sigma^2 + b_2^2 \sigma^2 + \dots + b_q^2 \sigma^2 \\ &= (1 + b_1^2 + b_2^2 + \dots + b_q^2) \sigma^2 \end{aligned} \quad (3.75)$$

En el caso de las covarianzas podemos utilizar una idea similar al caso del $AR(p)$, construir una expresión general para cualquier rezago τ :

$$\begin{aligned} Cov[X_t, X_{t+\tau}] &= \mathbb{E}[(X_t - \mu)(X_{t+\tau} - \mu)] \\ &= \mathbb{E}[(U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q}) \\ &\quad (U_{t+\tau} - b_1 U_{t+\tau-1} - b_2 U_{t+\tau-2} - \dots - b_q U_{t+\tau-q})] \end{aligned}$$

La expresión anterior se puede desarrollar para múltiples casos de $\tau = 1, 2, \dots, q$. De esta forma tenemos el siguiente sistema:

$$\begin{aligned} \tau = 1 & : \gamma(1) = (-b_1 + b_1 b_2 + \dots + b_{q-1} b_q) \sigma^2 \\ \tau = 2 & : \gamma(2) = (-b_2 + b_1 b_3 + \dots + b_{q-2} b_q) \sigma^2 \\ & \vdots \\ \tau = q & : \gamma(q) = b_q \sigma^2 \end{aligned}$$

Donde $\gamma(\tau) = 0$ para todo $\tau > q$. Es decir, todas las autocovarianzas y autocorrelaciones con órdenes superiores a q son cero (0). De esta forma, esta característica teórica permite identificar el orden de $MA(q)$ visualizando la función de autocorrelación y verificando a partir de cuál valor de rezago la autocorrelación es no significativa.

Regresando al problema original que es el de determinar una solución para la ecuación (3.72), tenemos que dicha solución estará dada por un $AR(\infty)$ en términos de U_t :

$$\begin{aligned} U_t = & -\frac{\mu}{1 - b_1 - b_2 - \dots - b_q} + \beta(L)^{-1} X_t \\ & -\frac{\mu}{1 - b_1 - b_2 - \dots - b_q} + \sum_{j=0}^{\infty} c_j X_{t-j} \end{aligned} \quad (3.76)$$

Donde se cumple que: $1 = (1 - b_1 L^1 - b_2 L^2 - \dots - b_q L^q)(1 - c_1 L - c_2 L^2 - \dots)$ y los coeficientes c_j se pueden determinar por un método de coeficientes indeterminados y en términos de los valores b_i . De igual forma que en el caso de la ecuación (3.53), en la ecuación (3.76) se deben cumplir condiciones de estabilidad asociadas con las raíces del polinomio característico dado por:

$$1 - b_1 x - b_2 x^2 - \dots - b_q x^q = 0 \quad (3.77)$$

El cual debe cumplir que $|x_i| < 1$ y que $1 - b_1 - b_2 - \dots - b_q < 1$.

Ejemplo. Ahora veamos un ejemplo del proceso $MA(q)$, para lo cual retomaremos la serie de Pasajeros transportados en el metro de la CDMX (*PaxMetro*). Estimaremos el $MA(q)$ mediante el método de máxima verosimilitud (ML). Antes de realizar el proceso de estimación, consideremos una transformación de la serie en logaritmos y una más en diferencias logarítmicas; lo anterior con el objeto de obtener una serie de tiempo suavizada y expresada en tasas de crecimiento, con un comportamiento parecido a un proceso estacionario.

La serie de Pasajeros transportados en el metro de la CDMX se muestra en la Figura 3.9. En esta, se muestra la gráfica de la serie en niveles (sin transformación logarítmica y con transformación logarítmica) y en diferencias logarítmicas mensuales (es decir, con una diferencia respecto del mes inmediato anterior). Utilizaremos la serie en diferencias, ya que es la que parece ser estacionaria.

Esta serie tiene la peculiaridad de que tiene un salto a la baja y uno al alza entre septiembre de 2017 y octubre de 2017. Para controlar ese efecto, en nuestro modelo $MA(q)$ incluiremos dos variables dummies para dichos meses.

A continuación, estimaremos una $MA(4)$ para la serie en diferencias:

Table 3.7: $MA(4)$ para la variable $DLPaxMetro_t$.

$DLPaxMetro_t$	=	0.0000 (0.0000)	—	0.7804 (0.0661)	U_{t-1}
	+	0.3591 (0.0826)	U_{t-2}		
	—	0.2775 (0.0908)	U_{t-3}		
	—	0.1120 (0.0769)	U_{t-4}		
	—	0.3789 (0.0436)	$DSep2017$		
	+	0.3695 (0.0434)	$DOct2017$		
$\hat{\sigma}^2$	=	0.002176	AIC	=	−749.94

Entre parentésis indicamos los errores estándar y al final reportamos el estadístico de Akaike, AIC. Finalmente, podemos determinar si la solución serán convergente, para ello en la Figura 3.14 mostramos las raíces asociadas a cada uno de los polinomios. De la inspección visual, podemos concluir que ambas propuestas de $MA(4)$ representan una solución convergente y estable.

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)

Datos <- read_excel("BD/Base_Transporte_ARIMA.xlsx",
                    sheet = "Datos", col_names = TRUE)

source("maroots.R")

source("plot.armaroots.R")

Pax_Metro <- ts(Datos$Pax_Metro,
                start = c(2000, 1),
                freq = 12)

LPax_Metro <- ts(log(Datos$Pax_Metro),
```

```

        start = c(2000, 1),
        freq = 12)

DLPax_Metro <- ts(log(Datos$Pax_Metro) -
                  lag(log(Datos$Pax_Metro), 1),
                  start = c(2000, 1),
                  freq = 12)

D_Sep2017 <- ts(Datos$D_Sep2017,
                start = c(2000, 1),
                freq = 12)

D_Abr2020 <- ts(Datos$D_Abr2020,
                start = c(2000, 1),
                freq = 12)

D_May2020 <- ts(Datos$D_May2020,
                start = c(2000, 1),
                freq = 12)

MA_LPax_Metro <- arima(LPax_Metro, order = c(0, 0, 4),
                       method = "ML")

MA_LPax_Metro_2 <- arima(LPax_Metro, order = c(0, 0, 4),
                        xreg = cbind(D_Abr2020, D_Sep2017,
                                      D_May2020),
                        method = "ML")

par(mfrow=c(1,2))

plot.armacroots(marroots(MA_LPax_Metro),
                main="Inverse MA roots of \nMA(p): LN PAx Metro")

plot.armacroots(marroots(MA_LPax_Metro_2),
                main="Inverse MA roots of \nMA(p): LN PAx Metro con Dummy")

par(mfrow=c(1,1))

```

3.6 Procesos ARMA(p, q) y ARIMA(p, d, q)

Hemos establecido algunas relaciones de los procesos AR y los procesos MA, es decir, cómo un $MA(q)$ de la serie X_t puede ser reexpresada como un $AR(\infty)$

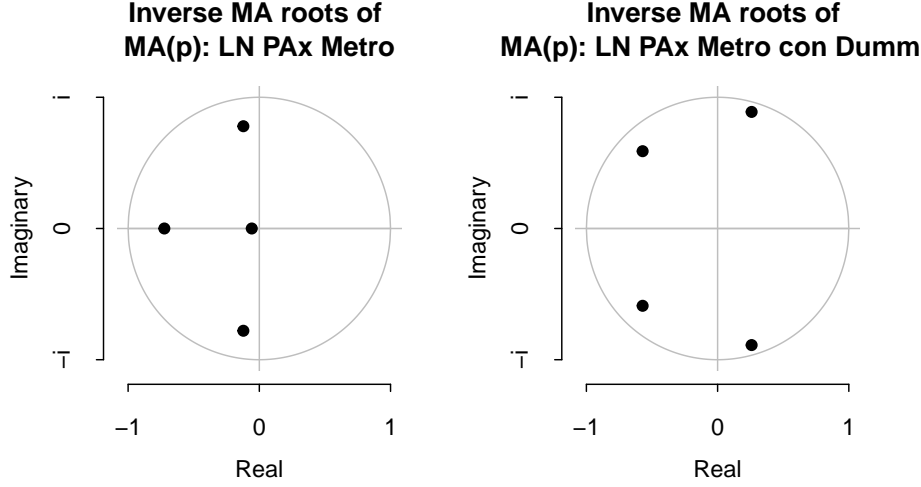


Figure 3.14: Inverso de las Raíces del polinomio característico

de la serie U_t , y viceversa un $AR(p)$ de la serie X_t puede ser reexpresada como un $MA(\infty)$.

En este sentido, para cerrar esta sección veamos el caso de la especificación que conjunta ambos modelos en un modelo general conocido como $ARMA(p, q)$ o $ARIMA(p, d, q)$. La diferencia entre el primero y el segundo es las veces que se tuvo que diferenciar la serie analizada, registro que se lleva en el índice d de los parámetros dentro del concepto $ARIMA(p, d, q)$. No obstante, en general nos referiremos al modelo como $ARMA(p, q)$ y dependerá del analista si modela la serie en niveles (por ejemplo, en logaritmos) o en diferencias logarítmicas (o diferencias sin logaritmos).

3.6.1 ARMA(1, 1)

Dicho lo anterior, vamos a empezar con el análisis de un $ARMA(1, 1)$. Un proceso $ARMA(1, 1)$ puede verse como:

$$X_t = \delta + a_1 X_{t-1} + U_t - b_1 U_{t-1} \quad (3.78)$$

Aplicando el operado rezago podemos rescribir la ecuación (3.78) como:

$$(1 - a_1 L)X_t = \delta + (1 - b_1 L)U_t \quad (3.79)$$

Donde U_t es un proceso puramente aleatorio como en los casos de $AR(p)$ y $MA(q)$, y X_t puede ser una serie en niveles o en diferencias (ambas, en términos logarítmicos).

Así, el modelo $ARIMA(p, q)$ también tiene una representación de Wold que estará dada por las siguientes expresiones:

$$X_t = \frac{\delta}{1 - a_1} + \frac{1 - b_1 L}{1 - a_1 L} U_t \quad (3.80)$$

Donde $a_1 \neq b_1$, puesto que en caso contrario X_t sería un proceso puramente aleatorio con una media $\mu = \frac{\delta}{1 - a_1}$. Así, podemos reescribir la descomposición de Wold a partir del componente de la ecuación (3.80):

$$\frac{1 - b_1 L}{1 - a_1 L} = \psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots \quad (3.81)$$

Esta ecuación es equivalente a la expresión:

$$\begin{aligned} (1 - b_1 L) &= (1 - a_1 L)(\psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots) \\ &= \psi_0 + \psi_1 L + \psi_2 L^2 + \psi_3 L^3 + \dots \\ &\quad - a_1 \psi_0 L - a_1 \psi_1 L^2 - a_2 \psi_2 L^3 - a_1 \psi_3 L^4 - \dots \end{aligned}$$

De esta forma podemos establecer el siguiente sistema de coeficientes indeterminados:

$$\begin{aligned} L^0 &: \Rightarrow \psi_0 = 1 \\ L^1 &: \psi_1 - a_1 \psi_0 = -b_1 \Rightarrow \psi_1 = a_1 - b_1 \\ L^2 &: \psi_2 - a_1 \psi_1 = 0 \Rightarrow \psi_2 = a_1(a_1 - b_1) \\ L^3 &: \psi_3 - a_1 \psi_2 = 0 \Rightarrow \psi_3 = a_1^2(a_1 - b_1) \\ &\vdots \\ L^j &: \psi_j - a_1 \psi_{j-1} = 0 \Rightarrow \psi_j = a_1^{j-1}(a_1 - b_1) \end{aligned}$$

Así, la solución a la ecuación (3.78) estará dada por la siguiente generalización:

$$X_t = \frac{\delta}{1 - a_1} + U_t + (a_1 - b_1)U_{t-1} + a_1(a_1 - b_1)U_{t-2} + a_1^2(a_1 - b_1)U_{t-3} + \dots \quad (3.82)$$

En la ecuación (3.82) las condiciones de estabilidad y de invertibilidad del sistema (de un MA a un AR, y viceversa) estarán dadas por: $|a_1| < 1$ y $|b_1| < 1$. Adicionalmente, la ecuación (3.82) expresa cómo una serie que tiene un comportamiento $ARMA(1, 1)$ es equivalente a una serie modelada bajo un $MA(\infty)$.

Al igual que en los demás modelos, ahora determinaremos los momentos del proceso $ARMA(1, 1)$. La media estará dada por:

$$\begin{aligned}
 \mathbb{E}[X_t] &= \mathbb{E}[\delta + a_1 X_{t-1} + U_t - b_1 U_{t-1}] \\
 &= \delta + a_1 \mathbb{E}[X_{t-1}] \\
 &= \frac{\delta}{1 - a_1} \\
 &= \mu
 \end{aligned} \tag{3.83}$$

Donde hemos utilizado que $\mathbb{E}[X_t] = \mathbb{E}[X_{t-1}] = \mu$. Es decir, la media de un $ARMA(1, 1)$ es idéntica a la de un $AR(1)$.

Para determinar la varianza tomaremos una estrategia similar a los casos de $AR(p)$ y $MA(q)$. Por lo que para todo $\tau \geq 0$, y suponiendo por simplicidad que $\delta = 0$ (lo que implica que $\mu = 0$) tendremos:

$$\begin{aligned}
 \mathbb{E}[X_{t-\tau} X_t] &= \mathbb{E}[(X_{t-\tau}) \cdot (a_1 X_{t-1} + U_t - b_1 U_{t-1})] \\
 &= a_1 \mathbb{E}[X_{t-\tau} X_{t-1}] + \mathbb{E}[X_{t-\tau} U_t] - b_1 \mathbb{E}[X_{t-\tau} U_{t-1}]
 \end{aligned} \tag{3.84}$$

De la ecuación (3.84) podemos determinar una expresión para el caso de $\tau = 0$:

$$\begin{aligned}
 \mathbb{E}[X_t X_t] &= \gamma(0) \\
 &= a_1 \gamma(1) + \mathbb{E}[U_t X_t] - b_1 \mathbb{E}[X_t U_{t-1}] \\
 &= a_1 \gamma(1) + \sigma^2 + b_1 \mathbb{E}[U_{t-1} (a_1 X_{t-1} + U_t - b_1 U_{t-1})] \\
 &= a_1 \gamma(1) + \sigma^2 - b_1 a_1 \sigma^2 + b_1 \sigma^2 \\
 &= a_1 \gamma(1) + (1 - b_1(a_1 - b_1)) \sigma^2
 \end{aligned} \tag{3.85}$$

Para el caso en que $\tau = 1$:

$$\begin{aligned}
 \mathbb{E}[X_{t-1} X_t] &= \gamma(1) \\
 &= a_1 \gamma(0) + \mathbb{E}[X_{t-1} U_t] - b_1 \mathbb{E}[X_{t-1} U_{t-1}] \\
 &= a_1 \gamma(0) - b_1 \sigma^2
 \end{aligned} \tag{3.86}$$

Estas últimas expresiones podemos resolverlas como sistema para determinar los siguientes valores:

$$\gamma(0) = \frac{1 + b_1^2 - 2a_1 b_1}{1 - a_1^2} \sigma^2 \tag{3.87}$$

$$\gamma(1) = \frac{(a_1 - b_1)(1 - a_1 b_1)}{1 - a_1^2} \sigma^2 \tag{3.88}$$

En general, para cualquier valor $\tau \geq 2$ tenemos que la autocovarianza y la función de autocorrelación serán:

$$\gamma(\tau) = a_1 \gamma(\tau - 1) \tag{3.89}$$

$$\rho(\tau) = a_1 \rho(\tau - 1) \tag{3.90}$$

Por ejemplo, para el caso de $\tau = 1$ tendríamos:

$$\rho(1) = \frac{(a_1 - b_1)(1 - a_1 b_1)}{1 + b_1^2 - 2a_1 b_1} \quad (3.91)$$

De esta forma, la función de autocorrelación oscilará en razón de los valores que tome a_1 y b_1 .

3.6.2 ARMA(p, q)

La especificación general de un $ARMA(p, q)$ (donde $p, q \in \mathbb{N}$) puede ser descrita por la siguiente ecuación:

$$\begin{aligned} X_t = & \delta + a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} \\ & + U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q} \end{aligned} \quad (3.92)$$

Donde U_t es un proceso puramente aleatorio, y X_t puede ser modelada en niveles o en diferencias (ya sea en logaritmos o sin transformación logarítmica).

Mediante el uso del operador rezago se puede escribir la ecuación (3.92) como:

$$(1 - a_1 L - a_2 L^2 - \dots - a_p L^p) X_t = \delta + (1 - b_1 L - b_2 L^2 - \dots - b_q L^q) U_t \quad (3.93)$$

En la ecuación (3.93) definamos dos polinomios: $\alpha(L) = (1 - a_1 L - a_2 L^2 - \dots - a_p L^p)$ y $\beta(L) = (1 - b_1 L - b_2 L^2 - \dots - b_q L^q)$. Así, podemos reescribir la ecuación (3.93) como:

$$\alpha(L) X_t = \delta + \beta(L) U_t \quad (3.94)$$

Asumiendo que existe el polinomio inverso tal que: $\alpha(L)^{-1} \alpha(L) = 1$. La solución entonces puede ser escrita como:

$$\begin{aligned} X_t &= \alpha(L)^{-1} \delta + \alpha(L)^{-1} \beta(L) U_t \\ &= \frac{\delta}{1 - a_1 - a_2 - \dots - a_p} + \frac{\beta(L)}{\alpha(L)} U_t \\ &= \frac{\delta}{1 - a_1 - a_2 - \dots - a_p} + U_t + \psi_1 L U_t + \psi_2 L^2 U_t + \dots \end{aligned} \quad (3.95)$$

Donde la ecuación (3.95) nos permite interpretar que un $ARMA(p, q)$ se puede reexpresar e interpretar como un $MA(\infty)$ y donde las condiciones para la estabilidad de la solución y la invertibilidad son que las raíces de los polinomios característicos $\alpha(L)$ y $\beta(L)$ son en valor absoluto menores a 1.

Adicionalmente, la fracción en la ecuación (3.95) se puede descomponer como en la forma de Wold:

$$\frac{\beta(L)}{\alpha(L)} = 1 + \psi_1 L + \psi_2 L^2 + \dots \quad (3.96)$$

Bajo los supuestos de estacionariedad del componente U_t , los valores de la media y varianza de un proceso $ARMA(p, q)$ serán como describimos ahora. Para el caso de la media, podemos partir de la ecuación (3.95) para generar:

$$\begin{aligned}\mathbb{E}[X_t] &= \mathbb{E}\left[\frac{\delta}{1 - a_1 - a_2 - \dots - a_p} + U_t + \psi_1 U_{t-1} + \psi_2 U_{t-2} + \dots\right] \\ &= \frac{\delta}{1 - a_1 - a_2 - \dots - a_p} \\ &= \mu\end{aligned}\tag{3.97}$$

Esta expresión indica que en general un proceso $ARMA(p, q)$ converge a una media idéntica a la de un proceso $AR(p)$. Para determinar la varianza utilizaremos la misma estrategia que hemos utilizado para otros modelos $AR(p)$ y $MA(q)$.

Sin pérdida de generalidad podemos asumir que $\delta = 0$, lo que implica que $\mu = 0$, de lo que podemos establecer una expresión de autocovarianzas para cualquier valor $\tau = 0, 1, 2, \dots$:

$$\begin{aligned}\gamma(\tau) &= \mathbb{E}[X_{t-\tau}X_t] \\ &= \mathbb{E}[X_{t-\tau}(\delta + a_1 X_{t-1} + a_2 X_{t-2} + \dots + a_p X_{t-p} \\ &\quad + U_t - b_1 U_{t-1} - b_2 U_{t-2} - \dots - b_q U_{t-q})] \\ &= a_1 \gamma(\tau - 1) + a_2 \gamma(\tau - 2) + \dots + a_p \gamma(\tau - p) \\ &\quad + \mathbb{E}[X_{t-\tau}U_t] - b_1 \mathbb{E}[X_{t-\tau}U_{t-1}] - \dots - b_q \mathbb{E}[X_{t-\tau}U_{t-q}]\end{aligned}\tag{3.98}$$

Ahora veámos un ejemplo. Utilizaremos la serie de Pasajeros en vuelos nacionales de salida para estimar un $ARMA(p, q)$ mediante el método de máxima verosimilitud (ML). Antes de realizar el proceso de estimación, consideremos una transformación de la serie en diferencias logarítmicas, ya que, según la gráfica en la Figura 3.10, esa es la que puede ser estacionaria.

A continuación, estimaremos una $ARMA(1, 1)$ para la serie en diferencias logarítmicas ($DLPaxNal_t$). También incorporaremos al análisis variables exógenas tales como dummies de estacionalidad. En particular, utilizaremos los meses de enero, febrero, julio y diciembre. No debe pasar desapercibido que un análisis de estacionalidad más formal debería considerar todos los meses para separar del término de error la parte que puede ser explicada por los ciclos estacionales.

Así obtenemos el siguiente resultado:

Table 3.8: ARMA(1, 1) para la variable $DLPaxNal_t$.

$DLPaxNal_t$	=	-0.0025	—	0.6018	$DLPaxNal_{t-1}$
		(0.0045)		(0.0964)	
	—	0.9064	U_{t-1}		

	(0.0397)			
—	0.0867	DNe_t		
	(0.0230)			
—	0.0409	$DFeb_t$		
	(0.0263)			
+	0.1529	$DJul_t$		
	(0.0245)			
+	0.0628	$DDic_t$		
	(0.0223)			
$\hat{\sigma}^2$	= 0.007096	AIC	=	—475.12

Donde entre paréntesis indicamos los errores estándar. Adicionalmente, reportamos el estadístico de Akaike (AIC). Finalmente, podemos determinar si las soluciones serán convergentes, para ello en la Figura 3.15 mostramos las raíces asociadas a cada uno de los polinomios. De la inspección visual, podemos concluir que tenemos una solución convergente y estable. Por su parte, la Figura 3.16 muestra los residuales de la estimación del $ARMA(1,1)$.

```
knitr::include_graphics("Plots/G_Roots_ARMA11.png")
```

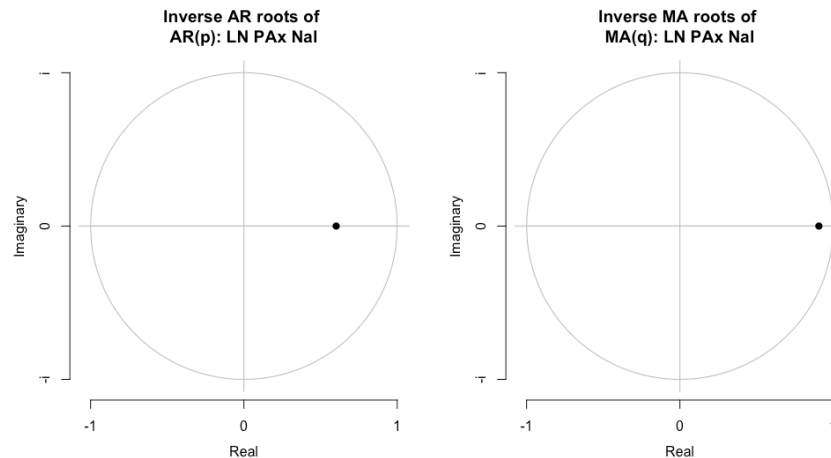


Figure 3.15: Inverso de las Raíces del polinomio característico de un $ARMA(1,1)$

```
knitr::include_graphics("Plots/G_Residuals_ARMA11.png")
```

En lo que resta de este capítulo, utilizaremos la serie en diferencias logarítmicas de los pasajeros en vuelos nacionales de salida, $DLPaxNaI_t$, para discutir

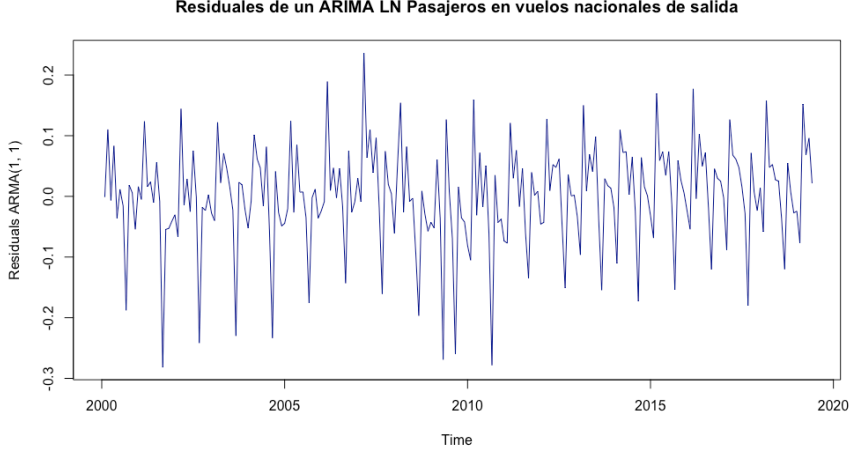


Figure 3.16: Residuales de un $ARMA(1,1)$ de la serie $DLPaxNal_t$

los ejemplos que ilustran cada uno de los puntos teóricos que a continuación exponemos.

3.7 Función de Autocorrelación Parcial

Ahora introduciremos el concepto de Función de Autocorrelación Parcial (PACF, por sus siglas en inglés). Primero, dadas las condiciones de estabilidad y de convergencia, si suponemos que los procesos AR, MA, ARMA o ARIMA tienen toda la información de los rezagos de la serie en conjunto y toda la información de los promedios móviles del término de error, resulta importante construir una métrica para distinguir el efecto de $X_{t-\tau}$ o el efecto de $U_{t-\tau}$ (para cualquier τ) sobre X_t de forma individual.

La idea es construir una métrica de la correlación que existe entre las diferentes variables aleatorias, si para tal efecto se ha controlado el efecto del resto de la información. Así, podemos definir la ecuación que puede responder a este planteamiento como:

$$X_t = \phi_{k1}X_{t-1} + \phi_{k2}X_{t-2} + \dots + \phi_{kk}X_{t-k} + U_t \quad (3.99)$$

Donde ϕ_{ki} es el coeficiente de la variable dada con el rezago i si el proceso tiene un orden k . Así, los coeficientes ϕ_{kk} son los coeficientes de la autocorrelación parcial (considerando un proceso $AR(k)$). Observemos que la autocorrelación parcial mide la correlación entre X_t y X_{t-k} que se mantiene cuando el efecto de las variables X_{t-1} , X_{t-2} , ... y X_{t-k-1} en X_t y X_{t-k} ha sido eliminado.

Dada la expresión considerada en la ecuación (3.99), podemos resolver el problema de establecer el valor de cada ϕ_{ki} mediante la solución del sistema que se representa en lo siguiente:

$$\begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(k) \end{bmatrix} = \begin{bmatrix} 1 & \rho(1) & \dots & \rho(k-1) \\ \rho(1) & 1 & \dots & \rho(k-2) \\ \rho(2) & \rho(1) & \dots & \rho(k-3) \\ \vdots & \vdots & \dots & \vdots \\ \rho(k-1) & \rho(k-2) & \dots & 1 \end{bmatrix} \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \vdots \\ \phi_{kk} \end{bmatrix} \quad (3.100)$$

Del cual se puede derivar una solución, resolviendo por cualquier método que consideremos y que permita calcular la solución de sistemas de ecuaciones.

Posterior al planteamiento analítico, plantearemos un enfoque para interpretar las funciones de autocorrelación y autocorrelación parcial. Este enfoque pretende aportar al principio de parsimonia, en el cual podemos identificar el número de parámetros que posiblemente puede describir mejor a la serie en un modelo ARMA(p, q).

En el siguiente cuadro mostramos un resumen de las características que debemos observar para determinar el número de parámetros de cada uno de los componentes AR y MA. Lo anterior por observación de las funciones de autocorrelación y autocorrelación parcial. Este enfoque no es el más formal, más adelante implementaremos uno más formal y que puede ser más claro de cómo determinar el número de parámetros.

Table 3.9: Relación entre la Función de Autocorrelación y la Función de Autocorrelación Parcial de una serie X_t .

Modelo	Función de Autocorrelación	Función de Autocorrelación Parcial
MA(q)	Rompimientos en la función	No hay rompimientos
AR(p)	No hay rompimientos	Rompimientos en la función

Ejemplo. Continuando con el ejemplo en la Figura 3.17 mostramos tanto la Función de Autocorrelación como la Función de Autocorrelación Parcial. En esta identificamos que ambas gráficas muestran que el modelo que explica a la variable $DLPaxNal_t$ tiene tanto componentes AR como MA. Sin embargo, dado lo errático del comportamiento de ambas funciones, resulta complicado determinar cuál sería un buen número de parámetros p y q a considerar en el $ARMA(p, q)$. Por esta razón, a continuación, plantearemos algunas pruebas más formales para determinar dichos parámetros.

3.8. SELECCIÓN DE LAS CONSTANTES P , Q , D EN UN $AR(P)$, UN $MA(Q)$, UN $ARMA(P, Q)$ O UN $ARIMA(P, D, Q)$

```
knitr::include_graphics("Plots/G_ACF_PACF.png")
```

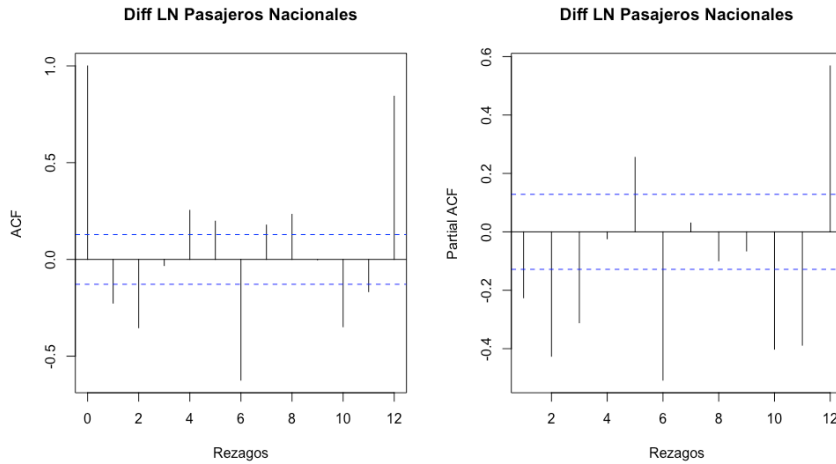


Figure 3.17: Función de Autocorrelación y la Función de Autocorrelación Parcial de una serie $DLPaxNal_t$

3.8 Selección de las constantes p , q , d en un $AR(p)$, un $MA(q)$, un $ARMA(p, q)$ o un $ARIMA(p, d, q)$

Respecto de cómo estimar un proceso $ARMA(p, q)$ —en general utilizaremos este modelo para discutir, pero lo planteado en esta sección es igualmente aplicable en cualquier otro caso como aquellos modelos que incluyen variables exógenas—existen diversas formas de estimar los parámetros a_i y b_i : i) por máxima verosimilitud y ii) por mínimos cuadrados ordinarios. El primer caso requiere que conozcamos la distribución del proceso aleatorio U_t . El segundo, por el contrario, no requiere el mismo supuesto. No obstante, para el curso utilizaremos el método de máxima verosimilitud.

Otra duda que debe quedar hasta el momento es ¿cómo determinar el orden p y q del proceso $ARMA(p, q)$? La manera más convencional y formal que existe para tal efecto es utilizar los criterios de información. Así, el orden se elige de acuerdo a aquel criterio de información que resulta ser el mínimo. En el caso de d se selecciona revisando la gráfica que parezca más estacionaria—más adelante mostraremos un proceso más formal para su selección.

Los criterios de información más comunes son los siguientes:

1. FPE (Final Prediction Error):

$$FPE = \frac{T+m}{T-m} \frac{1}{T} \sum_{t=1}^T \left(\hat{U}_t^{(p)} \right)^2 \quad (3.101)$$

2. Akaike:

$$AIC = \ln \left[\frac{1}{T} \sum_{t=1}^T \left(\hat{U}_t^{(p)} \right)^2 \right] + m \frac{2}{T} \quad (3.102)$$

3. Schwarz:

$$SC = \ln \left[\frac{1}{T} \sum_{t=1}^T \left(\hat{U}_t^{(p)} \right)^2 \right] + m \frac{\ln(T)}{T} \quad (3.103)$$

4. Hannan - Quinn:

$$HQ = \ln \left[\frac{1}{T} \sum_{t=1}^T \left(\hat{U}_t^{(p)} \right)^2 \right] + m \frac{2 \ln(\ln(T))}{T} \quad (3.104)$$

Donde $\hat{U}_t^{(p)}$ son los residuales estimados mediante un proceso ARIMA y m es el número de parámetros estimados: $m = p + q + 0 + 1$ (ya que asumimos que $d = 0$). Una propiedad que no se debe perder de vista es que los criterios de información cumplen la siguiente relación:

$$\text{orden}(SC) \leq \text{orden}(HQ) \leq \text{orden}(AIC) \quad (3.105)$$

Por esta razón, durante el curso solo utilizaremos el criterio de Akaike para determinar el orden óptimo del proceso ARMA, ya que ello garantiza el orden más grande posible.

Ejemplo. Ahora veamos un ejemplo de estimación del número de rezagos óptimo de un $ARMA(p, q)$. Retomemos la serie en diferencias logarítmicas de los pasajeros en vuelos nacionales de salidas, pero ahora incluiremos las variables exógenas de dummies estacionales: enero, febrero, julio y diciembre.

Como mencionamos, las gráficas de las funciones de autocorrelación permiten observar el valor de la correlación existente entre la variable en el momento t con cada uno de los rezagos. Incluso la Función de Autocorrelación Parcial puede ayudar a determinar el número máximo de rezagos que se debe incluir en el proceso $AR(p)$. No obstante, una métrica más formal es el uso de los criterios de información. En nuestro caso, dado lo discutido, sólo utilizaremos el criterio de Akaike.

Al respecto, en el siguiente cuadro reportamos el criterio de Akaike que resulta de aplicar dicho criterio a los residuales resultantes de cada combinación de procesos $ARMA(p, q)$. La forma de escoger será aquel modelo que reporta el

3.8. SELECCIÓN DE LAS CONSTANTES P , Q , D EN UN $AR(P)$, UN $MA(Q)$, UN $ARMA(P, Q)$ O UN $ARIMA(P, D, Q)$

criterio de Akaike menor. En la cuarta columna de la tabla se señala el valor del criterio de información que resulta ser el mínimo de todos los posibles.

Table 3.10: Criterio de Akaike para diferentes modelos $ARMA(p, q)$ de la serie $DLPaxNal_t$.

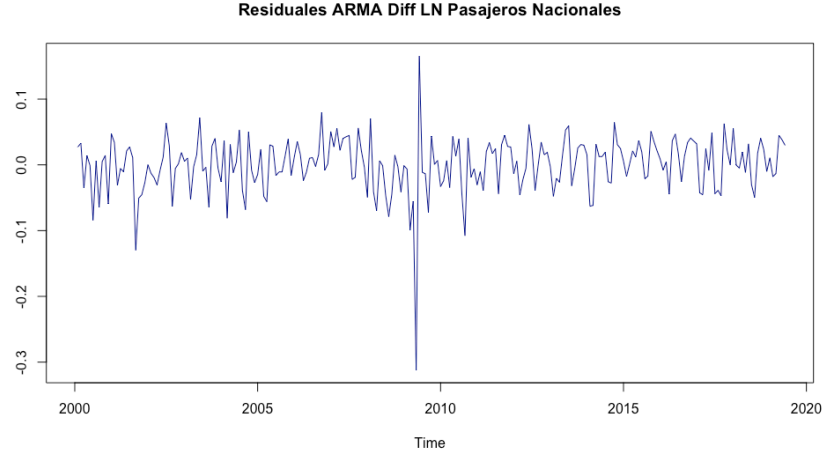
Modelo	AR(p)	MA(p)	Akaike (AIC)	Óptimo
1	1	1	-475.1190	
2	1	2	-473.4397	
3	1	3	-483.1239	
4	1	4	-482.4932	
5	1	5	-506.9880	
6	1	6	-533.9411	
7	2	1	-473.4758	
⋮	⋮	⋮	⋮	⋮
24	4	6	-752.8996	*
⋮	⋮	⋮	⋮	⋮
36	6	6	-752.2632	

El cuadro anterior reporta un resumen de los resultados para 36 diferentes modelos, todos incluyen variables exógenas. Como resultado del análisis concluimos que el modelo más adecuado es el 24, el cual considera un $ARMA(4, 6)$, con variables dummies para controlar la estacionalidad de los meses de enero, febrero, julio y diciembre. Más adelante mostramos los resultados del modelo.

No obstante, una inspección de los residuales del modelo nos permite sospechar que requiere de incluir un par de dummies más. Ambas, asociadas con la caída del transporte aéreo en 2009, principalmente asociadas con la crisis mundial de ese año. La Figura 3.18 muestra los residuales mencionados.

```
knitr::include_graphics("Plots/G_Residuals_ARMA46.png")
```

Una vez incluidas dos dummies más para mayo y junio de 2009, analizamos un total de 36 modelos ARMA y determinamos que el orden que minimiza el criterio de Akaike es un $ARMA(4, 6)$. El siguiente cuadro muestra los resultados para este nuevo modelo. No lo mostramos en esta sección, pero ambos modelos reportados tienen raíces de sus respectivos polinomios característicos menores a 1 en valor absoluto. En la Figura 3.19 mostramos los residuales ahora ajustados por las dummies de mayo y junio de 2009.

Figure 3.18: Residuales del ARMA(4, 6) de una serie $DLPaxNal_t$ Table 3.11: Modelo $ARMA(p, q)$ de la serie $DLPaxNal_t$.

Variable	Coeficiente	Error Est.	Coeficiente	Error Est.
	Modelo 1	Modelo 2		
Constante	0.0134	0.0033	0.0145	0.0029
$DLPaxNal_{t-1}$	0.0004	0.0011	0.0014	0.0016
$DLPaxNal_{t-2}$	0.9997	0.0019	0.9988	0.0024
$DLPaxNal_{t-3}$	0.0001	0.0012	0.0002	0.0014
$DLPaxNal_{t-4}$	-0.9997	0.0004	-0.9993	0.0009
\hat{U}_{t-1}	-0.2877	0.0670	-0.1472	0.0633
\hat{U}_{t-2}	-1.2028	0.0608	-1.3323	0.0625
\hat{U}_{t-3}	0.2302	0.0765	0.1201	0.0756
\hat{U}_{t-4}	1.2085	0.0645	1.3142	0.0642
\hat{U}_{t-5}	-0.2634	0.0708	-0.1034	0.0717
\hat{U}_{t-6}	-0.2171	0.0606	-0.3864	0.0657
$DEne_t$	-0.2904	0.0172	-0.2865	0.0156
$DFeb_t$	-0.1312	0.0170	-0.1273	0.0140
$DJul_t$	0.3303	0.0173	0.3164	0.0156
Dic_t	-0.0118	0.0169	-0.0138	0.0137
$DMay2009_t$			-0.3378	0.0360
$DJun2009_t$			0.2371	0.0359
$\hat{\sigma}^2$	0.001841		0.001283	
AIC	-752.9		-835.73	


```
knitr::include_graphics("Plots/G_Residuals_ARMA46_D.png")
```

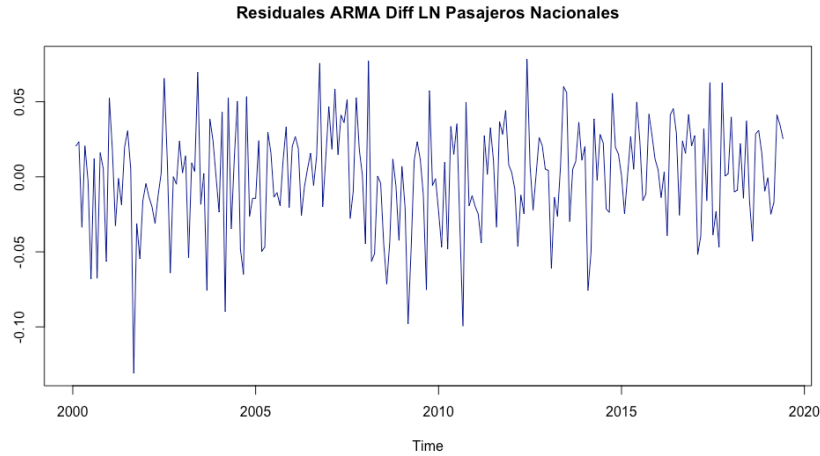


Figure 3.19: Residuales del ARMA(4, 6) de una serie $DLPaxNal_t$

3.9 Pronósticos

Para pronósticar el valor de la serie es necesario determinar cuál es el valor esperado de la serie en un momento $t + \tau$ condicional en que ésta se comporta como un $AR(p)$, un $MA(q)$ o un $ARMA(p, q)$ y a que los valores antes de t están dados. Por lo que el pronóstico de la serie estará dado por una expresión:

$$\mathbb{E}_t[X_{t+\tau}] = \delta + a_1 \mathbb{E}_t[X_{t+\tau-1}] + a_2 \mathbb{E}_t[X_{t+\tau-2}] + \dots + a_p \mathbb{E}_t[X_{t+\tau-p}] \quad (3.106)$$

Lo anterior para todo $\tau = 0, 1, 2, \dots$ y considerando que los componentes $MA(q)$ en la ecuación (3.106) son cero dado que para todo valor $t + \tau$ es cierto que $\mathbb{E}_t[U_{t+\tau}] = 0$.

Continuando con el ejemplo, en la Figura 3.20 mostramos el resultado del pronóstico de la serie a partir del modelo ARMA(4, 6) que hemos discutido anteriormente.

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)
```

```

Datos <- read_excel("BD/Base_Transporte_ARIMA.xlsx",
                    sheet = "Datos", col_names = TRUE)

Pax_Nal <- ts(Datos$Pax_Nal,
              start = c(2000, 1),
              freq = 12)

LPax_Nal <- ts(log(Datos$Pax_Nal),
               start = c(2000, 1),
               freq = 12)

DLPax_Nal <- ts(log(Datos$Pax_Nal) - lag(log(Datos$Pax_Nal), 1),
               start = c(2000, 1),
               freq = 12)

D_Mar2020 <- ts(Datos$D_Mar2020,
                start = c(2000, 1),
                freq = 12)

D_Abr2020 <- ts(Datos$D_Abr2020,
                start = c(2000, 1),
                freq = 12)

D_Jun2020 <- ts(Datos$D_Jun2020,
                start = c(2000, 1),
                freq = 12)

D_Jul2020 <- ts(Datos$D_Jul2020,
                start = c(2000, 1),
                freq = 12)

D_Mar2021 <- ts(Datos$D_Mar2021,
                start = c(2000, 1),
                freq = 12)

D_Ene <- ts(Datos$D_Ene,
            start = c(2000, 1),
            freq = 12)

D_Feb <- ts(Datos$D_Feb,
            start = c(2000, 1),
            freq = 12)

D_Jul <- ts(Datos$D_Jul,
            start = c(2000, 1),

```

```

        freq = 12)

D_Dic <- ts(Datos$D_Dic,
           start = c(2000, 1),
           freq = 12)

ARMA_Ex_DLPax_Nal_2 <- arima(DLPax_Nal, order = c(6, 0, 6),
                             xreg = cbind(D_Ene, D_Feb, D_Jul,
                                           D_Dic, D_Mar2020,
                                           D_Abr2020, D_Jun2020,
                                           D_Jul2020, D_Mar2021),
                             method = "ML")

Predict_Datos <- read_excel("BD/Predict_Base_Transporte_ARIMA.xlsx",
                           sheet = "Datos", col_names = TRUE)

D_Mar2020_f <- ts(Predict_Datos$D_Mar2020,
                  start = c(2022, 7),
                  freq = 12)

D_Abr2020_f <- ts(Predict_Datos$D_Abr2020,
                  start = c(2022, 7),
                  freq = 12)

D_Jun2020_f <- ts(Predict_Datos$D_Jun2020,
                  start = c(2022, 7),
                  freq = 12)

D_Jul2020_f <- ts(Predict_Datos$D_Jul2020,
                  start = c(2022, 7),
                  freq = 12)

D_Mar2021_f <- ts(Predict_Datos$D_Mar2021,
                  start = c(2022, 7),
                  freq = 12)

D_Ene_f <- ts(Predict_Datos$D_Ene,
              start = c(2022, 7),
              freq = 12)

D_Feb_f <- ts(Predict_Datos$D_Feb,
              start = c(2022, 7),
              freq = 12)

D_Jul_f <- ts(Predict_Datos$D_Jul,

```

```

        start = c(2022, 7),
        freq = 12)

D_Dic_f <- ts(Predict_Datos$D_Dic,
             start = c(2022, 7),
             freq = 12)

DLPax_Nal_f <- predict(ARMA_Ex_DLPax_Nal_2, n.ahead = 24,
                    newxreg = cbind(D_Ene_f, D_Feb_f, D_Jul_f,
                                     D_Dic_f, D_Mar2020_f,
                                     D_Abr2020_f, D_Jun2020_f,
                                     D_Jul2020_f, D_Mar2021_f))

Pronostico_Arima <- read_excel("BD/Pronostico_Arima.xlsx",
                              sheet = "Datos", col_names = TRUE)

Pronostico_Arima$Pax_Nal_f <- Pronostico_Arima$Pax_Nal

for(i in 1:24){
  Pronostico_Arima$Pax_Nal_f[281 + i] <-
    Pronostico_Arima$Pax_Nal_f[281 + i - 1]*(1 + DLPax_Nal_f$pred[i])
}

options(scipen = 999) # NO notacion cientifica

ggplot(data = Pronostico_Arima, aes(x = Periodo)) +
  geom_line(aes(y = Pax_Nal_f, color = "Pax_Nal_f")) +
  geom_line(aes(y = Pax_Nal, color = "Pax_Nal")) +
  scale_color_brewer(type = "qual", palette = 2) +
  #theme_bw() +
  theme(legend.position = "bottom") +
  theme(legend.title = element_blank()) +
  guides(col = guide_legend(nrow = 1, byrow = TRUE)) +
  xlab("Tiempo") +
  ylab("Pasajeros") +
  theme(plot.title = element_text(size = 11, face = "bold",
                                  hjust = 0)) +
  theme(plot.subtitle = element_text(size = 10, hjust = 0)) +
  theme(plot.caption = element_text(size = 10, hjust = 0)) +
  theme(plot.margin = unit(c(1,1,1,1), "cm")) +
  labs(
    title = "Pasajeros en vuelos nacionales (Salidas)",
    subtitle = "(Ene-2000 a Jun-2019)",
    caption = "Fuente: Elaboración propia"
  )

```

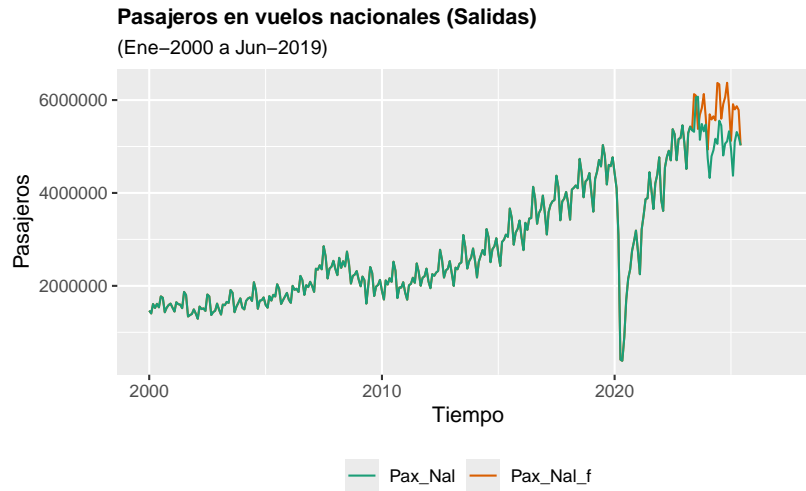


Figure 3.20: Pronóstico de la serie $PaxNAL_t$ a partir de una ARMA(4,6) en diferencias logarítmicas

```
ggsave("Pax_Nal_f.png", width = 20, height = 15, units = "cm")
```

3.10 Desestacionalización y filtrado de Series

3.11 Motivación

Existen múltiples enfoques para la desestacionalización de series. Algunos modelos, por ejemplo, pueden estar basados en modelos ARIMA como un conjunto de dummies. No obstante, en el caso particular que discutiremos en este curso, estará basado en un modelo ARIMA de la serie. Este enfoque está basado en el modelo X11 de la oficina del censo de Estados Unidos (Census Bureau) el cual es conocido como el modelo X13-ARIMA-SEATS.⁶ El modelo X13-ARIMA-SEATS es, como su nombre lo indica, la combinación de un modelo ARIMA con componentes estacionales (por la traducción literal de la palabra: seasonal).

Un modelo ARIMA estacional emplea la serie en diferencias y como regresores

⁶La información y material respecto del modelo esta disponible en la dirección <https://www.census.gov/srd/www/x13as/>

los valores rezagados de las diferencias de la serie tantas veces como procesos estacionales s existan en ésta, con el objeto de remover los efectos aditivos de la estacionalidad. Sólo para entender qué significa este mecanismo, recordemos que cuando se utiliza la primera diferencia de la serie respecto del periodo inmediato anterior se remueve la tendencia. Por su parte, cuando se incluye la diferencia respecto del mes s se está en el caso en que se modela la serie como una media móvil en términos del rezago s .

El modelo ARIMA estacional incluye como componentes autorregresivos y de medias móviles a los valores rezagados de la serie en el periodo s en diferencias. El ARIMA(p, d, q)(P, D, Q) estacional puede ser expresado de la siguiente manera utilizando el operador rezago L :

$$\Theta_P(L^s)\theta_p(L)(1-L^s)^D(1-L)^dX_t = \Psi_Q(L^s)\psi_q(L)U_t(\#eq : ARIMA_{seas}) \quad (3.107)$$

Donde $\Theta_P(\cdot)$, $\theta_p(\cdot)$, $\Psi_Q(\cdot)$ y $\psi_q(\cdot)$ son polinomios de L de orden P , p , Q y q respectivamente. En general, la representación es de una serie no estacionaria, aunque si $D = d = 0$ y las raíces del polinomio característico (de los polinomios del lado izquierdo de la ecuación @ref(eq:ARIMA_seas) todas son más grandes que 1 en valor absoluto, el proceso modelado será estacionario.

Si bien es cierto que existen otras formas de modelar la desestacionalización, como la modelación en diferencias con dummies para identificar ciertos patrones regulares, en los algoritmos disponibles como el X11 o X13-ARIMA-SEATS se emplea la formulación de la ecuación @ref(eq:ARIMA_seas). A continuación, implementaremos la desestacionalización de una serie.

Como ejemplo utilizaremos la serie del Índice Nacional de Precios al Consumidor (INPC).⁷ Podemos ver que la serie original del INPC y su ajuste estacional bajo una metodología X13-ARIMA-SEATS son como se muestra en la Figura 3.21.

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)
library(seasonal)
```

```
##
## Attaching package: 'seasonal'

## The following object is masked from 'package:tibble':
##
##      view
```

⁷Las definiciones de las variables ocupadas es: * INPC: Índice Nacional de Precios al Consumidor (2QJul2018 = 100) * TC: Tipo de Cambio FIX * CETE28: Tasa de rendimiento promedio mensual de los Cetes 28, en por ciento anual * IGAE: Indicador global de la actividad económica (2013 = 100) * IPI: Industrial Production Index (2012 = 100)

```
library(seasonalview)
```

```
##  
## Attaching package: 'seasonalview'  
  
## The following object is masked from 'package:seasonal':  
##  
##      view  
  
## The following object is masked from 'package:tibble':  
##  
##      view
```

```
library(shiny)
```

```
Datos <- read_excel("BD/Base_VAR.xlsx", sheet = "Datos", col_names = TRUE)
```

```
INPC <- ts(Datos$INPC,  
          start = c(2000, 1),  
          freq = 12)
```

```
Seas_INPC <- seas(INPC)
```

```
INPC_Ad <- final(Seas_INPC)
```

```
#names(Seas_INPC)
```

```
#summary(Seas_INPC)
```

```
plot(Seas_INPC)
```

```
# Forma Facil:
```

```
#view(Seas_INPC)
```

```
TC <- ts(Datos$TC,  
        start = c(2000, 1),  
        freq = 12)
```

```
Seas_TC <- seas(TC)
```

```
TC_Ad <- final(Seas_TC)
```

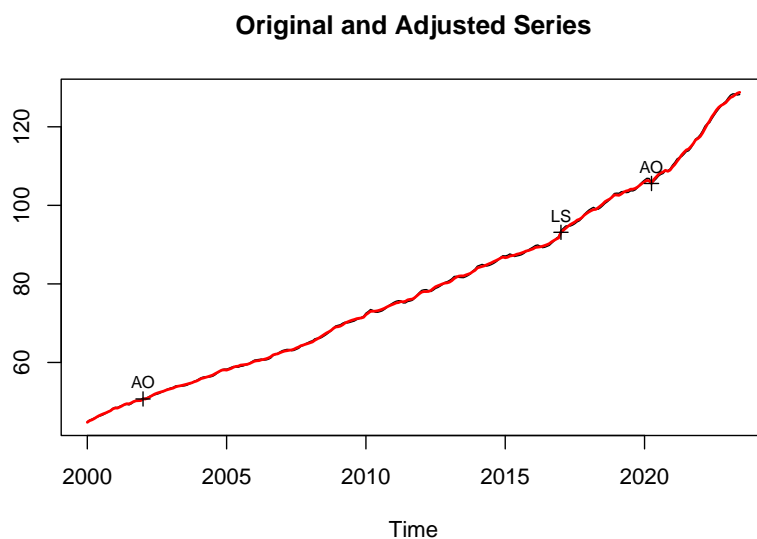


Figure 3.21: Índice Nacional de Precios al Consumidor ($INPC_t$) y su serie desestacionalizada utilizando un proceso X13-ARIMA-SEATS

```
CETE28 <- ts(Datos$CETE28,
             start = c(2000, 1),
             freq = 12)

Seas_CETE28 <- seas(CETE28)

CETE28_Ad <- final(Seas_CETE28)

IGAE <- ts(Datos$IGAE,
           start = c(2000, 1),
           freq = 12)

Seas_IGAE <- seas(IGAE)

IGAE_Ad <- final(Seas_IGAE)

IPI <- ts(Datos$IPI,
         start = c(2000, 1),
         freq = 12)

Seas_IPI <- seas(IPI)
```



```

IPI_Ad <- final(Seas_IPI)

Datos_Ad <- data.frame(cbind(INPC_Ad, TC_Ad, CETE28_Ad, IGAE_Ad, IPI_Ad))

Datos_Ad <- cbind(Datos, Datos_Ad)

save(Datos_Ad, file = "BD/Datos_Ad0.RData")

```

El mismo procesamiento puede ser seguido para todas las series que busquemos analizar. En particular, en adelante, además del INPC que incluimos en la lista, utilizaremos las siguientes series, así como su versión desestacionalizada:

- Índice Nacional de Precios al Consumidor (base 2QJul2018 = 100), $INPC_t$.
- Tipo de Cambio FIX, TC_t
- Tasa de rendimiento promedio mensual de los Cetes 28, en por ciento anual, $CETE28_t$
- Indicador global de la actividad económica (base 2013 = 100), $IGAE_t$
- Industrial Production Index o Índice de Producción Industrial de los Estados Unidos (base 2012 = 100), IPI_t

3.12 Filtro Hodrick-Prescott

Como último tema de los procesos univariados y que no necesariamente aplican a series estacionarias, a continuación desarrollaremos el procedimiento conocido como filtro de Hodrick y Prescott (1997). El trabajo de estos autores era determinar una técnica de regresión que permitiera utilizar series agregadas o macroeconómicas para separarlas en dos componentes: uno de ciclo de negocios y otro de tendencia. En su trabajo original, Hodrick y Prescott (1997) utilizaron datos trimestrales de algunas series como el Gross Domestic Product (GNP, por sus siglas en Inglés), los agregados monetarios M1, empleo, etc., de los Estados Unidos que fueron observados posteriormente a la Segunda Guerra Mundial.

El marco conceptual de Hodrick y Prescott (1997) parte de suponer una serie X_t que se puede descomponer en la suma de componente de crecimiento tendencial, g_t , y su componente de ciclo de negocios, c_t , de esta forma para $t = 1, 2, \dots, T$ tenemos que:

$$X_t = g_t + c_t (\#eq : HP_{Eq}) \quad (3.108)$$

En la ecuación @ref(eq:HP_Eq) se asume que el ajuste de la ruta seguida por g_t es resultado de la suma de los cuadrados de su segunda diferencia. En esa

misma ecuación asumiremos que c_t son las desviaciones de g_t , las cuales en el largo plazo tienen una media igual a cero (0). Por esta razón, se suele decir que el filtro de Hodrick y Prescott representa una descomposición de la serie en su componente de crecimiento natural y de sus desviaciones transitorias que en promedio son cero, en el largo plazo.

Estas consideraciones que hemos mencionado señalan que el procesamiento de Hodrick y Prescott (1997) implican resolver el siguiente problema minimización para determinar cada uno de los componentes en que X_t se puede descomponer:

$$\min_{\{g_t\}_{t=1}^T} \left[\sum_{t=1}^T c_t^2 + \lambda \sum_{t=1}^T [\Delta g_t - \Delta g_{t-1}]^2 \right] \quad (3.109)$$

Donde $\Delta g_t = g_t - g_{t-1}$ y $\Delta g_{t-1} = g_{t-1} - g_{t-2}$; $c_t = X_t - g_t$, y el parámetro λ es un número positivo que penaliza la variabilidad en el crecimiento de las series. De acuerdo con el trabajo de Hodrick y Prescott (1997) la constante λ debe tomar valores específicos de acuerdo con la periodicidad de las series. Así, λ será:

- 100 si la serie es de datos anuales
- 1,600 si la serie es de datos trimestrales
- 14,400 si la serie es de datos mensuales

En resumen, podemos decir que el filtro de Hodrick y Prescott (1997) es un algoritmo que minimiza las distancias o variaciones de la trayectoria de largo plazo. De esta forma, determina una trayectoria estable de largo plazo, por lo que las desviaciones respecto de esta trayectoria serán componentes de ciclos de negocio o cambios transitorios (tanto positivos como negativos).

A continuación, ilustraremos el filtro de Hodrick y Prescott (1997) para dos series desestacionalizadas: $INPC_t$ y TC_t . Las Figura 3.22 y Figura 3.23 muestran los resultados de la implementación del filtro.

```
library(ggplot2)
library(dplyr)
library(readxl)
library(stats)
library(mFilter)
library(plm)

##
## Attaching package: 'plm'
```

```
## The following objects are masked from 'package:dplyr':
##
##   between, lag, lead
```

```
load("BD/Datos_Ad0.RData")

INPC <- ts(Datos_Ad$INPC_Ad, start = c(2000, 1), freq = 12)
TC <- ts(Datos_Ad$TC_Ad, start = c(2000, 1), freq = 12)

INPC_hpf <- hpfilter(INPC, freq = 14400)

#plot(INPC_hpf)

knitr::include_graphics("Plots/G_INPC_HP.png")
```

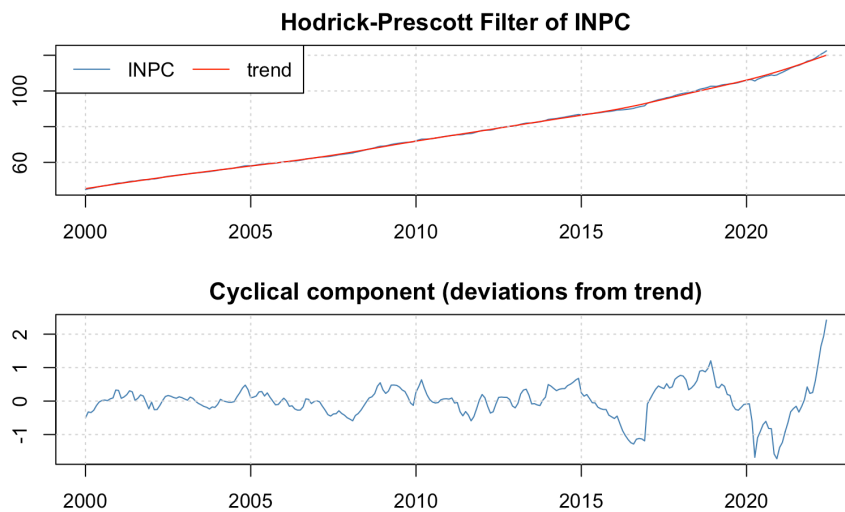


Figure 3.22: Descomposición del Índice Nacional de Precios al Consumidor ($INPC_t$) en su tendencia o trayectoria de largo plazo y su ciclo de corto plazo utilizando un filtro Hodrick y Prescott (1997)

```
TC_hpf <- hpfilter(TC, freq = 14400)

#plot(TC_hpf)

knitr::include_graphics("Plots/G_TC_HP.png")
```

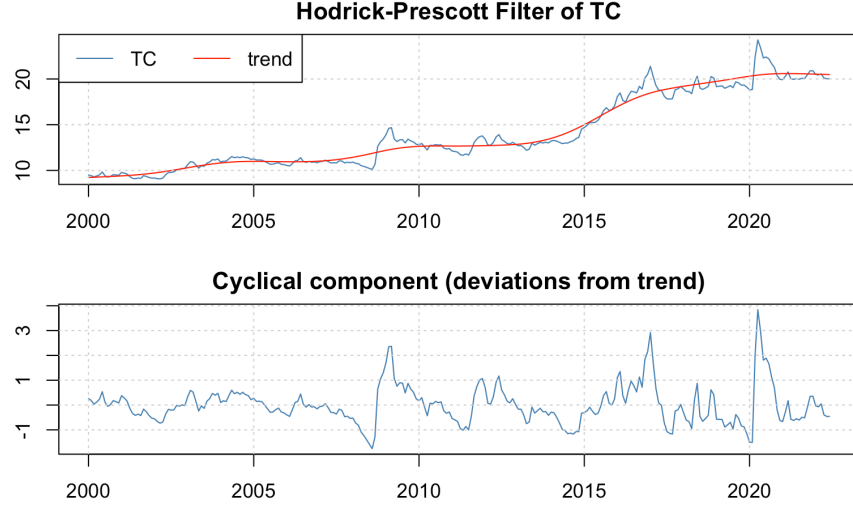


Figure 3.23: Descomposición del Tipo de Cambio FIX (TC_t) en su tendencia o trayectoria de largo plazo y su ciclo de corto plazo utilizando un filtro Hodrick y Prescott (1997)

3.12.1 Filtro Hodrick-Prescott planteado por St-Amant & van Norden

Existen una técnica adicional que soluciona un problema de distribuciones de colas pesadas es el planteado por St-Amant & van Norden. Analicemos ese caso. Método tradicional de HP consiste en minimizar la serie $\{\tau_t\}_{t=-1}^T$:

$$\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=1}^T [(\tau_t - \tau_{t-1}) - (\tau_{t-1} - \tau_{t-2})]^2 \quad (3.110)$$

Donde λ es un parámetro fijo (determinado ex-ante) y τ_t es un componente de tendencia de y_t .

Sin pérdida de generalidad, asumiremos que τ_{-1} y τ_0 son cero (0). De esta manera, la forma matricial del filtro HP es:

$$(Y - G)'(Y - G) + \lambda G' K' K G \quad (3.111)$$

La derivada de los anteriores:

$$-2Y + 2G + \lambda 2K' K G = 0 \quad (3.112)$$

Despejando:

$$G_{hp} = [I_T + \lambda K' K]^{-1} Y \quad (3.113)$$

Donde G es el vector de tendencia, Y es el vector de la serie de datos, λ es la constante tradicional, y K es de dimensión $T \times T$ y está dada por la expresión:

$$K = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots & 0 \\ -2 & 1 & 0 & 0 & \dots & 0 \\ 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (3.114)$$

Así:

$$K' = \begin{pmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ 0 & 0 & 1 & -2 & \dots & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (3.115)$$

Método modificado de HP consiste en minimizar los valores de la serie $\{\tau_t\}_{t=1}^T$:

$$\sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2 + \lambda_{ss} \sum_{t=T-j}^T [\Delta \tau_t - u_{ss}] \quad (3.116)$$

Donde λ es un parámetro fijo (determinado ex-ante), τ_t es un componente de tendencia de y_t , y los nuevos parámetros son u_{ss} y λ_{ss} ajustadas por el procedimiento de Marcet y Ravn (2004).

Este procedimiento asume que parte del filtro HP y que esta versión tiene el problema de pérdida de información al final y al principio de la muestra. La razón es que es un procedimiento univariado que requiere de mucha información futura y pasada para mejorar el ajuste.

El componente adicional al filtro HP es un componente de castigo por desviaciones de la tasa de crecimiento de largo plazo, u_{ss} .

El proceso de selección de λ_{ss} es propuesto por Marcet y Ravn (2004), el cual consiste en utilizar un λ convencional y el filtro HP convencional para estimar la siguiente función:

$$F(\lambda) = \frac{\sum_{t=2}^{T-1} ((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2}{\sum_{t=1}^T (y_t - \tau_t)^2} \quad (3.117)$$

Entonces el valor de λ_{ss} será aquel que:

$$F(\lambda_{ss}) = \frac{\sum_{t=2}^{T-1} ((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2}{\sum_{t=1}^T (y_t - \tau_t)^2} = F(\lambda) \quad (3.118)$$

Nota: Antón (2009) estimó $\lambda_{ss} = 1096$ para datos trimestrales del PIB.

La forma matricial del filtro HP-SAVN es:

$$(Y - G)'(Y - G) + \lambda G' K' K G + \lambda_{ss}(L^j G + \bar{u}_{ss} M^j) \quad (3.119)$$

Donde $L^j = (0, 0, \dots, 0, -1, 0, \dots, 0, 1)$, en el cual el valor -1 es en la posición $T - j - 1$ -ésima, y M^j es un vector que toma valores de cero hasta antes de $T - j$ y de 1 después.

La derivada de los anteriores:

$$-2Y + 2G + \lambda 2K' K G + \lambda_{ss} L'^j = 0 \quad (3.120)$$

Despejando:

$$G_{SAVN} = \frac{1}{2} [I_T + \lambda K' K]^{-1} (2Y - \lambda_{ss} L'^j) \quad (3.121)$$

$$G_{SAVN} = [I_T + \lambda K' K]^{-1} Y - \frac{1}{2} [I_T + \lambda K' K]^{-1} \lambda_{ss} L'^j \quad (3.122)$$

Donde G es el vector de tendencia, Y es el vector de la serie de datos, λ es la constante tradicional, y K es de dimensión $T \times T$ y está dada por la expresión:

$$K' = \begin{pmatrix} 1 & -2 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 1 \end{pmatrix} \quad (3.123)$$

Dicho lo anterior, podemos modificar $F(\lambda)$ para el filtro HP convencional como en forma matricial:

$$F(\lambda) = \frac{G' K' K G}{(Y - G)'(Y - G)} \quad (3.124)$$

Chapter 4

Procesos No Estacionarios Univariados

4.1 Definición y formas de No Estacionariedad

Hasta ahora hemos planteado una serie de técnicas de regresión (AR, MA, ARMA, ARIMA, ARIMAX, etc.) que aplican sólo a procesos o series estacionarias. Sin embargo, no hemos proporcionado una prueba estadística para determinar si una serie es estacionaria. Ese es uno de los objetivos de esta sección.

Adicionalmente, en esta sección relajaremos la definición de estacionariedad y plantearemos pruebas para determinar cuándo una serie es estacionaria bajo tres diferentes especificaciones:

1. Estacionariedad alrededor de una tendencia determinística;
2. Estacionariedad alrededor de una media, y
3. Estacionariedad alrededor del cero.

A continuación, discutiremos cómo es posible que una serie sea estacionaria al rededor de una tendencia determinística. Diremos que una tendencia es determinística si ésta puede ser aproximada o modelada por un polinomio en función de t , la cual incluye posibles transformaciones logarítmicas.

Bajo este enfoque, el proceso está lejos de cumplir con la definición de estacionariedad que hemos establecido en capítulos previos, pero relajaremos el supuesto y reconoceremos que una serie puede ser estacionaria en varianza bajo una tendencia determinística. Así, diremos que la serie será descrita por una ecuación

dada por:

$$Y_t = \sum_{j=0}^m \delta_j t^j + X_t \quad (4.1)$$

Donde X_t es un proceso $ARMA(p, q)$ con media cero, que se puede ver como:

$$\alpha(L)X_t = \beta(L)U_t \quad (4.2)$$

Entonces, los momentos y variación de la ecuación (4.1) estarán dados por:

$$\mathbb{E}[Y_t] = \sum_{j=0}^m \delta_j t^j = \mu_t \quad (4.3)$$

Dada la ecuación (4.3) podemos plantear la siguiente ecuación de covarianzas:

$$\begin{aligned} \mathbb{E}[(Y_t - \mu_t) \cdot (Y_{t+\tau} - \mu_{t+\tau})] &= \mathbb{E}[X_t \cdot X_{t+\tau}] \\ &= \gamma_X(\tau) \end{aligned} \quad (4.4)$$

Utilizando el resultado de la ecuación (4.4) podemos establecer que:

$$\mathbb{E}[(Y_t - \mu_t)^2] = \mathbb{E}[X_t^2] = \sigma_X^2 \quad (4.5)$$

Así, las ecuaciones (4.3) y (4.5) significan que el proceso descrito por la ecuación (4.1) es estacionario pero en varianza. De esta forma, a partir de este momento diremos que una serie será estacionaria al rededor de una tendencia determinística si cumple con las condiciones establecidas en las ecuaciones (4.1), (4.3) y (4.5).

Dicho lo anterior, estudiaremos el concepto de raíz unitaria de un proceso estocástico o de una serie de tiempo. Partamos de plantear que un proceso $AR(1)$ tiene raíz unitaria cuando el coeficiente $a_1 = 1$, es decir:

$$Y_t = Y_{t-1} + U_t \quad (4.6)$$

Donde U_t es un proceso puramente aleatorio con media cero, varianza constante y autocovarianza cero (0), al cual nos referiremos simplemente como ruido blanco. Supongamos ahora que incluimos un término constante en la ecuación (4.6), de forma que tenemos:

$$Y_t = \delta + Y_{t-1} + U_t \quad (4.7)$$

Tomando a la ecuación (4.7) y suponiendo que existe un valor inicial Y_0 de la serie, podemos plantear la siguiente secuencia de expresiones:

$$\begin{aligned} Y_1 &= \delta + Y_0 + U_1 \\ Y_2 &= \delta + Y_1 + U_2 \\ &= \delta + (\delta + Y_0 + U_1) + U_2 \\ &= 2 \times \delta + Y_0 + U_1 + U_2 \end{aligned}$$

Si repitiéramos la sustitución sucesiva anterior hasta el momento t encontraríamos que la ecuación de la solución general que describe a un $AR(1)$ con término constante que tiene raíz unitaria es de la forma:

$$Y_t = t\delta + Y_0 + \sum_{i=1}^t U_i \quad (4.8)$$

La ecuación (4.8) es equivalente a la ecuación (4.1). A la ecuación (4.8) también se le conoce como proceso con Drift o con término constante, indistintamente, ya que el componente de Drift suele asociarse a la posibilidad de incorporar el efecto de los residuales pasados, lo cual es posible simplemente agregando una constante.s

Si revisamos el comportamiento de sus momentos y varianza de la ecuación (4.8) encontramos que:

$$\begin{aligned} \mathbb{E}[Y_t] &= Y_0 + \delta t = \mu_t \\ \text{Var}[Y_t] &= t\sigma^2 = \gamma(0, t) \\ \text{Cov}(Y_t, Y_{t+\tau}) &= (t - \tau)\sigma^2 = \gamma(t, \tau) \end{aligned}$$

De esta forma, la ecuación (4.8) no describe un proceso estacionario, sólo podría ser estacionario si $t = 1$, en cualquier otro caso sería no estacionario en varianza. Ahora hagamos un resumen y acordemos notación que se utilizará en esta sección. Supongamos un proceso o serie de tiempo que es descrito por la siguiente ecuación:

$$Y_t = \delta + Y_{t-1} + X_t \quad (4.9)$$

Donde X_t es un $ARMA(p, q)$ con media cero. Si definimos a $\Delta Y_t = Y_t - Y_{t-1}$, entonces la ecuación (4.9) la podemos escribir como:

$$\Delta Y_t = \delta + X_t \quad (4.10)$$

A la ecuación (4.10) la denominaremos como un proceso estacionario en diferencias o simplemente como un proceso integrado. Así, utilizaremos la siguiente definición.

Sea un proceso estocástico Y , decimos que es un **proceso integrado de orden d** , $I(d)$, si este puede transformarse a uno estacionario, que sea invertible, mediante la diferenciación del mismo d -veces, es decir:

$$(1 - L)^d Y_t = \delta + X_t \quad (4.11)$$

Donde X_t es un proceso $ARMA(p, q)$. De lo cual se infiere que en la ecuación (4.11) Y_t será una $ARIMA(p, d, q)$, el cual contiene d raíces unitarias. A estos procesos también se les conoce como procesos con tendencia estocástica.

Dada la discusión anterior, a continuación, plantearemos un resumen de cuáles son los dos casos a los cuales nos referiremos como procesos que no son estacionarios en media, pero que sí lo son en varianza. Estos casos son:

$$Y_t = Y_0 + \delta t + U_t \quad (4.12)$$

$$Y_t = Y_0 + \delta + \sum_{i=1}^t U_i \quad (4.13)$$

Ambos casos no son estacionarios en media, pero sí lo son en varianza. De ambos podemos decir que los choques o innovaciones del término de error tienen un efecto transitorio en el primero, pero permanentes en el segundo.

4.2 Pruebas de Raíces Unitarias

En esta sección plantearemos una serie de pruebas estadísticas para determinar cuándo una serie puede ser estacionaria bajo tres posibles casos:

1. Estacionariedad alrededor de una tendencia determinística;
2. Estacionariedad alrededor de una media, y
3. Estacionariedad alrededor del cero.

A continuación, mostraremos como ejemplo la aplicación de las pruebas de raíces unitarias a la serie de Tipo de Cambio en forma logarítmica y de diferencias logarítmicas. Asumamos que determinamos el valor de los rezagos de la prueba con el criterio de $p = \text{int}\{4(T/100)^{1/4}\}$. En la práctica, existen otras formas de determinar el valor de p , como el criterio de AIC, pero es decisión del investigador cuál usar.

```
library(ggplot2)
library(dplyr)
library(stats)
library(MASS)
library(strucchange)
library(zoo)
library(sandwich)
library(urca)
library(lmtest)
library(vars)

#
load("BD/Datos_Ad.RData")
```

```
#
## Conversion a series de tiempo:
Datos <- ts(Datos_Ad[7: 11],
            start = c(2000, 1),
            freq = 12)

LDatos <- log(Datos)

DLDatos <- diff(log(Datos, base = exp(1)),
               lag = 1,
               differences = 1)

DaLDatos <- diff(log(Datos, base = exp(1)),
                 lag = 12,
                 differences = 1)
```

Como primer paso, exploraremos las series en niveles (bajo una transformación logarítmica), como en la Figura 4.1, y en diferencias logarítmicas, como en la Figura 4.2. De dicha exploración, podemos concluir que las series parecen estacionarias en diferencias, pero no en niveles.

```
plot(LDatos,
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     main = "Series en Logaritmos",
     xlab = "Tiempo")
```

```
plot(DLDatos,
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     main = "Series en Diferencias Logarítmicas",
     xlab = "Tiempo")
```

```
plot(DaLDatos,
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     main = "Series en Diferencias Anuales Logarítmicas",
     xlab = "Tiempo")
```

```
plot(cbind(DLDatos, DaLDatos),
     plot.type = "m", nc = 2,
```

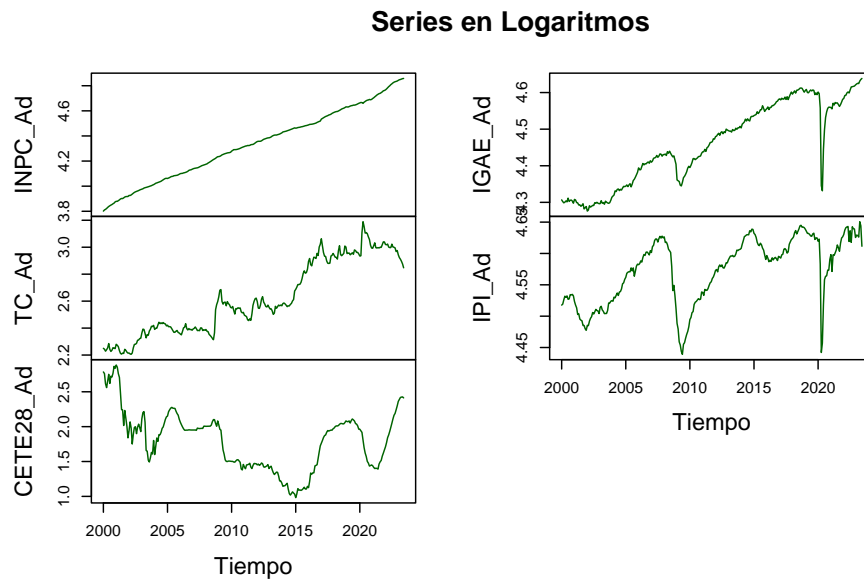


Figure 4.1: Series en Logaritmos

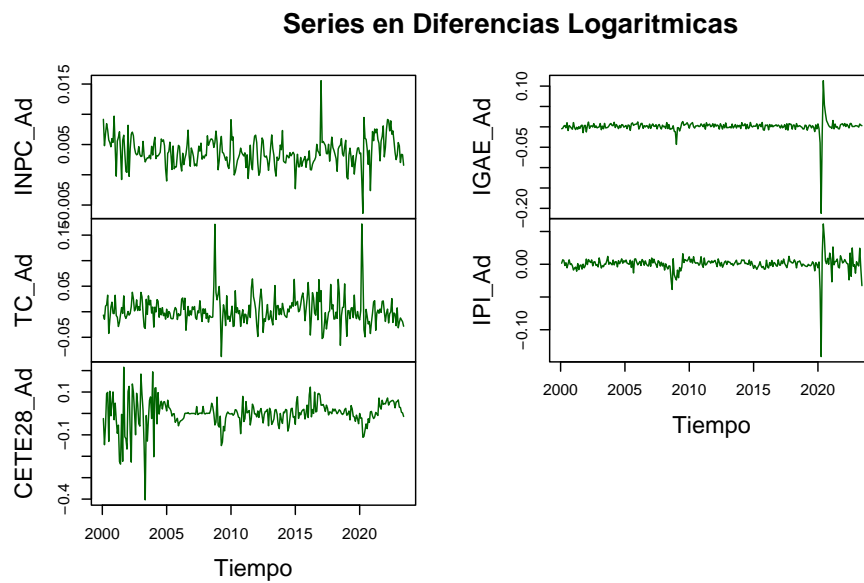


Figure 4.2: Series en Diferencias Logaritmicas

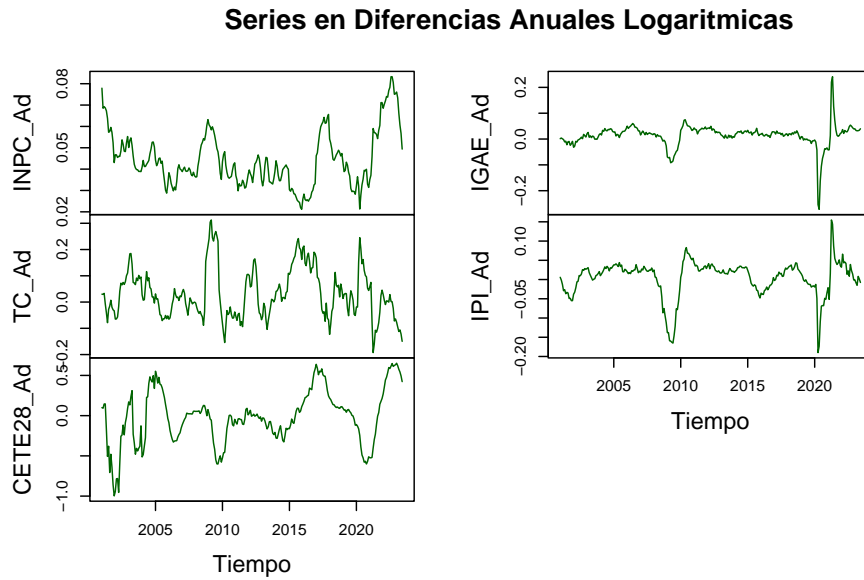


Figure 4.3: Series en Diferencias Anuales Logaritmicas

```
col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
main = "Comparacion de Series en Diferencias",
xlab = "Tiempo")
```

```
plot(cbind(LDatos, DLDatos),
plot.type = "m", nc = 2,
col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
main = "Comparacion de Series en Diferencias",
xlab = "Tiempo")
```

4.2.1 Dickey - Fuller (DF)

Partamos de una forma del proceso Y_t dada por:

$$Y_t = \sum_{j=0}^m \delta_j t^j + X_t \quad (4.14)$$

Donde X_t es un $ARMA(p, q)$ con media cero. Esta prueba asume que $m = 1$, por lo que utilizaremos un modelo del tipo:

$$Y_t = \alpha + \delta t + \rho Y_{t-1} + U_t \quad (4.15)$$

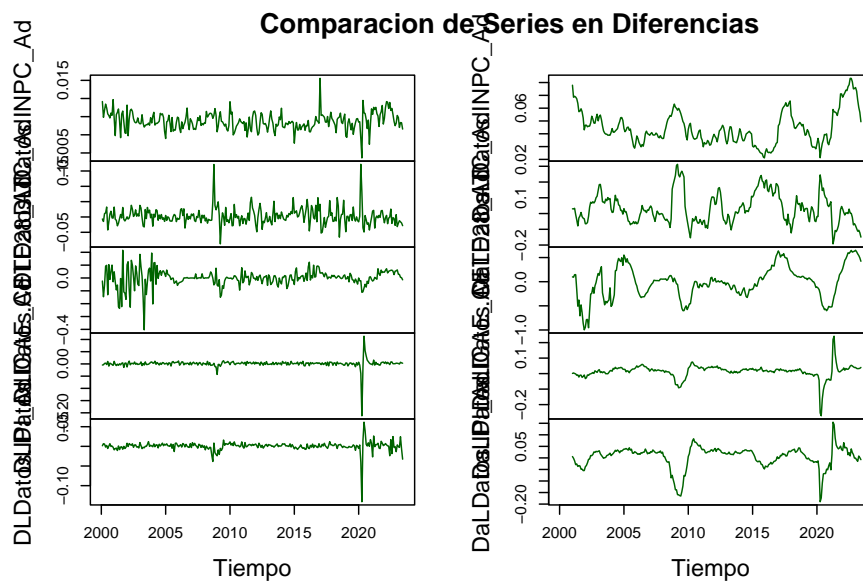


Figure 4.4: Comparacion de Series en Diferencias Anuales

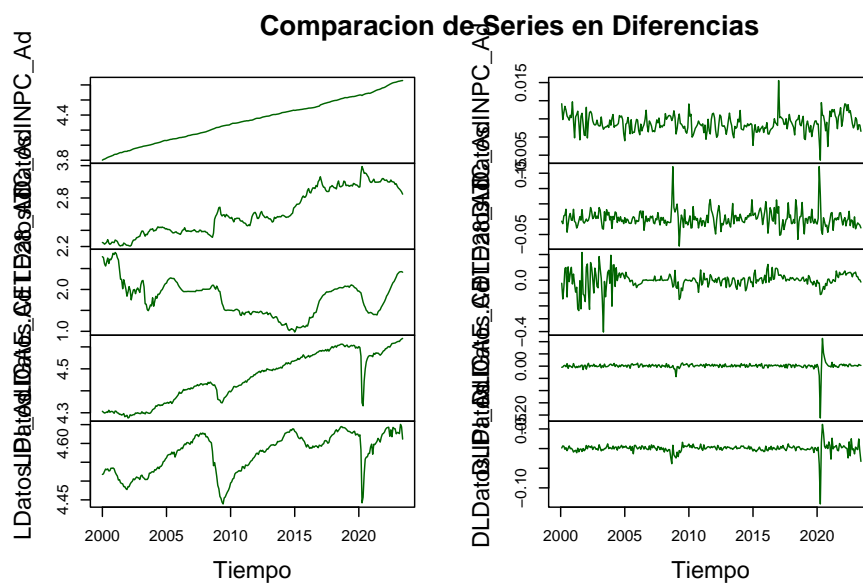


Figure 4.5: Comparacion de Series en Diferencias

Si, el $AR(1)$ planteado tiene raíz unitaria, es decir, $\rho = 1$, entonces tendríamos:

$$\begin{aligned} Y_t &= \alpha + \delta t + Y_{t-1} + U_t \\ \Delta Y_t &= \alpha + \delta t + U_t \end{aligned}$$

De esta forma, para determinar si una serie tiene raíz unitaria basta con probar la hipótesis nula de que $H_0 : \rho = 1$, junto con las diferentes combinaciones que impliquen restricciones respecto a δ y α .

En resumen, la prueba DF consiste en asumir un modelo general dado por la ecuación (4.15) y probar tres especificaciones distintas que serían válidas bajo $H_0 : \rho = 1$:

1. Modelo A: con intercepto y tendencia:

$$\Delta Y_t = \alpha + \delta t + \beta Y_{t-1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una tendencia determinística.

2. Modelo B: con intercepto:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una constante.

3. Modelo C: sin intercepto y tendencia:

$$\Delta Y_t = \beta Y_{t-1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario sin considerar una constante o una tendencia determinística, es decir, es un proceso puramente aleatorio.

Ejemplo. En los siguientes mostramos los resultados de aplicar las pruebas de raíces unitarias a la serie $LT C_t$, considerando $p = \text{int}\{4(234/100)^{(1/4)}\} = 4$ para aquellas pruebas que requieren de rezagos: ADF, PP y KPSS. La convención es implementar las pruebas de raíces unitarias iniciando con el modelo más general hasta llegar al más restringido. Es decir, iniciar con el Modelo A, pasar al Modelo B y cerrar con el Modelo C.

Iniciaremos implementando una prueba de DF simple (o una ADF sin rezagos). En este caso obtenemos los resultados mostrados en la siguiente Tabla. En esta se reportan los resultados de la prueba para cada uno de los 3 modelos (A, B y C) y los valores críticos (CV, por sus siglas en inglés) de la prueba para los niveles 1%, 5% y 10%—estos valores son los reportados por MacKinnon (1996) para una prueba con distribución de 1 cola (izquierda). La convención es utilizar el nivel del 5% de significancia.

Los resultados deben leerse de la siguiente forma. La siguiente Tabla muestra los resultados de la prueba para la serie en niveles, LTC_t , y en diferencias ΔLTC_t . De esta forma, podremos comparar si la serie es estacionaria en niveles o en diferencias. La manera de determinarlo es comparar el estadístico con el valor crítico. Así, rechazaremos la hipótesis nula en los casos en que el estadístico t se ubique por encima del valor crítico. En este caso, podemos observar que la serie se vuelve estacionaria hasta su primera diferencia—caso en el que rechazamos la hipótesis nula en todos los modelos.

Table 4.1: Resumen de resultados de la prueba Dickey-Fuller (DF).

Model	Estadístico	1% CV	5% CV	10% CV
LTC_t				
Modelo A	$t = -2.0752$	-3.98	-3.42	-3.13
Modelo B	$t = -1.1724$	-3.44	-2.87	-2.57
Modelo C	$t = 1.1815$	-2.58	-1.95	-1.62
ΔLTC_t				
Modelo A	$t = -12.6805$	-3.98	-3.42	-3.13
Modelo B	$t = -12.6899$	-3.44	-2.87	-2.57
Modelo C	$t = -12.6515$	-2.58	-1.95	-1.62

```
## Dickey-Fuller:
```

```
### NIVELES: Tipo de cambio
```

```
summary(ur.df(LDatos[, 2], type = "trend", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
```



```
## lm(formula = z.diff ~ z.lag.1 + 1 + tt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.083973 -0.015415 -0.002526  0.009332  0.170645
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.08700923  0.04029878   2.159   0.0317 *
## z.lag.1      -0.03844435  0.01852595  -2.075   0.0389 *
## tt           0.00011203  0.00006246   1.794   0.0740 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02705 on 278 degrees of freedom
## Multiple R-squared:  0.01629,    Adjusted R-squared:  0.00921
## F-statistic: 2.301 on 2 and 278 DF,  p-value: 0.102
##
##
## Value of test-statistic is: -2.0752 2.1124 2.3013
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36
```

```
summary(ur.df(LDatos[, 2], type = "drift", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.089442 -0.014571 -0.002497  0.009620  0.171964
##
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.020316   0.015600   1.302   0.194
## z.lag.1      -0.006946   0.005925  -1.172   0.242
##
## Residual standard error: 0.02716 on 279 degrees of freedom
## Multiple R-squared:  0.004903,   Adjusted R-squared:  0.001336
## F-statistic: 1.375 on 1 and 279 DF,  p-value: 0.242
##
##
## Value of test-statistic is: -1.1724 1.5477
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79
```

```
summary(ur.df(LDatos[, 2], type = "none", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.089736 -0.014701 -0.003293  0.010359  0.169748
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## z.lag.1  0.0007280   0.0006161   1.181   0.238
##
## Residual standard error: 0.0272 on 280 degrees of freedom
## Multiple R-squared:  0.004961,   Adjusted R-squared:  0.001407
## F-statistic: 1.396 on 1 and 280 DF,  p-value: 0.2384
##
##
## Value of test-statistic is: 1.1815
##
## Critical values for test statistics:
```

```
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

```
### DIFERENCIAS: Tipo de cambio
```

```
summary(ur.df(DLdatos[, 2], type = "trend", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09095 -0.01306 -0.00158  0.01055  0.17069
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.002881589  0.003163345   0.911      0.363
## z.lag.1      -0.736373275  0.058071315 -12.680 <0.0000000000000002 ***
## tt           -0.000009351  0.000019483  -0.480      0.632
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02634 on 277 degrees of freedom
## Multiple R-squared:  0.3673, Adjusted R-squared:  0.3628
## F-statistic: 80.41 on 2 and 277 DF, p-value: < 0.00000000000000022
##
##
## Value of test-statistic is: -12.6805 53.6071 80.4095
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36
```

```
summary(ur.df(DLDatos[, 2], type = "drift", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-0.090664	-0.013492	-0.001331	0.010702	0.169754

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.001566	0.001577	0.993	0.322
z.lag.1	-0.735658	0.057972	-12.690	<0.0000000000000002 ***

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02631 on 278 degrees of freedom
## Multiple R-squared:  0.3668, Adjusted R-squared:  0.3645
## F-statistic: 161 on 1 and 278 DF, p-value: < 0.00000000000000022
##
##
## Value of test-statistic is: -12.6899 80.5185
##
## Critical values for test statistics:
```

	1pct	5pct	10pct
tau2	-3.44	-2.87	-2.57
phi1	6.47	4.61	3.79

```
summary(ur.df(DLDatos[, 2], type = "none", lags = 0))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
```

```
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.089121 -0.011916  0.000311  0.012192  0.171310
##
## Coefficients:
##      Estimate Std. Error t value      Pr(>|t|)
## z.lag.1 -0.73093    0.05777  -12.65 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02631 on 279 degrees of freedom
## Multiple R-squared:  0.3646, Adjusted R-squared: 0.3623
## F-statistic: 160.1 on 1 and 279 DF,  p-value: < 0.00000000000000022
##
##
## Value of test-statistic is: -12.6515
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

4.2.2 Dickey - Fuller Aumentada (ADF)

A diferencia de un modelo AR(1) para el caso de una prueba DF como en la ecuación (4.15), en una prueba ADF se asume que el proceso es un AR(p) de la forma (por simplicidad, hemos omitido el término constante y el término de tendencia determinística):

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p} + U_t \quad (4.16)$$

Haciendo una sustitución de términos similar a las que hemos planteado en otras secciones, podemos reexpresar la ecuación (4.16) en su versión en diferencias siguiendo el proceso:

$$Y_t = \rho Y_{t-1} + \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \dots + \theta_{p-1} \Delta Y_{t-p+1} + U_t \quad (4.17)$$

Donde $\rho = \theta_0 = \sum_{j=1}^p a_j$, $\theta_i = -\sum_{j=i+1}^p a_j$, $i = 1, 2, \dots, p-1$. Así, si el proceso AR(p) tiene raíz unitaria, entonces, veremos que:

$$\begin{aligned} 1 - a_1 - a_2 - \dots - a_p &= 0 \\ \rho &= 1 \end{aligned}$$

De donde podemos establecer que el modelo general de una prueba ADF será:

$$\Delta Y_{t-1} = \alpha + \beta t + (\rho - 1)Y_{t-1} + \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \dots + \theta_k \Delta Y_{t-k} + U_t \quad (4.18)$$

Donde U_t es un proceso puramente aleatorio y k es elegido de tal manera que los residuales sean un proceso puramente aleatorio. En resumen, la prueba ADF consiste en asumir un modelo general dado por la ecuación (4.16), que incluya constante y tendencia, y probar tres especificaciones distintas que serían válidas bajo $H_0 : \rho = 1$:

1. Modelo A: con intercepto y tendencia:

$$\Delta Y_t = \alpha + \delta t + \beta Y_{t-1} + \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \dots + \theta_{p-1} \Delta Y_{t-p+1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una tendencia determinística.

2. Modelo B: con intercepto:

$$\Delta Y_t = \alpha + \beta Y_{t-1} + \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \dots + \theta_{p-1} \Delta Y_{t-p+1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una constante.

3. Modelo C: sin intercepto y tendencia:

$$\Delta Y_t = \beta Y_{t-1} + \theta_1 \Delta Y_{t-1} + \theta_2 \Delta Y_{t-2} + \dots + \theta_{p-1} \Delta Y_{t-p+1} + U_t$$

Buscamos probar si $H_0 : \beta = \rho - 1 = 0$ contra $H_a : \beta < 0$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario sin considerar una constante o una tendencia determinística, es decir, es un proceso puramente aleatorio.

Ejemplo. La siguiente Tabla muestra los resultados de la prueba para la serie en niveles, LTC_t , y en diferencias ΔLTC_t de una prueba ADF. De esta forma, podremos comparar si la serie es estacionaria en niveles o en diferencias. La manera de determinarlo es, al igual que en el caso anterior, comparar el estadístico con el valor crítico. Así, rechazaremos la hipótesis nula en los casos en que el estadístico t se ubique por encima del valor crítico. En este caso, podemos observar que la serie se vuelve estacionaria hasta su primera diferencia—caso en el que rechazamos la hipótesis nula en todos los modelos.

Table 4.2: Resumen de resultados de la prueba Augmented Dickey-Fuller (ADF) con 5 rezagos (resultado de $p = \text{int}\{4(282/100)^{(1/4)}\} = 5$).

Model	Estadístico	1% CV	5% CV	10% CV
LTC_t				
Modelo A	$t = -2.4972$	-3.98	-3.42	-3.13
Modelo B	$t = -1.1265$	-3.44	-2.87	-2.57
Modelo C	$t = 1.0273$	-2.58	-1.95	-1.62
ΔLTC_t				
Modelo A	$t = -7.4125$	-3.98	-3.42	-3.13
Modelo B	$t = -7.4034$	-3.44	-2.87	-2.57
Modelo C	$t = -7.2639$	-2.58	-1.95	-1.62

```
## Augmented Dickey - Fuller
```

```
### NIVELES: Tipo de cambio
```

```
summary(ur.df(LDatos[, 2], type = "trend", lags = 5))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.077054 -0.014133 -0.002281  0.011551  0.167611
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.10899655  0.04255838   2.561    0.0110 *
## z.lag.1      -0.04919274  0.01969939  -2.497    0.0131 *
## tt           0.00015118  0.00006679   2.263    0.0244 *
## z.diff.lag1  0.32887766  0.06112000   5.381 0.000000162 ***
## z.diff.lag2 -0.12845492  0.06432008  -1.997    0.0468 *
## z.diff.lag3 -0.01064885  0.06462312  -0.165    0.8692
## z.diff.lag4  0.05690285  0.06350714   0.896    0.3711
```

```
## z.diff.lag5 -0.01668349  0.06151762  -0.271      0.7864
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02597 on 268 degrees of freedom
## Multiple R-squared:  0.1196, Adjusted R-squared:  0.09662
## F-statistic: 5.202 on 7 and 268 DF,  p-value: 0.00001414
##
##
## Value of test-statistic is: -2.4972 2.5862 3.2058
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36
```

```
summary(ur.df(LDatos[, 2], type = "drift", lags = 5))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.081102 -0.013688 -0.002598  0.011536  0.167963
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.019078   0.015382   1.240    0.2160
## z.lag.1      -0.006575   0.005837  -1.126    0.2610
## z.diff.lag1   0.306455   0.060772   5.043 0.000000845 ***
## z.diff.lag2  -0.156749   0.063575  -2.466    0.0143 *
## z.diff.lag3  -0.034594   0.064238  -0.539    0.5907
## z.diff.lag4   0.038455   0.063463   0.606    0.5451
## z.diff.lag5  -0.041998   0.060954  -0.689    0.4914
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



```
##
## Residual standard error: 0.02617 on 269 degrees of freedom
## Multiple R-squared:  0.1028, Adjusted R-squared:  0.08277
## F-statistic: 5.136 on 6 and 269 DF,  p-value: 0.00005131
##
##
## Value of test-statistic is: -1.1265 1.2978
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau2 -3.44 -2.87 -2.57
## phi1  6.47  4.61  3.79
```

```
summary(ur.df(LDatos[, 2], type = "none", lags = 5))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.080975 -0.014773 -0.002006  0.011681  0.165789
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## z.lag.1      0.0006247  0.0006081   1.027    0.3052
## z.diff.lag1  0.3052416  0.0608250   5.018 0.000000947 ***
## z.diff.lag2 -0.1599301  0.0635869  -2.515    0.0125 *
## z.diff.lag3 -0.0366726  0.0642802  -0.571    0.5688
## z.diff.lag4  0.0367316  0.0635108   0.578    0.5635
## z.diff.lag5 -0.0449750  0.0609677  -0.738    0.4613
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0262 on 270 degrees of freedom
## Multiple R-squared:  0.1027, Adjusted R-squared:  0.08271
## F-statistic: 5.148 on 6 and 270 DF,  p-value: 0.00004979
##
```

```
##
## Value of test-statistic is: 1.0273
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

DIFERENCIAS: Tipo de cambio

```
summary(ur.df(DLDatos[, 2], type = "trend", lags = 5))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.069846 -0.013941 -0.001969  0.011796  0.166553
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.00398042  0.00326372   1.220      0.2237
## z.lag.1      -0.96600985  0.13032157  -7.413 0.00000000000164 ***
## tt           -0.00001242  0.00001981  -0.627      0.5310
## z.diff.lag1   0.27428705  0.11800103   2.324      0.0209 *
## z.diff.lag2   0.11502927  0.10628825   1.082      0.2801
## z.diff.lag3   0.07837533  0.09095430   0.862      0.3896
## z.diff.lag4   0.09775877  0.07423509   1.317      0.1890
## z.diff.lag5   0.07280186  0.06082666   1.197      0.2324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02605 on 267 degrees of freedom
## Multiple R-squared:  0.3901, Adjusted R-squared:  0.3741
## F-statistic: 24.39 on 7 and 267 DF,  p-value: < 0.0000000000000022
##
##
## Value of test-statistic is: -7.4125 18.3632 27.5402
```

```
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau3 -3.98 -3.42 -3.13
## phi2  6.15  4.71  4.05
## phi3  8.34  6.30  5.36

summary(ur.df(DLDatos[, 2], type = "drift", lags = 5))

##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.070858 -0.014250 -0.001793  0.011998  0.165361
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.002197   0.001601   1.372      0.1711
## z.lag.1      -0.963143   0.130094  -7.403 0.000000000000172 ***
## z.diff.lag1   0.272137   0.117818   2.310      0.0217 *
## z.diff.lag2   0.113463   0.106139   1.069      0.2860
## z.diff.lag3   0.077366   0.090837   0.852      0.3951
## z.diff.lag4   0.097369   0.074148   1.313      0.1903
## z.diff.lag5   0.072725   0.060758   1.197      0.2324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02602 on 268 degrees of freedom
## Multiple R-squared:  0.3892, Adjusted R-squared:  0.3755
## F-statistic: 28.46 on 6 and 268 DF,  p-value: < 0.00000000000000022
##
##
## Value of test-statistic is: -7.4034 27.41
##
## Critical values for test statistics:
##      1pct  5pct 10pct
```

```
## tau2 -3.44 -2.87 -2.57
## phi1 6.47 4.61 3.79
```

```
summary(ur.df(DLDatos[, 2], type = "none", lags = 5))
```

```
##
## #####
## # Augmented Dickey-Fuller Test Unit Root Test #
## #####
##
## Test regression none
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.069424 -0.012002  0.000317  0.013820  0.167787
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## z.lag.1      -0.92780    0.12773  -7.264 0.000000000000407 ***
## z.diff.lag1   0.24302    0.11608   2.094    0.0372 *
## z.diff.lag2   0.08889    0.10479   0.848    0.3970
## z.diff.lag3   0.05875    0.08997   0.653    0.5143
## z.diff.lag4   0.08492    0.07371   1.152    0.2503
## z.diff.lag5   0.06500    0.06060   1.073    0.2844
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02607 on 269 degrees of freedom
## Multiple R-squared:  0.3849, Adjusted R-squared:  0.3712
## F-statistic: 28.05 on 6 and 269 DF,  p-value: < 0.00000000000000022
##
##
## Value of test-statistic is: -7.2639
##
## Critical values for test statistics:
##      1pct  5pct 10pct
## tau1 -2.58 -1.95 -1.62
```

4.2.3 Phillips - Perron (PP)

Una tercera prueba es la de PP, la cual también está basada en un AR(1) descrito por la ecuación:

$$Y_t = d\eta + \rho Y_{t-1} + U_t \quad (4.19)$$

Donde d incluye a cualquiera de los componentes determinísticos como constante y tendencia.

Al igual que los casos pasados, la hipótesis a probar es $H_0 : \rho = 1$ contra la alternativa $H_a : |\rho| < 1$, y asumimos una estructura de MA(q) que es un término de error de la forma:

$$U_t = \psi(L)\varepsilon_t = \psi_0\varepsilon_t + \psi_1\varepsilon_{t-1} + \dots + \psi_p\varepsilon_{t-p} \quad (4.20)$$

Con ε_t que es un ruido blanco con media cero y varianza σ^2 . En este modelo se elige el valor p que hace que el componente sea un MA(p). Las tablas estadísticas de PP para esta prueba pueden utilizar una estadística Z_τ o Z_ρ , las cuales se pueden emplear indistintamente.

En resumen, la prueba PP consiste en asumir un modelo general dado por la ecuación (4.19) y probar dos especificaciones distintas que serían válidas bajo $H_0 : \rho = 1$, ambas considerando un componente Drift:

1. Modelo A: con intercepto y tendencia:

$$Y_t = \alpha + \delta t + \rho Y_{t-1} + U_t$$

Buscamos probar si $H_0 : \rho = 1$ contra $H_a : |\rho| < 1$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una tendencia determinística.

2. Modelo B: con intercepto:

$$Y_t = \alpha + \rho Y_{t-1} + U_t$$

Buscamos probar si $H_0 : \rho = 1$ contra $H_a : |\rho| < 1$, por lo que es una prueba de una cola. Otra forma de decirlo es: probamos si el proceso tiene raíz unitaria contra si el proceso es estacionario alrededor de una constante.

Ejemplo. La siguiente Tabla muestra los resultados de la prueba para la serie en niveles, LTC_t , y en diferencias ΔLTC_t de una prueba PP. De esta forma, podremos comparar si la serie es estacionaria en niveles o en diferencias. La manera de determinarlo es, al igual que en los casos anteriores, comparar el

estadístico con el valor crítico (Z_α o Z_τ). Así, rechazaremos la hipótesis nula en los casos en que el estadístico Z_α o Z_τ se ubique por encima del valor crítico. En este caso, podemos observar que la serie se vuelve estacionaria hasta su primera diferencia—caso en el que rechazamos la hipótesis nula en todos los modelos.

Table 4.3: Resumen de resultados de la prueba Phillips-Perron (PP) con 5 rezagos (resultado de $p = \text{int}\{4(282/100)^{(1/4)}\} = 5$).

Model	Estadístico	1% CV	5% CV	10% CV
LTC_t				
Modelo A	$Z_\tau = -2.5039$	-3.994127	-3.427199	-3.136601
Modelo B	$Z_\tau = -1.2297$	-3.455219	-2.87195	-2.572274
ΔLTC_t				
Modelo A	$Z_\tau = -12.3906$	-3.994127	-3.427199	-3.136601
Modelo B	$Z_\tau = -12.4037$	-3.455219	-2.87195	-2.572274

```
## PP: Phillips - Perron Test

### NIVELES: Tipo de cambio

summary(ur.pp(LDatos[, 2], type = "Z-tau", model = "trend", use.lag = 5))

##
## #####
## # Phillips-Perron Unit Root Test #
## #####
##
## Test regression with intercept and trend
##
##
## Call:
## lm(formula = y ~ y.l1 + trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.083973 -0.015415 -0.002526  0.009332  0.170645
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.10275012  0.04851421   2.118    0.0351 *
## y.l1         0.96155565  0.01852595  51.903 <0.0000000000000002 ***
## trend        0.00011203  0.00006246   1.794    0.0740 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.02705 on 278 degrees of freedom
## Multiple R-squared:  0.9903, Adjusted R-squared:  0.9902
## F-statistic: 1.416e+04 on 2 and 278 DF,  p-value: < 0.000000000000000022
##
##
## Value of test-statistic, type: Z-tau  is: -2.5039
##
##          aux. Z statistics
## Z-tau-mu          1.1638
## Z-tau-beta        2.2617
##
## Critical values for Z statistics:
##              1pct      5pct      10pct
## critical values -3.994127 -3.427199 -3.136601

summary(ur.pp(LDatos[, 2], type = "Z-tau", model = "constant", use.lag = 5))

##
## #####
## # Phillips-Perron Unit Root Test #
## #####
##
## Test regression with intercept
##
##
## Call:
## lm(formula = y ~ y.l1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.089442 -0.014571 -0.002497  0.009620  0.171964
##
## Coefficients:
##              Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.020316   0.015600   1.302      0.194
## y.l1         0.993054   0.005925 167.617 <0.00000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02716 on 279 degrees of freedom
## Multiple R-squared:  0.9902, Adjusted R-squared:  0.9901
## F-statistic: 2.81e+04 on 1 and 279 DF,  p-value: < 0.000000000000000022
##
##
```

```
## Value of test-statistic, type: Z-tau is: -1.2297
##
##          aux. Z statistics
## Z-tau-mu          1.3459
##
## Critical values for Z statistics:
##          1pct      5pct      10pct
## critical values -3.455219 -2.87195 -2.572274
```

DIFERENCIAS: Tipo de cambio

```
summary(ur.pp(DLDatos[, 2], type = "Z-tau", model = "trend", use.lag = 5))
```

```
##
## #####
## # Phillips-Perron Unit Root Test #
## #####
##
## Test regression with intercept and trend
##
##
## Call:
## lm(formula = y ~ y.l1 + trend)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.09095 -0.01306 -0.00158  0.01055  0.17069
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.001572438  0.001579672   0.995   0.320
## y.l1         0.263626725  0.058071315   4.540 0.00000841 ***
## trend       -0.000009351  0.000019483  -0.480   0.632
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02634 on 277 degrees of freedom
## Multiple R-squared:  0.07036, Adjusted R-squared:  0.06365
## F-statistic: 10.48 on 2 and 277 DF, p-value: 0.00004089
##
##
## Value of test-statistic, type: Z-tau is: -12.3906
##
##          aux. Z statistics
## Z-tau-mu          1.2352
```



```
## Z-tau-beta          -0.4753
##
## Critical values for Z statistics:
##              1pct      5pct      10pct
## critical values -3.994237 -3.427252 -3.136632
```

```
summary(ur.pp(DLDatos[, 2], type = "Z-tau", model = "constant", use.lag = 5))
```

```
##
## #####
## # Phillips-Perron Unit Root Test #
## #####
##
## Test regression with intercept
##
##
## Call:
## lm(formula = y ~ y.l1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.090664 -0.013492 -0.001331  0.010702  0.169754
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)  0.001566   0.001577   0.993      0.322
## y.l1         0.264342   0.057972   4.560 0.00000768 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02631 on 278 degrees of freedom
## Multiple R-squared:  0.06959,    Adjusted R-squared:  0.06624
## F-statistic: 20.79 on 1 and 278 DF,  p-value: 0.000007682
##
##
## Value of test-statistic, type: Z-tau is: -12.4037
##
##      aux. Z statistics
## Z-tau-mu          0.969
##
## Critical values for Z statistics:
##              1pct      5pct      10pct
## critical values -3.455298 -2.871985 -2.572293
```

4.2.4 Kwiatkowski - Phillips - Schmidt - Shin (KPSS)

La prueba KPSS considera que el proceso es estacionario bajo la hipótesis nula, lo cual hace una diferencia respecto de las anteriores pruebas. El modelo considerado es:

$$Y_t = \delta t + \xi_t + U_t \quad (4.21)$$

Donde U_t es un proceso estacionario y ξ_t es un ruido blanco descrito por la forma: $\xi_t = \xi_{t-1} + \varepsilon_t$, donde ε_t es un proceso normalmente distribuido con media cero y varianza σ_ε^2 .

Así, bajo la hipótesis nula $H_0 : \sigma_\varepsilon^2 = 0$, ξ se vuelve una constante y el proceso puede tener una tendencia estacionaria. Dado el planteamiento de la prueba, los valores críticos al 95% son:

1. **Modelo A:** 0.146, para un modelo con tendencia (estadístico τ).
2. **Modelo B:** 0.463, para un modelo con constante (estadístico μ).

Ejemplo. Finalmente, la siguiente Tabla muestra los resultados de la prueba para la serie en niveles, LTC_t , y en diferencias ΔLTC_t de una prueba KPSS. De esta forma, podremos comparar si la serie es estacionaria en niveles o en diferencias. La manera de determinarlo es, al igual que en los casos anteriores, comparar el estadístico con el valor crítico (μ y τ). No obstante, en este caso buscaremos aceptar la hipótesis nula en los casos en que los estadísticos μ y τ se ubique por debajo del valor crítico. En este caso, podemos observar que la serie se vuelve estacionaria hasta su primera diferencia-caso en el que rechazamos la hipótesis nula en todos los modelos.

Table 4.4: Resumen de resultados de la prueba Kwiatkowski - Phillips - Schmidt - Shin (KPSS).

Model	Estadístico	1% CV	5% CV	10% CV
LTC_t				
Modelo A	$\tau = 0.3433$	0.216	0.146	0.119
Modelo B	$\mu = 4.429$	0.739	0.463	0.347
ΔLTC_t				
Modelo A	$\tau = 0.0701$	0.216	0.146	0.119
Modelo B	$\mu = 0.0801$	0.739	0.463	0.347

KPSS:

NIVELES: Tipo de cambio

```
summary(ur.kpss(LDatos[, 2], type = "tau"))
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: tau with 5 lags.
##
## Value of test-statistic is: 0.3433
##
## Critical value for a significance level of:
##          10pct  5pct  2.5pct  1pct
## critical values 0.119 0.146  0.176 0.216
```

```
summary(ur.kpss(LDatos[, 2], type = "mu"))
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: mu with 5 lags.
##
## Value of test-statistic is: 4.429
##
## Critical value for a significance level of:
##          10pct  5pct  2.5pct  1pct
## critical values 0.347 0.463  0.574 0.739
```

```
### DIFERENCIAS: Tipo de cambio
```

```
summary(ur.kpss(DLDatos[, 2], type = "tau"))
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: tau with 5 lags.
##
## Value of test-statistic is: 0.0701
##
## Critical value for a significance level of:
##          10pct  5pct  2.5pct  1pct
## critical values 0.119 0.146  0.176 0.216
```

```
summary(ur.kpss(DLDatos[, 2], type = "mu"))
```

```
##
## #####
## # KPSS Unit Root Test #
## #####
##
## Test is of type: mu with 5 lags.
##
## Value of test-statistic is: 0.0801
##
## Critical value for a significance level of:
##          10pct  5pct 2.5pct  1pct
## critical values 0.347 0.463 0.574 0.739
```

Chapter 5

Causalidad de Granger, Modelos Multivariados de Vectores Autoregresivos y de Cointegración, Modelos ARDL y Filtro Multivariado

En este capítulo removeremos el supuesto de que el análisis es univariado, ya que introduciremos la posibilidad de que los procesos generadores de datos compartan información entre dos o más series. Como primera aproximación desarrollaremos el concepto de Causalidad de Granger. Mediante esta metodología discutiremos cuándo dos series se causan estadísticamente. Posteriormente, introduciremos una técnica más sofisticada conocida como la metodología de Vectores Autoregresivos (VAR), la cual es una generalización de los procesos Autoregresivos (AR) que analizamos al principio del curso. Finalmente, introduciremos la técnica de cointegración y de rezagos distribuidos (ARDL) para los casos en que las series que analicemos sean procesos no estacionarios.

A partir de este punto, asumiremos que las series empleadas son estacionarias en sus primeras diferencias y solo nos preocuparemos por su estacionariedad en los casos particulares de Cointegración y los modelos ARDL.

5.1 Causalidad de Granger

Hasta ahora hemos supuesto que una serie puede ser explicada únicamente con la información contenida en ella misma. No obstante, en adelante trataremos de analizar el caso en el que buscamos determinar relaciones entre variables y cómo el comportamiento de una serie influye en las demás. Algunas relaciones más importantes son las llamadas: causalidad. En este caso analizaremos el procedimiento de Granger (1969), conocido como causalidad de Granger. En adelante asumiremos que las series involucradas son débilmente estacionarias.

Sean X y Y dos series débilmente estacionarias. Definamos a I_t un conjunto de toda la información disponible hasta el momento t . Asimismo, digamos que \bar{X}_t y \bar{Y}_t son los conjuntos de toda la información disponible (actual y pasada) de X y Y , respectivamente. Es decir:

$$\begin{aligned}\bar{X}_t &:= \{X_t, X_{t-1}, X_{t-2}, \dots\} \\ \bar{Y}_t &:= \{Y_t, Y_{t-1}, Y_{t-2}, \dots\} \\ I_t &:= \bar{X}_t + \bar{Y}_t\end{aligned}$$

Adicionalmente, definamos $\sigma^2(*)$ como la varianza del término de error estimado de una regresión dada. Dicho lo anterior, digamos que:

1. Existe Causalidad de Granger o X causa a Y si y solo si, una regresión lineal da como resultado que:

$$\sigma^2(Y_{t+1}|I_t) < \sigma^2(Y_{t+1}|I_t - X_t) \quad (5.1)$$

Es decir, que la variabilidad del término de error de una regresión lineal de Y sobre el conjunto de toda la información aplicada a un pronóstico de Y_{t+1} es MENOR que la variabilidad del término de error de una regresión lineal de Y sobre el conjunto de la información de Y aplicada a un pronóstico de Y_{t+1} .

2. Existe Causalidad de Granger Instantánea o X causa de forma instantánea a Y si y solo si, una regresión lineal da como resultado:

$$\sigma^2(Y_{t+1}|\{I_t, X_{t+1}\}) < \sigma^2(Y_{t+1}|I_t) \quad (5.2)$$

La definición anterior aplica de igual forma si se reemplaza a X por Y y a Y por X , respectivamente. De acuerdo a la definición anterior, existen 5 diferentes posibilidades de relaciones causales entre las dos series:

1. X y Y son independientes: (X, Y) ;
2. Existe solo causalidad instantánea: $(X - Y)$;

3. X causa a Y : $(X \longrightarrow Y)$;
4. Y causa a X : $(X \longleftarrow Y)$, y
5. Ambas series se causan: $(X \longleftrightarrow Y)$.

Por lo anterior, representaremos mediante un $AR(p)$ con variables exógenas lo siguiente:

$$A(L) \begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} a_{11}(L) & a_{12}(L) \\ a_{21}(L) & a_{22}(L) \end{bmatrix} \begin{bmatrix} Y_t \\ X_t \end{bmatrix} = \begin{bmatrix} V_t \\ U_t \end{bmatrix} \quad (5.3)$$

O en su versión $MA(q)$ con variables exógenas:

$$\begin{bmatrix} Y_t \\ X_t \end{bmatrix} = B(L) \begin{bmatrix} V_t \\ U_t \end{bmatrix} = \begin{bmatrix} b_{11}(L) & b_{12}(L) \\ b_{21}(L) & b_{22}(L) \end{bmatrix} \begin{bmatrix} V_t \\ U_t \end{bmatrix} \quad (5.4)$$

Para determinar el test de causalidad utilizaremos una especificación similar a la de la ecuación (5.3). Para probar si X causa a Y , consideraremos la siguiente regresión:

$$Y_t = \alpha_0 + \sum_{k=1}^{k_1} a_{11}^k Y_{t-k} + \sum_{k=k_0}^{k_2} a_{12}^k X_{t-k} + U_{1,t} \quad (5.5)$$

Donde $k_0 = 1$ y, en general, se asume que $k_1 = k_2$. Asimismo, el valor de estas constantes se puede determinar con el criterio de Akaike (o cualquier otro criterio de información). No obstante, algunos autores sugieren que una buena práctica es considerar valores de k_1 y k_2 que recorran al 4, 8, 12 y 16.

Dicho lo anterior, el test de causalidad de Granger se establece con una prueba F , en la cual se prueba la siguiente hipótesis nula:

$$H_0 : a_{12}^1 = a_{12}^2 = \dots = a_{12}^{k_2} = 0 \quad (5.6)$$

Ejemplo. Consideremos como variables analizadas al Índice Nacional de Precios al Consumidor ($INPC_t$), al Tipo de Cambio (TDC_t) y al rendimiento anual de los Cetes a 28 días ($CETE28_t$), todas desestacionalizadas para el periodo de enero de 2000 a julio de 2019. Dado que la metodología de Granger supone que las series son estacionarias, utilizaremos las diferencias logarítmicas de cada una de las tres series (es decir, utilizaremos una transformación del tipo $\ln(X_t) - \ln(X_{t-1})$). La Figura 5.1 muestra las series en su transformación de diferencias logarítmicas.

```
library(ggplot2)
library(dplyr)
library(stats)
library(MASS)
library(strucchange)
```

```

library(zoo)
library(sandwich)
library(urca)
library(lmtest)
library(vars)

#
load("BD/Datos_Ad.RData")

#
INPC <- ts(Datos_Ad$INPC_Ad,
           start = c(2000, 1),
           freq = 12)

DLINPC <- diff(log( ts(Datos_Ad$INPC_Ad, start = c(2000, 1), freq = 12) ))

TC <- ts(Datos_Ad$TC_Ad,
         start = c(2000, 1),
         freq = 12)

DLTC <- diff(log( ts(Datos_Ad$TC_Ad, start = c(2000, 1), freq = 12) ))

CETE28 <- ts(Datos_Ad$CETE28_Ad,
            start = c(2000, 1),
            freq = 12)

DLCETE28 <- diff(log( ts(Datos_Ad$CETE28_Ad, start = c(2000, 1), freq = 12) ))

#
#png("Plots/DLGranger.png", width = 800, height = 1200)

par(mfrow=c(3, 1))

plot(DLINPC, xlab = "Tiempo",
     main = "Diferencias Logarítmicas del INPC",
     col = "darkgreen")

plot(DLTC, xlab = "Tiempo",
     main = "Diferencias Logarítmicas del Tipo de Cambio",
     col = "darkblue")

plot(DLCETE28, xlab = "Tiempo",
     main = "Diferencias Logarítmicas de los Cetes a 28 dias",
     col = "darkred")

```

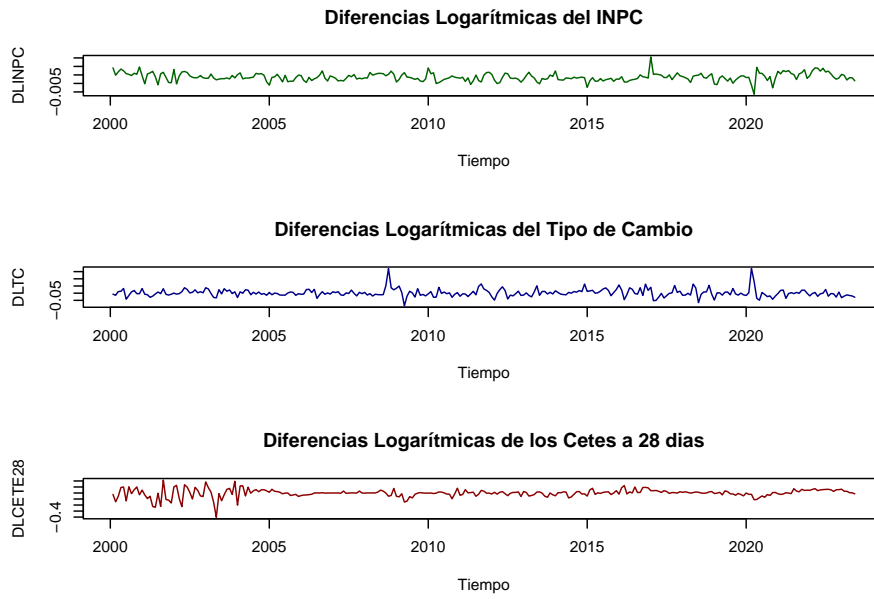



Figure 5.1: Series en diferencias logarítmicas dadas por las siguientes expresiones: $DLINPC_t = \ln(DLINPC_t) - \ln(DLINPC_{t-1})$, $DLTC_t = \ln(TC_t) - \ln(TC_{t-1})$ y $DLCETE28_t = \ln(CETE28_t) - \ln(CETE28_{t-1})$.

```
par(mfrow=c(1, 1))

#dev.off()
```

Por simplicidad, en el Cuadro 5.1 se muestra el resultado de aplicar el test de Granger a diferentes especificaciones, con rezagos 4, 8, 12 y 16, sólo para la serie de Tipo de Cambio en diferencias logarítmicas. En cada una de las pruebas se compara el modelo considerado como regresor a la variable que es candidata de causar, respecto del modelo sin considerar a dicha variable.

Table 5.1: Prueba de si $DLINPC_t$ Granger causa a $DLTC_t$.

Rezagos	Estadística F	Probabilidad (>F)	Significancia
4	3.2621	0.01265	*
8	1.9079	0.06030	
12	2.2577	0.01067	*
16	1.6735	0.05495	*
Notas:	***	signif.	al 0.1%
	**	signif.	al 1%
	*	signif.	al 5%

De acuerdo con el Cuadro 5.1, podemos concluir que existe información estadísticamente significativa para concluir que la inflación causa a la tasa de depreciación cambiaria, ambas medidas como las diferencias logarítmicas. El resto de los resultados para las otras combinaciones de causalidad se encuentran en el R Markdown llamado Clase 13 ubicado en el repositorio de GitHub.

5.2 Definición y representación del Sistema o Modelo de Vectores Autorregresivos (VAR(p))

En esta sección ampliaremos la discusión planteada en el apartado anterior. En el sentido de que en la sección pasada nuestra discusión se limitó al análisis de causalidad entre dos variables a la vez, que si bien es posible extenderlo a más variables, es un procedimiento limitado a casos particulares por las siguientes razones.

El procedimiento de causalidad de Granger supone que es posible identificar un sistema de ecuaciones que debe conformarse una vez que se ha identificado el sentido de la causalidad. Así, el proceso anterior necesita del conocimiento previo de las relaciones que existen entre las variables.

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (VAR)

Adicionalmente, no resuelve el problema más general que está relacionado con cómo identificar la causalidad cuando se tienen múltiples variables con múltiples sentidos de causalidad. En esta sección analizaremos una mejor aproximación al problema de cómo identificar la causalidad múltiple. Por lo tanto, como mecanismo para solucionar el problema planteado, analizaremos el caso de un Sistema o Modelo de Vectores Autoregresivos conocido como VAR.

El primer supuesto del que partiremos es que existe algún grado de endogeneidad entre las variables consideradas en el análisis. Adicionalmente, el segundo supuesto que estableceremos es que requerimos que las variables que tengamos consideradas sean estacionarias.

Por lo anterior, diremos que un modelo de Vectores Autoregresivos (VAR) es un procedimiento que sigue fundado en el supuesto de que las variables consideradas son estacionarias. Así, hasta este momento del curso hemos pasado de modelos univariados a modelos multivariados, pero no hemos podido dejar de asumir que las series son estacionarias.

En lo subsecuente asumiremos que las series empleadas son estacionarias y sólo lo demostraremos cuando, en su caso, sea necesario. Esto no significa que el lector deba asumir estacionariedad. Por el contrario, siempre debe probar que las series son estacionarias antes de iniciar la implementación de cualquier técnica de series de tiempo.

Ahora bien, iniciaremos con el establecimiento de la representación del proceso. Digamos que tenemos un proceso estocástico \mathbf{X}_t estacionario vectorial de dimensión k :

$$\mathbf{X}_t = \begin{bmatrix} X_{1t} \\ X_{2t} \\ \vdots \\ X_{kt} \end{bmatrix}$$

Para cualquier $i = 1, 2, \dots, p$:

$$\mathbf{X}_{t-i} = \begin{bmatrix} X_{1t-i} \\ X_{2t-i} \\ \vdots \\ X_{kt-i} \end{bmatrix}$$

Donde cada X_{kt} en \mathbf{X}_t es una serie de tiempo por sí misma. De esta forma, la expresión reducida del modelo o el proceso $VAR(p)$ estará dado por:

$$\mathbf{X}_t = \delta + A_1 \mathbf{X}_{t-1} + A_2 \mathbf{X}_{t-2} + \dots + A_p \mathbf{X}_{t-p} + \mathbf{U}_t \quad (5.7)$$

Donde cada uno de las A_i , $i = 1, 2, \dots, p$, son matrices cuadradas de dimensión k y \mathbf{U}_t representa un vector de dimensión $k \times 1$ con los residuales en el momento del tiempo t que son, por individual, un proceso puramente aleatorio. También se incorpora un vector de términos constantes denominado como δ , el cual es de dimensión $k \times 1$.

Así, la ecuación (5.7) supone la siguiente estructura del vector δ :

$$\delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_k \end{bmatrix}$$

También, la ecuación (5.7) supone que cada matriz A_i , $i = 1, 2, \dots, p$ está definida de la siguiente forma:

$$\mathbf{A}_i = \begin{bmatrix} a_{11}^{(i)} & a_{12}^{(i)} & \dots & a_{1k}^{(i)} \\ a_{21}^{(i)} & a_{22}^{(i)} & \dots & a_{2k}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1}^{(i)} & a_{k2}^{(i)} & \dots & a_{kk}^{(i)} \end{bmatrix}$$

Donde $i = 1, 2, \dots, p$.

Retomando la ecuación (5.7) y considerando que podemos ocupar el operador rezago L^j de forma análoga al caso del modelo $AR(p)$, pero aplicado a un vector, tenemos las siguientes ecuaciones:

$$\begin{aligned} \mathbf{X}_t - A_1 \mathbf{X}_{t-1} - A_2 \mathbf{X}_{t-2} - \dots - A_p \mathbf{X}_{t-p} &= \delta + \mathbf{U}_t \\ \mathbf{X}_t - A_1 L \mathbf{X}_t - A_2 L^2 \mathbf{X}_t - \dots - A_p L^p \mathbf{X}_t &= \delta + \mathbf{U}_t \\ (I_k - \mathbf{A}_1 L - \mathbf{A}_2 L^2 - \dots - \mathbf{A}_p L^p) \mathbf{X}_t &= \delta + \mathbf{U}_t \\ \mathbf{A}(L) \mathbf{X}_t &= \delta + \mathbf{U}_t \end{aligned} \quad (5.8)$$

Adicionalmente, requeriremos que dado que \mathbf{U}_t es un proceso puramente aleatorio, este debe cumplir con las siguientes condiciones:

1. El valor esperado del término de error es cero:

$$\mathbb{E}[\mathbf{U}_t] = 0 \quad (5.9)$$

2. Existe una matriz de varianzas y covarianzas entre los términos de error

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

contemporáneos dada por:

$$\begin{aligned}
 \mathbb{E}[\mathbf{U}_t \mathbf{U}_t'] &= \mathbb{E} \left[\begin{bmatrix} U_1^{(t)} \\ U_2^{(t)} \\ \vdots \\ U_k^{(t)} \end{bmatrix} \begin{bmatrix} U_1^{(t)} & U_2^{(t)} & \dots & U_k^{(t)} \end{bmatrix} \right] \\
 &= \mathbb{E} \begin{bmatrix} U_1^{(t)} U_1^{(t)} & U_1^{(t)} U_2^{(t)} & \dots & U_1^{(t)} U_k^{(t)} \\ U_2^{(t)} U_1^{(t)} & U_2^{(t)} U_2^{(t)} & \dots & U_2^{(t)} U_k^{(t)} \\ \vdots & \vdots & \dots & \vdots \\ U_k^{(t)} U_1^{(t)} & U_k^{(t)} U_2^{(t)} & \dots & U_k^{(t)} U_k^{(t)} \end{bmatrix} \\
 &= \begin{bmatrix} \sigma_1^2 & \rho_{12} & \dots & \rho_{1k} \\ \rho_{21} & \sigma_2^2 & \dots & \rho_{2k} \\ \vdots & \vdots & \dots & \vdots \\ \rho_{k1} & \rho_{k2} & \dots & \sigma_k^2 \end{bmatrix} \\
 &= \mathbf{U} \mathbf{U}
 \end{aligned} \tag{5.10}$$

3. La matriz de varianzas y covarianzas no contemporáneas es nula. Es decir, que para todo $t \neq s$:

$$\begin{aligned}
 \mathbb{E}[\mathbf{U}_t \mathbf{U}_s'] &= \mathbb{E} \left[\begin{bmatrix} U_1^{(t)} \\ U_2^{(t)} \\ \vdots \\ U_k^{(t)} \end{bmatrix} \begin{bmatrix} U_1^{(s)} & U_2^{(s)} & \dots & U_k^{(s)} \end{bmatrix} \right] \\
 &= \mathbb{E} \begin{bmatrix} U_1^{(t)} U_1^{(s)} & U_1^{(t)} U_2^{(s)} & \dots & U_1^{(t)} U_k^{(s)} \\ U_2^{(t)} U_1^{(s)} & U_2^{(t)} U_2^{(s)} & \dots & U_2^{(t)} U_k^{(s)} \\ \vdots & \vdots & \dots & \vdots \\ U_k^{(t)} U_1^{(s)} & U_k^{(t)} U_2^{(s)} & \dots & U_k^{(t)} U_k^{(s)} \end{bmatrix} \\
 &= \mathbf{0}
 \end{aligned} \tag{5.11}$$

Las ecuaciones (5.10) y (5.11) significan que los residuales \mathbf{U}_t pueden estar correlacionados entre ellos solo en el caso de que la información sea contemporánea, pero no tienen información en común entre residuales de otros periodos.

Al igual que en el caso del modelo o especificación $AR(p)$ en la especificación del modelo $VAR(p)$ existen condiciones de estabilidad. Dichas condiciones están dadas por lo siguiente, definamos el siguiente polinomio que resulta de tomar la matriz $\mathbf{A}(L)$ en la ecuación (5.8):

$$\text{Det}[I_t - A_1 z - A_2 z^2 - \dots - A_p z^p] \neq 0 \tag{5.12}$$

Donde las raíces del polinomio cumplen que $|z| \leq 1$, es decir, se ubican dentro del círculo unitario.

La ecuación (5.8) puede ser reexpresada en una forma similar al un proceso de MA. Al respecto, de forma similar a la siguiente ecuación podemos construir un modelo $VARMA(p, q)$, el cual no estudiamos en este curso.

Recomiendo el primer planteamiento, podemos escribir:

$$\begin{aligned}\mathbf{X}_t &= \mathbf{A}^{-1}(L)\delta + \mathbf{A}^{-1}(L)\mathbf{U}_t \\ &= \mu + \beta(L)\mathbf{U}_t\end{aligned}\quad (5.13)$$

Donde μ es un vector de $k \times 1$ constantes y $\beta(L)$ es una matriz que depende de L .

Por el lado de las matrices que representan la autocovarianza, éstas resultan de resolver lo siguiente:

$$\Gamma_X(\tau) = E[(\mathbf{X}_t - \mu)(\mathbf{X}_{t-\tau} - \mu)'] \quad (5.14)$$

Ahora, sin pérdida de generalidad digamos que la especificación $VAR(p)$ en la ecuación (5.7) no tiene constante, por lo que $\delta = \mathbf{0}$, lo que implica que $\mu = \mathbf{0}$. De esta forma las matrices de autocovarianza resultan de:

$$\begin{aligned}\Gamma_{\mathbf{X}}(\tau) &= E[(\mathbf{X}_t)(\mathbf{X}_{t-\tau})'] \\ &= \mathbf{A}_1 E[(\mathbf{X}_{t-1})(\mathbf{X}_{t-\tau})'] + \mathbf{A}_2 E[(\mathbf{X}_{t-2})(\mathbf{X}_{t-\tau})'] \\ &\quad + \dots + \mathbf{A}_p E[(\mathbf{X}_{t-p})(\mathbf{X}_{t-\tau})'] + E[(\mathbf{U}_t)(\mathbf{X}_{t-\tau})']\end{aligned}$$

Finalmente, al igual que en el caso $AR(p)$, requerimos de una métrica que nos permita determinar el número de rezagos óptimo p en el $VAR(p)$. Así, establecemos criterios de información similares a los del $AR(p)$ dados por:

1. Final Prediction Error (FPE):

$$FPE(p) = \left[\frac{T + kp + 1}{T - kp - 1} \right]^k |\hat{\mathbf{U}}\hat{\mathbf{U}}(p)| \quad (5.15)$$

2. Akaike Criterion (AIC):

$$AIC(p) = \ln |\hat{\mathbf{U}}\hat{\mathbf{U}}(p)| + (k + pk^2) \frac{2}{T} \quad (5.16)$$

3. Hannan - Quinn Criterion (HQ):

$$HQ(p) = \ln |\hat{\mathbf{U}}\hat{\mathbf{U}}(p)| + (k + pk^2) \frac{2 \ln(\ln(2))}{T} \quad (5.17)$$

4. Schwartz Criterion (SC):

$$SC(p) = \ln |\hat{\mathbf{U}}\hat{\mathbf{U}}(p)| + (k + pk^2) \frac{\ln(T)}{T} \quad (5.18)$$

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

Donde la matriz de varianzas y covarianzas contemporáneas estará dada por:

$$\hat{U}\hat{U}(p) = \mathbb{E} \left[\begin{bmatrix} U_1^{(t)} \\ U_2^{(t)} \\ \vdots \\ U_k^{(t)} \end{bmatrix} \begin{bmatrix} U_1^{(t)} & U_2^{(t)} & \dots & U_k^{(t)} \end{bmatrix} \right]$$

Ejemplo. Ahora veámos un ejemplo de estimación de $VAR(p)$. Para el ejemplo utilizaremos las series de INPC, Tipo de Cambio, rendimiento de los Cetes a 28 días, el IGAE y el Índice de Producción Industrial de los Estados Unidos, todas desestacionalizadas y para el período de enero de 2000 a julio de 2019. Dado que el supuesto estacionariedad sigue presente en nuestro análisis, emplearemos cada una de las series en su versión de diferencias logaritmicas. Las Figuras 5.2 y 5.3 muestra las series referidas.

```
library(ggplot2)
library(dplyr)
library(stats)
library(MASS)
library(strucchange)
library(zoo)
library(sandwich)
library(urca)
library(lmtest)
library(vars)

#
load("BD/Datos_Ad.RData")

#
DLINPC <- diff(log( ts(Datos_Ad$INPC_Ad, start = c(2000, 1), freq = 12) ))

DLTC <- diff(log( ts(Datos_Ad$TC_Ad, start = c(2000, 1), freq = 12) ))

DLCETE28 <- diff(log( ts(Datos_Ad$CETE28_Ad, start = c(2000, 1), freq = 12) ))

DLIGAE <- diff(log( ts(Datos_Ad$IGAE_Ad, start = c(2000, 1), freq = 12) ))

DLIPI <- diff(log( ts(Datos_Ad$IPI_Ad, start = c(2000, 1), freq = 12) ))

Datos <- data.frame(cbind(DLINPC, DLTC, DLCETE28, DLIGAE, DLIPI))

Datos <- ts(Datos,
            start = c(2000, 2), freq = 12)
```

```
plot(Datos, plot.type = "s",
     col = c("darkgreen", "darkblue", "darkred", "black", "purple"),
     main = "Series en Diferencias logaritmicas",
     xlab = "Tiempo", ylab = "Variacion")

legend("bottomright", c("INPC", "TC", "CETES28", "IGAE", "IPI"),
     cex = 0.6, lty = 1:1,
     col = c("darkgreen", "darkblue", "darkred", "black", "purple"))
```

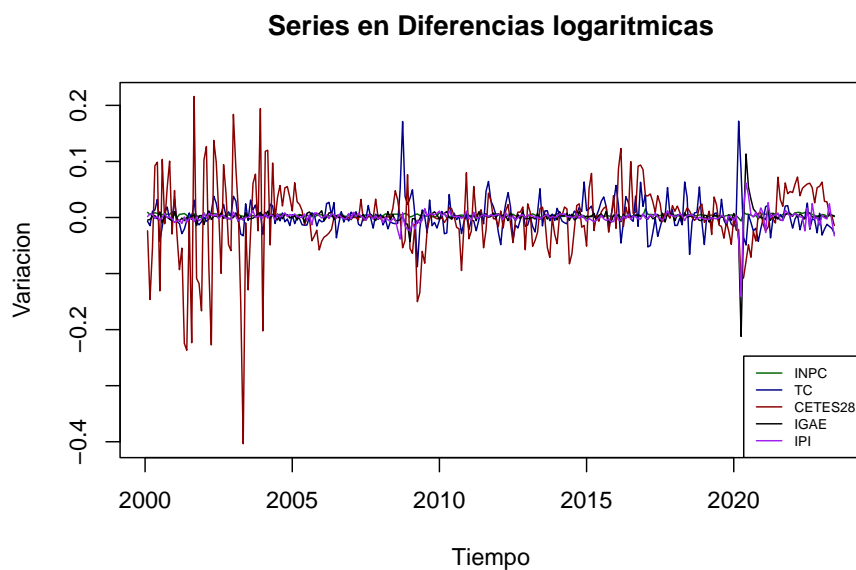


Figure 5.2: Series en diferencias logarítmicas (Forma 1)

```
plot(Datos, plot.type = "m",
     col = "darkgreen",
     main = "Series en Diferencias logaritmicas", xlab = "Tiempo")
```

Dicho lo anterior, a continuación mostraremos la tabla que resume el valor de los distintos criterios de información para una especificación de un $VAR(p)$ con constante. Nótese que es posible especificar un $VAR(p)$ con tendencia, siempre que exista evidencia de que algunas de las series sean estacionarias alrededor de una tendencia. Caso que no aplica hasta este momento, ya que nuestro análisis de estacionariedad es claro respecto a la media constante (más adelante aportaremos la evidencia de esto), lo cual elimina la posibilidad de incluir una tendencia.

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

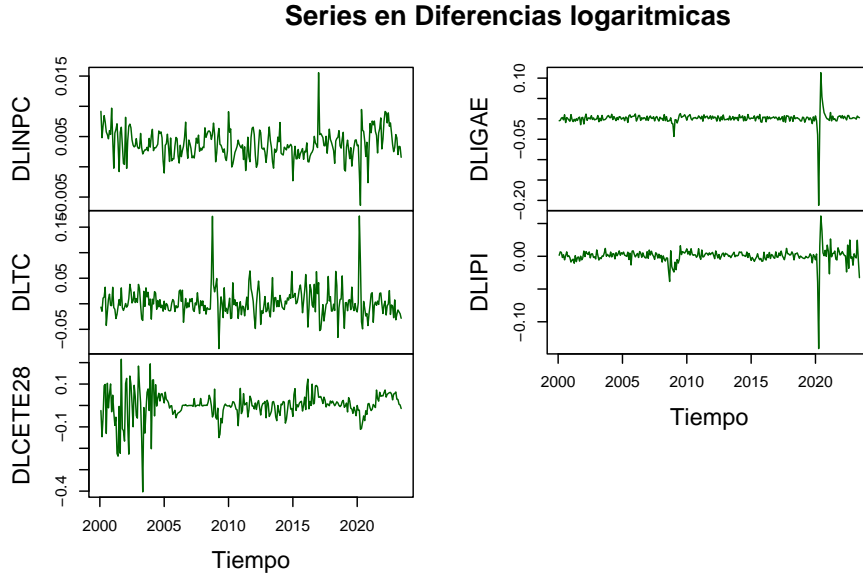


Figure 5.3: Series en diferencias logarítmicas (Forma 2)

En el Cuadro 5.2 reportamos el número de rezagos propuesto a partir de cada criterio de información y en el Cuadro 5.3 reportamos los resultados de aplicar una prueba de criterios de información para diferentes valores de rezagos. Del cual se concluye que el número óptimo de rezagos es 2 (según el criterio AIC y el FPE) y 1 (según el criterio HQ y el SC). Recordemos que es común que el criterio AIC siempre reporte el mayor valor de rezagos, por lo que es una buena práctica utilizarlo como referente principal.

Table 5.2: Número de rezagos determinados por cada uno de los criterios de información para diferentes especificaciones de modelos $VAR(p)$ con término constante de la series $DLINPC_t$, $DLTC_t$, $DLCETE28_t$, $DLIGAE_t$ y $DLIPI_t$.

AIC	HQ	SC	FPE
2	1	1	2

Table 5.3: Criterios de información para diferentes especificaciones de modelos VAR(p) con término constante de la serie $DLINPC_t$, $DLTC_t$, $DLCEFE28_t$, $DLIGAE_t$ y $DLIPI_t$.

Rezagos	AIC	HQ	SC	FPE
1	-4.636412e+01	-4.617847e+01	-4.590430e+01	7.317262e-21
2	-4.639541e+01	-4.605506e+01	-4.555241e+01	7.094216e-21
3	-4.635305e+01	-4.585799e+01	-4.512686e+01	7.407479e-21
⋮	⋮	⋮	⋮	⋮

```
VARselect(Datos, lag.max = 12, type = "const")
```

```
## $selection
## AIC(n) HQ(n) SC(n) FPE(n)
##      2      1      1      2
##
## $criteria
##                                1                                2
## AIC(n) -43.4369385930444451560106245 -43.4878280980991220872056147
## HQ(n)  -43.2759374580961733158801508 -43.1926593506939511257769482
## SC(n)  -43.0360414131631969780755753 -42.7528499349835016118959174
## FPE(n)  0.000000000000000001366449  0.000000000000000001298899
##                                3                                4
## AIC(n) -43.4268489491758202802884625 -43.3868843889051234441467386
## HQ(n)  -42.9975125893137573029889609 -42.8233804165861613455490442
## SC(n)  -42.3577898028258275076041173 -41.9837442593207583740877453
## FPE(n)  0.000000000000000001381225  0.000000000000000001438821
##                                5                                6
## AIC(n) -43.293163285544032703455741 -43.2431704133633161291072611
## HQ(n)  -42.595491700768178588987212 -42.4113312161305699987678963
## SC(n)  -41.555942172725302441449458 -41.1718683173102135697263293
## FPE(n)  0.00000000000000000158246  0.000000000000000001667106
##                                7                                8
## AIC(n) -43.2902076769675190348607430 -43.2419352439015938216471113
## HQ(n)  -42.3242008672778737832231855 -42.1417608217550565541387186
## SC(n)  -40.8848245976800441781051632 -40.5024711813797466675168835
## FPE(n)  0.000000000000000001595179  0.000000000000000001680601
##                                9                                10
## AIC(n) -43.2201406224138153788771888 -43.1788072276712782127106038
## HQ(n)  -41.9857985878103789900706033 -41.8102975806109569134605408
## SC(n)  -40.1465955766575959273723129 -39.7711811986806935692584375
## FPE(n)  0.000000000000000001726236  0.000000000000000001810365
##                                11                                12
## AIC(n) -43.1191364389135785017970193 -43.0502244962885995960277796
```

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

```
## HQ(n) -41.6164591793963580812487635 -41.4133796243144800541813311
## SC(n) -39.3774294266886144555428473 -38.9744365008292632523989596
## FPE(n) 0.0000000000000000001936459 0.0000000000000000002093856
```

De esta forma, justificamos la estimación de un $VAR(2)$. Los resultados del mismo se reportan en los siguientes cuadros, en los que se muestra el resultado de una de las ecuaciones. Los resultados restantes se encuentran en el código de R mostrado más abajo. Primero mostraremos los resultados de las raíces del polinomio característico en el Cuadro 5.4, seguido de un cuadro para la ecuación del IGAE en el Cuadro 5.5 (por simplicidad se omiten las otras cuatro ecuaciones del $VAR(2)$), y del Cuadro 5.6 con la matriz $\hat{U}\hat{U}$ estimada del VAR.

Table 5.4: Raíces del polinomio característico de un $VAR(2)$.

0.7452	0.4403	0.4403	0.3503	0.3503
0.3342	0.3342	0.3339	0.3339	0.06951

Table 5.5: Criterios de información para diferentes especificaciones de modelos $VAR(p)$ con término constante de la series $DLINPC_t$, $DLTC_t$, $DLCETE28_t$, $DLIGAE_t$ y $DLIPI_t$.

Variable	Coficiente	Error Est.	Estad. t	Prob.(> t)
$DLINPC_{t-1}$	-0.2584978	0.1658396	-1.559	0.120493
$DLTC_{t-1}$	0.0022016	0.0152876	0.144	0.885620
$DLCETE28_{t-1}$	0.0009547	0.0049115	0.194	0.846054
$DLIGAE_{t-1}$	-0.2351453	0.0699797	-3.360	0.000917 ***
$DLIPI_{t-1}$	0.2442406	0.0600502	4.067	6.62e-05 ***
$DLINPC_{t-2}$	-0.0775039	0.1694809	-0.457	0.647904
$DLTC_{t-2}$	-0.0413316	0.0144650	-2.857	0.004680 **
$DLCETE28_{t-2}$	0.0005341	0.0048058	0.111	0.911612
$DLIGAE_{t-2}$	-0.0646890	0.0693711	-0.933	0.352092
$DLIPI_{t-2}$	0.1796286	0.0620861	2.893	0.004195 **
δ_4	0.0030377	0.0008077	3.761	0.000217 ***
Notas:		***	signif.	al 0.1%
		**	signif.	al 1%
		*	signif.	al 5%

Table 5.6: Matriz $\hat{U}\hat{U}$ estimada del $VAR(2)$.

	$DLINPC_t$	$DLTC_t$	$DLCE28_t$	$DLIGAE_t$	$DLIGAE_t$
$DLINPC_t$	3.95e-06	3.19e-06	-1.83e-06	-5.29-07	1.34e-06
$DLTC_t$	3.19e-06	5.04e-04	4.27e-04	9.81e-06	1.61e-05

$DLCE28_t$	-1.83e-06	4.27e-04	4.63e-03	1.26e-05	2.76e-05
$DLIGAE_t$	-5.29e-07	9.81e-06	1.26e-05	2.43e-05	8.75e-06
$DLIGAE_t$	1.34e-06	1.61e-05	2.76e-05	8.75e-06	3.13e-05

```
VAR_p <- VAR(Datos, p = 2, type = "const")
```

```
summary(VAR_p)
```

```
##
## VAR Estimation Results:
## =====
## Endogenous variables: DLINPC, DLTC, DLCETE28, DLIGAE, DLIPI
## Deterministic variables: const
## Sample size: 279
## Log Likelihood: 4142.895
## Roots of the characteristic polynomial:
## 0.53 0.53 0.4501 0.4501 0.4425 0.4425 0.3251 0.3251 0.1677 0.1677
## Call:
## VAR(y = Datos, p = 2, type = "const")
##
##
## Estimation results for equation DLINPC:
## =====
## DLINPC = DLINPC.l1 + DLTC.l1 + DLCETE28.l1 + DLIGAE.l1 + DLIPI.l1 + DLINPC.l2 + DLTC.l2 + DLCETE28.l2 + DLIGAE.l2 + DLIPI.l2 + const
##
##              Estimate Std. Error t value      Pr(>|t|)
## DLINPC.l1    0.387061   0.063100   6.134 0.0000000030610259 ***
## DLTC.l1      -0.004611   0.005286  -0.872      0.3839
## DLCETE28.l1  0.001229   0.002078   0.592      0.5546
## DLIGAE.l1    -0.022088   0.015077  -1.465      0.1441
## DLIPI.l1      0.009878   0.019363   0.510      0.6104
## DLINPC.l2    -0.006707   0.063621  -0.105      0.9161
## DLTC.l2       0.009608   0.005496   1.748      0.0816 .
## DLCETE28.l2  0.001827   0.002051   0.891      0.3737
## DLIGAE.l2     0.006544   0.014289   0.458      0.6473
## DLIPI.l2     -0.015245   0.019632  -0.777      0.4381
## const         0.002325   0.000294   7.908 0.0000000000000688 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.002176 on 268 degrees of freedom
## Multiple R-Squared: 0.1783, Adjusted R-squared: 0.1476
## F-statistic: 5.815 on 10 and 268 DF, p-value: 0.00000006289
```

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

```
##
##
## Estimation results for equation DLTC:
## =====
## DLTC = DLINPC.11 + DLTC.11 + DLCETE28.11 + DLIGAE.11 + DLIPI.11 + DLINPC.12 + DLTC.12 + DLCETE
##
##          Estimate Std. Error t value   Pr(>|t|)
## DLINPC.11  -1.344803   0.745382  -1.804   0.07233 .
## DLTC.11     0.315786   0.062442   5.057 0.00000079 ***
## DLCETE28.11 -0.039912   0.024542  -1.626   0.10506
## DLIGAE.11   0.379306   0.178102   2.130   0.03411 *
## DLIPI.11    -0.514827   0.228727  -2.251   0.02521 *
## DLINPC.12   0.071779   0.751538   0.096   0.92398
## DLTC.12     -0.183813   0.064926  -2.831   0.00499 **
## DLCETE28.12  0.030125   0.024228   1.243   0.21481
## DLIGAE.12   0.070713   0.168797   0.419   0.67561
## DLIPI.12    -0.166610   0.231909  -0.718   0.47312
## const       0.006463   0.003473   1.861   0.06388 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.0257 on 268 degrees of freedom
## Multiple R-Squared: 0.1425, Adjusted R-squared: 0.1105
## F-statistic: 4.454 on 10 and 268 DF, p-value: 0.000008102
##
##
## Estimation results for equation DLCETE28:
## =====
## DLCETE28 = DLINPC.11 + DLTC.11 + DLCETE28.11 + DLIGAE.11 + DLIPI.11 + DLINPC.12 + DLTC.12 + DL
##
##          Estimate Std. Error t value Pr(>|t|)
## DLINPC.11   3.933798   1.861561   2.113   0.0355 *
## DLTC.11     0.112997   0.155945   0.725   0.4693
## DLCETE28.11  0.113300   0.061292   1.849   0.0656 .
## DLIGAE.11   -0.588758   0.444803  -1.324   0.1868
## DLIPI.11     1.174496   0.571237   2.056   0.0407 *
## DLINPC.12   -2.213087   1.876936  -1.179   0.2394
## DLTC.12     0.117661   0.162151   0.726   0.4687
## DLCETE28.12  0.063398   0.060509   1.048   0.2957
## DLIGAE.12   -0.237974   0.421565  -0.565   0.5729
## DLIPI.12     1.110203   0.579183   1.917   0.0563 .
## const       -0.007514   0.008674  -0.866   0.3872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```

##
## Residual standard error: 0.06419 on 268 degrees of freedom
## Multiple R-Squared: 0.08823, Adjusted R-squared: 0.05421
## F-statistic: 2.593 on 10 and 268 DF, p-value: 0.005122
##
##
## Estimation results for equation DLIGAE:
## =====
## DLIGAE = DLINPC.11 + DLTC.11 + DLCETE28.11 + DLIGAE.11 + DLIPI.11 + DLINPC.12 + DLTC.12
##
##              Estimate Std. Error t value      Pr(>|t|)
## DLINPC.11    0.719855   0.393617   1.829      0.0685 .
## DLTC.11      -0.218508   0.032974  -6.627 0.000000000188 ***
## DLCETE28.11  0.007718   0.012960   0.596      0.5520
## DLIGAE.11    0.037559   0.094051   0.399      0.6900
## DLIPI.11     0.389894   0.120785   3.228      0.0014 **
## DLINPC.12   -0.737653   0.396868  -1.859      0.0642 .
## DLTC.12      0.073830   0.034286   2.153      0.0322 *
## DLCETE28.12 -0.013542   0.012794  -1.058      0.2908
## DLIGAE.12   -0.142845   0.089138  -1.603      0.1102
## DLIPI.12    -0.154835   0.122465  -1.264      0.2072
## const       0.001637   0.001834   0.893      0.3729
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.01357 on 268 degrees of freedom
## Multiple R-Squared: 0.3272, Adjusted R-squared: 0.3021
## F-statistic: 13.03 on 10 and 268 DF, p-value: < 0.00000000000000022
##
##
## Estimation results for equation DLIPI:
## =====
## DLIPI = DLINPC.11 + DLTC.11 + DLCETE28.11 + DLIGAE.11 + DLIPI.11 + DLINPC.12 + DLTC.12
##
##              Estimate Std. Error t value      Pr(>|t|)
## DLINPC.11    0.042383   0.326611   0.130      0.8968
## DLTC.11      -0.163742   0.027361  -5.985 0.000000000693 ***
## DLCETE28.11  0.012143   0.010754   1.129      0.2598
## DLIGAE.11    0.098821   0.078041   1.266      0.2065
## DLIPI.11     0.067742   0.100224   0.676      0.4997
## DLINPC.12   -0.524835   0.329309  -1.594      0.1122
## DLTC.12      0.053353   0.028449   1.875      0.0618 .
## DLCETE28.12 -0.016237   0.010616  -1.529      0.1273
## DLIGAE.12   -0.088483   0.073964  -1.196      0.2326
## DLIPI.12    -0.082982   0.101618  -0.817      0.4149

```

5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

```
## const          0.002361    0.001522    1.551          0.1220
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##
## Residual standard error: 0.01126 on 268 degrees of freedom
## Multiple R-Squared: 0.1978, Adjusted R-squared: 0.1679
## F-statistic: 6.61 on 10 and 268 DF, p-value: 0.000000003715
##
##
##
## Covariance matrix of residuals:
##              DLINPC          DLTC          DLCETE28          DLIGAE          DLIPI
## DLINPC    0.0000047345 -0.0000001939 0.000004559 0.000006652 0.00000700
## DLTC      -0.0000001939 0.0006606619 0.000369751 -0.000035441 -0.00003174
## DLCETE28 0.0000045591 0.0003697510 0.004120747 0.000063618 0.00008525
## DLIGAE    0.0000066522 -0.0000354410 0.000063618 0.000184234 0.00011641
## DLIPI     0.0000069998 -0.0000317443 0.000085254 0.000116407 0.00012685
##
## Correlation matrix of residuals:
##              DLINPC          DLTC          DLCETE28          DLIGAE          DLIPI
## DLINPC    1.000000 -0.003466 0.03264 0.22524 0.2856
## DLTC      -0.003466 1.000000 0.22409 -0.10159 -0.1097
## DLCETE28 0.032640 0.224095 1.00000 0.07301 0.1179
## DLIGAE    0.225240 -0.101585 0.07301 1.00000 0.7615
## DLIPI     0.285629 -0.109656 0.11792 0.76147 1.0000
```

Finalmente, en el Cuadro 5.7 reportamos las pruebas de diagnóstico del VAR(2). Incluimos las pruebas de correlación serial (o autocorrelación), normalidad y de heterocedasticidad. De acuerdo con esa información, la correlación serial muestra que existe relación de la matriz de covarianzas no contemporánea considerando pocos rezagos, pero se elimina conforme los rezagos se incrementan. En cuanto a normalidad, se observa que los residuales no lo son, por lo que se requeriría mejorar la especificación del VAR. Finalmente, se observa que los residuales no son homocedásticos.

Table 5.7: Pruebas de diagnóstico sobre los residuales del VAR(2).

Estadística (rezagos)	Coefficiente	p-value	Conclusión
Correlación Serial ($\chi^2(2)$)	59.436	0.1696	Existe autocorrelación serial

Estadística (rezagos)	Coficiente	p-value	Conclusión
Correlación Serial ($\chi^2(4)$)	127.17	0.03461	No existe autocorrelación serial
Correlación Serial ($\chi^2(6)$)	183.14	0.03393	No existe autocorrelación serial
Normalidad - JB (χ^2)	2335	0.0000	Los residuales no son normales
ARCH ($\chi^2(2)$)	691.58	0.0000	Los residuales no son homocedásticos

Diagnostic tests

Normalidad:

```
normality.test(VAR_p)
```

```
## $JB
##
## JB-Test (multivariate)
##
## data: Residuals of VAR object VAR_p
## Chi-squared = 36186, df = 10, p-value < 0.00000000000000022
##
##
## $Skewness
##
## Skewness only (multivariate)
##
## data: Residuals of VAR object VAR_p
## Chi-squared = 1261, df = 5, p-value < 0.00000000000000022
##
##
## $Kurtosis
##
## Kurtosis only (multivariate)
##
## data: Residuals of VAR object VAR_p
## Chi-squared = 34925, df = 5, p-value < 0.00000000000000022
```


5.2. DEFINICIÓN Y REPRESENTACIÓN DEL SISTEMA O MODELO DE VECTORES AUTORREGRESIVOS (

```
#### Autocorrelacion Serial:
```

```
#### LAGS = 2:
```

```
serial.test(VAR_p, lags.bg = 2, type = "BG")
```

```
##  
## Breusch-Godfrey LM test  
##  
## data: Residuals of VAR object VAR_p  
## Chi-squared = 74.839, df = 50, p-value = 0.013
```

```
#### LAGS = 4:
```

```
serial.test(VAR_p, lags.bg = 4, type = "BG")
```

```
##  
## Breusch-Godfrey LM test  
##  
## data: Residuals of VAR object VAR_p  
## Chi-squared = 131.69, df = 100, p-value = 0.01849
```

```
#### LAGS = 6:
```

```
serial.test(VAR_p, lags.bg = 6, type = "BG")
```

```
##  
## Breusch-Godfrey LM test  
##  
## data: Residuals of VAR object VAR_p  
## Chi-squared = 201.63, df = 150, p-value = 0.003147
```

```
#### Homocedasticidad:
```

```
arch.test(VAR_p, lags.multi = 6)
```

```
##  
## ARCH (multivariate)  
##  
## data: Residuals of VAR object VAR_p  
## Chi-squared = 2122.4, df = 1350, p-value < 0.00000000000000022
```

5.3 Análisis de Impulso-Respuesta

Una de las grandes ventajas que aporta el análisis de los modelos VAR es el análisis de Impulso-Respuesta. Dicho análisis busca cuantificar el efecto que tiene en \mathbf{X}_t una innovación o cambio en los residuales de cualquiera de las variables en un momento definido. Partamos de la ecuación (5.13) de forma que tenemos:

$$\begin{aligned}\mathbf{X}_t &= \mathbf{A}^{-1}(L)\delta + \mathbf{A}^{-1}(L)\mathbf{U}_t \\ &= \mu + \mathbf{B}(L)\mathbf{U}_t \\ &= \mu + \Psi_0\mathbf{U}_t + \Psi_1\mathbf{U}_{t-1} + \Psi_2\mathbf{U}_{t-2} + \Psi_3\mathbf{U}_{t-3} + \dots\end{aligned}\quad (5.19)$$

Donde $\Psi_0 = I$ y cada una de las $\Psi_i = -\mathbf{B}_i$, $i = 1, 2, \dots$. De esta forma se verifica el efecto que tiene en \mathbf{X}_t cada una de las innovaciones pasadas. Por lo que el análisis de Impulso-Respuesta cuantifica el efecto de cada una de esas matrices en las que hemos descompuesto a $\mathbf{B}(L)$.

Ejemplo. Retomando el modelo $VAR(2)$ anteriormente estimado, en las siguientes figuras reportamos las gráficas de Impulso-Respuesta de la serie $DLTC_t$ ante cambios en los residuales del resto de las series y de la propia serie.

```
IR_DLINPC <- irf(VAR_p, n.ahead = 12, boot = TRUE,
                 ci = 0.95, response = "DLINPC")
```

```
IR_DLINPC
```

```
##
## Impulse response coefficients
## $DLINPC
##           DLINPC
## [1,] 0.0021759002410
## [2,] 0.0008094393250
## [3,] 0.0002429936391
## [4,] 0.0001048999556
## [5,] 0.0000430519256
## [6,] 0.0000257195308
## [7,] 0.0000108405478
## [8,] 0.0000011471582
## [9,] -0.0000014256627
## [10,] -0.0000009999499
## [11,] -0.0000003443770
## [12,] -0.0000001512886
## [13,] -0.0000001157354
##
## $DLTC
```

```
##                               DLINPC
## [1,] 0.000000000000000
## [2,] -0.0000826837093
## [3,] 0.0003117546911
## [4,] 0.0002648974979
## [5,] 0.0000587162871
## [6,] -0.0000289491380
## [7,] -0.0000105463736
## [8,] 0.0000105785893
## [9,] 0.0000076216336
## [10,] -0.0000004038885
## [11,] -0.0000021138204
## [12,] -0.0000004691794
## [13,] 0.0000003771599
```

```
##
## $DLCETE28
##                               DLINPC
## [1,] 0.000000000000000
## [2,] 0.00006486928725
## [3,] 0.00013182636385
## [4,] 0.00003487943931
## [5,] 0.00004740702286
## [6,] 0.00003336371156
## [7,] 0.00001571609856
## [8,] 0.00000377046957
## [9,] 0.00000005672766
## [10,] 0.00000025569500
## [11,] 0.00000033794064
## [12,] 0.00000004626880
## [13,] -0.00000011376620
```

```
##
## $DLIGAE
##                               DLINPC
## [1,] 0.000000000000000
## [2,] -0.00021156780483
## [3,] -0.00017900748374
## [4,] -0.00001453803984
## [5,] 0.00004387550079
## [6,] 0.00001348412556
## [7,] -0.00000660541189
## [8,] -0.00000344927548
## [9,] 0.00000223865243
## [10,] 0.00000206681740
## [11,] 0.00000008258756
## [12,] -0.00000046595182
## [13,] -0.00000009698431
```

```

##
## $DLIPI
##          DLINPC
## [1,]  0.000000000000000
## [2,]  0.00007038187903
## [3,] -0.00011077803202
## [4,] -0.00003218007799
## [5,] -0.00001697303009
## [6,]  0.00000048565947
## [7,]  0.00001045292187
## [8,]  0.00000469541763
## [9,]  0.00000035996398
## [10,] -0.00000051048646
## [11,]  0.00000005294854
## [12,]  0.00000026647657
## [13,]  0.00000009015073
##
##
## Lower Band, CI= 0.95
## $DLINPC
##          DLINPC
## [1,]  0.001864170847
## [2,]  0.000499776775
## [3,]  0.000002450675
## [4,] -0.000078104730
## [5,] -0.000075950781
## [6,] -0.000037958817
## [7,] -0.000025483439
## [8,] -0.000012332298
## [9,] -0.000011329253
## [10,] -0.000007641493
## [11,] -0.000004043746
## [12,] -0.000001841703
## [13,] -0.000001172170
##
## $DLTC
##          DLINPC
## [1,]  0.000000000000000
## [2,] -0.000286990353
## [3,]  0.000061870696
## [4,]  0.000078542144
## [5,] -0.000039820370
## [6,] -0.000095870183
## [7,] -0.000050900347
## [8,] -0.000019789658
## [9,] -0.000005574237

```

```
## [10,] -0.000008547257
## [11,] -0.000009189057
## [12,] -0.000004530429
## [13,] -0.000001945539
##
```

```
## $DLCETE28
```

```
##          DLINPC
## [1,]  0.000000000000
## [2,] -0.000230649362
## [3,] -0.000196536071
## [4,] -0.000126449052
## [5,] -0.000058700460
## [6,] -0.000023095154
## [7,] -0.000007659372
## [8,] -0.000008557329
## [9,] -0.000008242220
## [10,] -0.000005060508
## [11,] -0.000002432306
## [12,] -0.000001136291
## [13,] -0.000001066601
##
```

```
## $DLIGAE
```

```
##          DLINPC
## [1,]  0.000000000000
## [2,] -0.000410323921
## [3,] -0.000418493922
## [4,] -0.000146714896
## [5,] -0.000044408954
## [6,] -0.000035327961
## [7,] -0.000035119635
## [8,] -0.000022423812
## [9,] -0.000012260613
## [10,] -0.000003350492
## [11,] -0.000003212534
## [12,] -0.000004039474
## [13,] -0.000002483970
##
```

```
## $DLIPI
```

```
##          DLINPC
## [1,]  0.000000000000
## [2,] -0.0001982996896
## [3,] -0.0003103493634
## [4,] -0.0001445269721
## [5,] -0.0001147657072
## [6,] -0.0000494580078
## [7,] -0.0000148574620
```

```

## [8,] -0.0000103989132
## [9,] -0.0000059299583
## [10,] -0.0000046364719
## [11,] -0.0000029746270
## [12,] -0.0000012355125
## [13,] -0.0000008043852
##
##
## Upper Band, CI= 0.95
## $DLINPC
##          DLINPC
## [1,] 0.002423055334
## [2,] 0.001069138548
## [3,] 0.000514403106
## [4,] 0.000299906999
## [5,] 0.000185297033
## [6,] 0.000101768840
## [7,] 0.000056864142
## [8,] 0.000028728717
## [9,] 0.000018624195
## [10,] 0.000012817976
## [11,] 0.000006987585
## [12,] 0.000004523406
## [13,] 0.000002520736
##
## $DLTC
##          DLINPC
## [1,] 0.000000000000
## [2,] 0.000137344529
## [3,] 0.000583804216
## [4,] 0.000433186162
## [5,] 0.000171212748
## [6,] 0.000036307348
## [7,] 0.000034282255
## [8,] 0.000038723787
## [9,] 0.000024964010
## [10,] 0.000009783037
## [11,] 0.000003517692
## [12,] 0.000003748501
## [13,] 0.000003845745
##
## $DLCETE28
##          DLINPC
## [1,] 0.000000000000
## [2,] 0.000304305857
## [3,] 0.000373465696

```

```

## [4,] 0.000167853443
## [5,] 0.000127305791
## [6,] 0.000089751434
## [7,] 0.000049983292
## [8,] 0.000025579016
## [9,] 0.000012448007
## [10,] 0.000006485072
## [11,] 0.000004515005
## [12,] 0.000002515525
## [13,] 0.000001179168
##
## $DLIGAE
##          DLINPC
## [1,] 0.000000000000
## [2,] -0.000006014034
## [3,] 0.000035752798
## [4,] 0.000101880979
## [5,] 0.000107697317
## [6,] 0.000067291001
## [7,] 0.000019302644
## [8,] 0.000009260325
## [9,] 0.000010330953
## [10,] 0.000010351320
## [11,] 0.000004299667
## [12,] 0.000002280664
## [13,] 0.000001077557
##
## $DLIPI
##          DLINPC
## [1,] 0.000000000000
## [2,] 0.0002938283073
## [3,] 0.0001164391542
## [4,] 0.0000642643310
## [5,] 0.0000621769713
## [6,] 0.0000441663489
## [7,] 0.0000332198998
## [8,] 0.0000181489496
## [9,] 0.0000082995640
## [10,] 0.0000034179891
## [11,] 0.0000034672138
## [12,] 0.0000024110405
## [13,] 0.0000009791902

```

```
#plot(IR_DLINPC)
```

```
IR_DLTC <- irf(VAR_p, n.ahead = 12, boot = TRUE,
               ci = 0.95, response = "DLTC")

plot(IR_DLTC)
```

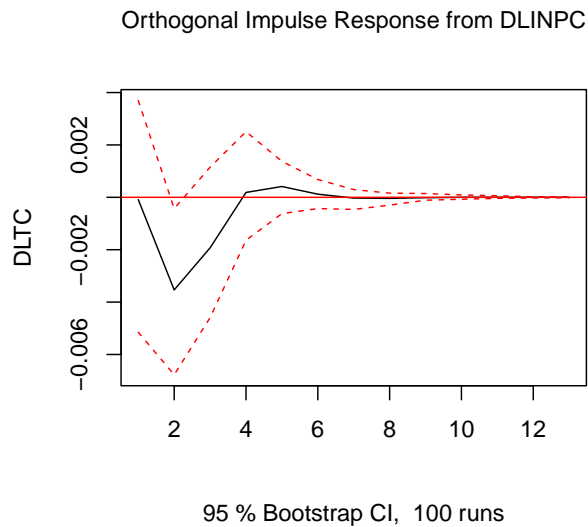


Figure 5.4: Impulso - Respuesta en $DLTC_t$

Los resultados muestran que la respuesta de $DLTC_t$ ante impulsos en los términos de error fue estadísticamente significativo sólo para algunos de los casos y en periodos cortos de tiempo. El resto de los resultados de Impulso-Respuesta se encuentra en el Scrip llamado Clase 15 que se ubica en el repositorio de GitHub.

5.4 Identificación de los Choques Estructurales en los Modelos VAR

En su trabajo seminal “*Macroeconomics and Reality*”, Sims (1980) cuestionó los modelos macroeconómicos utilizados en ese momento. Antes de la aparición de los modelos de referencia modernos, como los Vectores Autorregresivos (VAR) y los modelos estocásticos dinámicos de equilibrio general (DSGE), se estimaban modelos de ecuaciones simultáneas derivados de la escuela keynesiana, especificados con n cantidad de ecuaciones. Por lo tanto, para entender

5.4. IDENTIFICACIÓN DE LOS CHOQUES ESTRUCTURALES EN LOS MODELOS VAR169

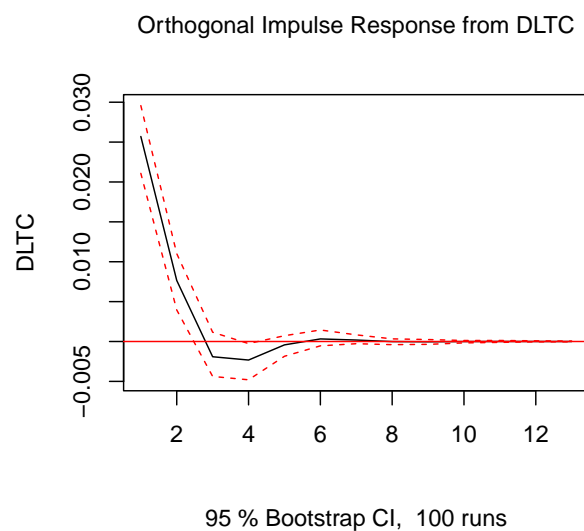


Figure 5.5: Impulso - Respuesta en $DLTC_t$

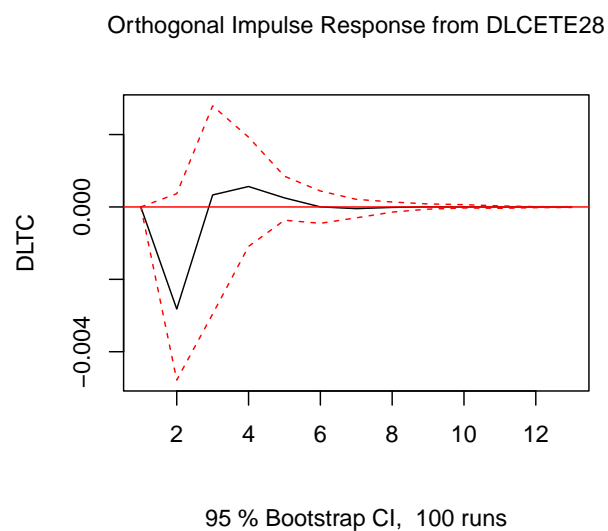


Figure 5.6: Impulso - Respuesta en $DLTC_t$

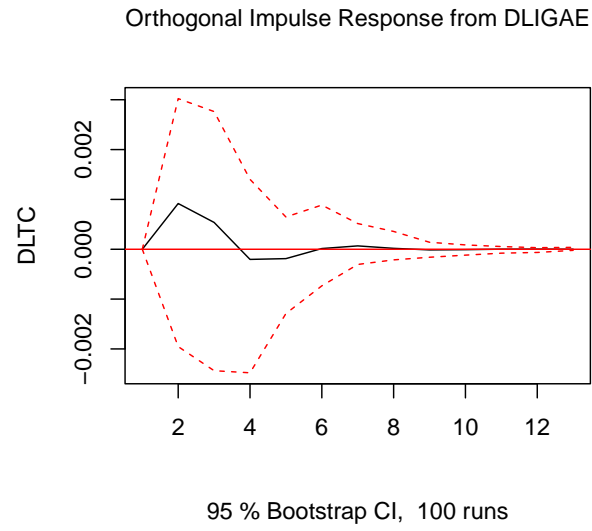


Figure 5.7: Impulso - Respuesta en $DLTC_t$

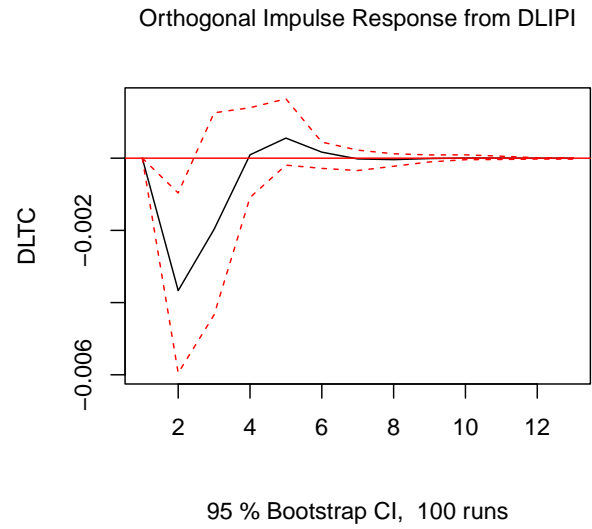


Figure 5.8: Impulso - Respuesta en $DLTC_t$

el surgimiento de los modelos VAR en la macroeconometría, es necesario partir del contexto en el que se escribió el trabajo seminal de Sims.

A finales de los años 70 e inicios de los 80, comenzó una crisis en torno a los modelos estructurales tradicionales. Los modelos macroeconómicos de gran escala –como los de la Cowles Commission– utilizaban restricciones teóricas para identificar relaciones estructurales que, en muchos casos, eran impuestas de manera arbitraria. Estos modelos estaban altamente parametrizados y basados en supuestos teóricos fuertes –como las expectativas adaptativas–, lo que los hacía sensibles a errores específicos. De esta situación surge en parte la *Crítica de Lucas*, en la cual Robert Lucas (1976) mostró que este tipo de modelos fallaba al capturar cambios en el comportamiento cuando las políticas cambiaban, debido a que no trataban adecuadamente las expectativas racionales.

Ante este panorama, Sims propuso los modelos VAR, lo que eventualmente le valió el Premio Nobel de Economía. El objetivo de estos modelos era ofrecer una alternativa más flexible y empíricamente realista, sin imponer supuestos estructurales fuertes a priori. De esta manera, los VAR permiten modelar un conjunto de variables macroeconómicas como funciones de sus propios rezagos, sin imponer una estructura teórica rígida. Además, todos los choques y relaciones se tratan inicialmente de forma simétrica y empírica. Así, los modelos VAR en su forma reducida son modelos completamente ateóricos.

Sin embargo, aunque los modelos VAR permiten capturar la dinámica conjunta entre variables sin imponer restricciones teóricas fuertes a priori, tienen una limitación que no es menor: los choques estimados a partir de la forma reducida –forma que hemos trabajado hasta ahora– están correlacionados entre sí. Es decir, los errores u_t de la forma reducida reflejan combinaciones lineales de múltiples choques estructurales y esto nos impide identificar la interacción contemporánea de las variables. Los choques u_t no se pueden interpretar directamente porque no son ortogonales. En este sentido, para encontrar choques estructurales, debemos encontrar choques w_t que sí sean ortogonales y tengan una interpretación económica.

Comencemos con la forma estructural del VAR para comprender sus parámetros:

$$B_0 Y_t = B_1 Y_{t-1} + \dots B_p Y_{t-p} + w_t \quad (5.20)$$

donde w_t es un término de error con media cero y no correlacionado en el tiempo, también conocido como innovación estructural o shock estructural. Además, se asume que el término de error es incondicionalmente homocedástico a menos de que se indique lo contrario. La matriz B_0 es no singular y establece la interacción contemporánea entre las variables del modelo. De este modo, el modelo puede escribirse de manera compacta como:

$$B(L)y_t = w_t \quad (5.21)$$

donde $B(L) \equiv B_0 - B_1 L - B_2 L^2 - \dots B_p L^p$ es el polinomio autorregresivo en rezagos. La matriz de covarianza del término de error estructural es normalizada

tal que:

$$\mathbb{E}(w_t w_t') \equiv \Sigma_w = I_K. \quad (5.22)$$

Con esto, sabemos que existen tantos shocks estructurales como variables en el modelo. Asimismo, los shocks estructurales, por definición, son mutuamente no correlacionados lo que implica que Σ_w es diagonal. Por último, también sabemos que si normalizamos la varianza de todos los shocks estructurales a uno, no implica una pérdida de generalidad siempre y cuando los elementos diagonales B_0 permanezcan sin restricciones.

Sin embargo, para que el modelo $B_0 Y_t = B_1 Y_{t-1} + \dots + B_p Y_{t-p} + w_t$ sea considerado como un VAR estructural no es suficiente con que los elementos de w_t no se encuentren correlacionados, si no que los choques también deben ser económicamente interpretables. Para esto, derivamos la forma reducida de este modelo VAR estructural de tal modo que y_t es una función de los rezagos de y_t únicamente. Si multiplicamos ambos lados de la representación estructural del VAR por B_0^{-1}

$$B_0^{-1} B_0 y_t = B_0^{-1} B_1 y_{t-1} + \dots + B_0^{-1} B_p y_{t-p} + B_0^{-1} w_t \quad (5.23)$$

de tal modo que puede ser representado como

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \quad (5.24)$$

donde $A_i = B_0^{-1} B_i$, $i = 1, \dots, p$, y $u_t = B_0^{-1} w_t$. Así, las innovaciones en forma reducida de u_t son un promedio ponderado de los shocks estructurales w_t . De manera compacta, el modelo puede ser expresado como:

$$A(L)y_t = u_t \quad (5.25)$$

donde $A(L) = I_K - A_1 L - A_2 L^2 - \dots - A_p L^p$ es el polinomio autorregresivo en rezagos. Los métodos de estimación estándar permiten obtener estimaciones de los parámetros de la forma reducida A_i , para $i = 1, \dots, p$, de las innovaciones reducidas u_t y de su matriz de covarianza $\mathbb{E}(u_t u_t') \equiv \Sigma_u$.

Sin embargo, el debate que se ha generado desde entonces es, una vez estimada la forma reducida, de qué manera se puede recuperar la representación estructural del modelo VAR. Es decir, queremos recuperar la matriz B_0 -o bien, su inversa- que contiene las relaciones estructurales contemporáneas de las variables, recordando que $u_t = B_0^{-1} w_t$.

El método para recuperar dicha matriz ha generado gran debate en el análisis macroeconómico, dando pie incluso a tesis doctorales al respecto. Con esto en mente, ahora veremos algunos de los métodos clásicos para recuperar las relaciones estructurales de las variables. Nos concentraremos en los siguientes:

1. Identificación Recursiva

2. Restricción de largo plazo

3. Imposición de restricciones de signo

Nota: *RStudio* es un poco limitado en cuanto a los métodos de identificación de las innovaciones estructurales relacionados con la teoría económica. En realidad, en los trabajos actuales de macroeconometría suele ocuparse, mayoritariamente, el método de Identificación Recursiva. Sin embargo, para ver los métodos de Restricción de largo plazo e Imposición de restricciones de signo, ver el *Script de MatLab* de estos.

5.4.1 VAR Recursivo - Identificación Recursiva

Una de las formas más populares para recuperar las innovaciones estructurales es mediante la *identificación recursiva*, un caso particular de las restricciones de corto plazo que se implementa utilizando la **descomposición de Cholesky**.

La idea detrás de este método es simple e intuitivo: queremos recuperar las innovaciones estructurales w_t a partir de las innovaciones de forma reducida u_t . Para lograr esto, buscamos *ortogonalizar* los errores de forma reducida, lo que en este contexto significa transformarlos en un conjunto de choques no correlacionados contemporáneamente. Es decir, deseamos que $\mathbb{E}[w_t w_t'] = I$.

Para lograr esto, definimos una matriz P de tamaño $K \times K$, triangular inferior y con diagonal principal positiva, de tal modo que:

$$\Sigma_u = \mathbb{E}[u_t u_t'] = P P' \quad (5.26)$$

La matriz P es conocida como la **descomposición de Cholesky inferior de Σ_u** . Por ejemplo, en un VAR con tres variables, P tendría la forma:

$$P = \begin{bmatrix} p_{11} & 0 & 0 \\ p_{21} & p_{22} & 0 \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

Ahora, si recordamos que en el modelo estructural los errores están relacionados por $u_t = B_0^{-1} w_t$ y que

$$\Sigma_u = B_0^{-1} \mathbb{E}[w_t w_t'] B_0^{-1'} = B_0^{-1} B_0^{-1'} \quad (5.27)$$

entonces una solución válida al problema de identificación estructural es asumir que $B_0^{-1} = P$. Dado que P es triangular inferior, contiene exactamente $\frac{K(K-1)}{2}$ ceros impuestos, lo que cumple la condición de orden para identificar todos los elementos libres de la matriz B_0^{-1} . Por simetría, si B_0^{-1} es triangular inferior, entonces B_0 también lo es -aunque triangular superior-.

Sin embargo, es importante destacar que esta ortogonalización de los errores solo es válida si la estructura recursiva impuesta por la matriz P se puede justificar con fundamentos económicos. El método impone que la primera variable del sistema responde solo a su propio shock contemporáneo, la segunda puede responder al shock de la primera, la tercera a los dos primeros, y así sucesivamente. Por eso se dice que el modelo estructural resultante es recursivo: se impone una estructura jerárquica de causalidad contemporánea, en lugar de inferirla directamente de los datos.

En este sentido, la **descomposición de Cholesky** no descubre las relaciones estructurales sino que las impone. Por lo tanto, el orden de las variables en el VAR **si importa**, ya que determina la interpretación de los shocks estructurales. Por esta razón, se recomienda justificar el orden mediante teoría económica. Por ejemplo, en aplicaciones del Banco de México, el orden se suele establecer de acuerdo con la hipótesis de una economía pequeña y abierta, ordenando las variables de la más exógena a la más endógena.

Veamos un ejemplo rápido de la *identificación recursiva* mediante la **descomposición de Cholesky**:

```
# install.packages("svars")
# install.packages("ggplot2")
# install.packages("ggfortify")
library(svars)
library(ggplot2)
library(ggfortify)

var.reducido <- vars::VAR(USA, lag.max = 10, ic = "AIC")
var.estructural <- id.chol(var.reducido)
summary(var.estructural)
```

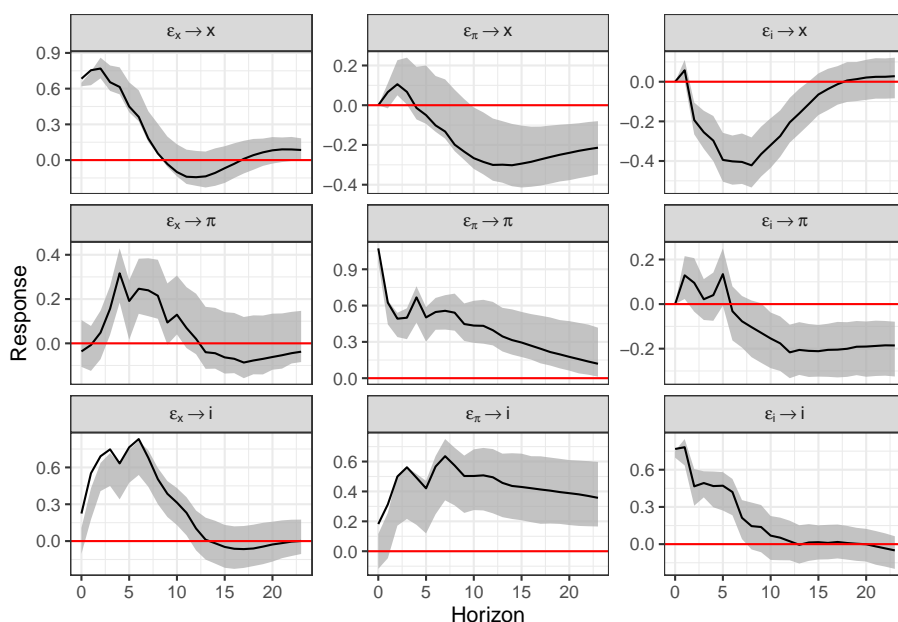
```
##
## Identification Results
## -----
##
## Method: Cholesky
## Sample size: 169
##
## Estimated B Matrix (unique decomposition of the covariance matrix):
##           [,1]      [,2]      [,3]
## x    0.68338360 0.0000000 0.0000000
## pi -0.03642697 1.0726563 0.0000000
## i    0.22450359 0.1818031 0.7672318
```

5.4. IDENTIFICACIÓN DE LOS CHOQUES ESTRUCTURALES EN LOS MODELOS VAR175

```
# x = Porcentaje de desviación logarítmica del PIB real con respecto a la
#      estimación del producto potencial
# pi = Crecimiento anualizado trimestre a trimestre del deflactor del PIB
# i = Tasa de interés de los fondos federales

cores <- parallel::detectCores() - 1
boot.svar <- wild.boot(var.estructural, n.ahead = 24, nboot = 1000, nc = cores)

plot(boot.svar)
```



5.4.2 VAR Estructural - Restricción de largo plazo

Otra manera de identificar las innovaciones estructurales de un modelo VAR, es mediante el método propuesto por Blanchard y Quah (1989) en *“The Dynamic Effects of Aggregate Demand and Supply Disturbances”*.

La intuición de este método es simple: Blanchard y Quah querían medir el efecto de choques de oferta agregada w_t^{AS} y demanda agregada w_t^{AD} sobre el Producto y el Desempleo, imponiendo restricciones en la respuesta acumulada, es decir, en el efecto acumulado permanente (horizonte $h \rightarrow \infty$). La lógica económica detrás de este método es, básicamente, que el choque de demanda agregada no tiene efectos de largo plazo sobre el nivel del PIB real.

Ahora bien, sean:

- ur_t : tasa de desempleo en EE.UU.
- gdp_t : logaritmo del PIB real de EE.UU.

Y definimos el vector z_t

$$z_t = \begin{bmatrix} \Delta gdp_t \\ ur_t \end{bmatrix} \sim I(0) \quad (5.28)$$

Aunque $gdp_t \sim I(1)$, su primera diferencia es estacionaria ($I(0)$), lo cual justifica que $z_t \sim I(0)$. Por otro lado, podemos suponer que este vector es generado por un VAR reducido:

$$A(L)z_t = u_t \quad (5.29)$$

donde $A(L) = I_2 - A_1L - \dots - A_pL^p$, y $u_t \sim (0, \Sigma_u)$ es ruido blanco. Si lo representamos como su forma estructural, tenemos que:

$$B(L)z_t = w_t \quad (5.30)$$

donde $B(L) = B_0 - B_1L - \dots - B_pL^p = B_0A(L)$, y si recordamos a qué es igual w_t , tenemos que:

$$w_t = B_0u_t \rightarrow u_t = B_0^{-1}w_t \rightarrow \Sigma_u = B_0^{-1}(B_0^{-1})' \quad (5.31)$$

Ahora, pasemos a la representación MA estructural para entender dónde se impone la restricción. Desde el modelo MA, tenemos que:

$$z_t = B(L)^{-1}w_t = \Theta(L)w_t \quad (5.32)$$

Esto implica que, dado que $z_t \sim I(0)$, el efecto de un choque estructural sobre z_t se desvanece con el tiempo y, por tanto, tanto Δgdp_t como ur_t eventualmente vuelven a su nivel original tras un choque. Sin embargo, el nivel del PIB real -y no su crecimiento- no necesariamente vuelve a su nivel inicial, por lo que la suma acumulada de las respuestas al choque nos da su efecto permanente. Ahora, pasemos a la matriz de efectos acumulados de largo plazo, definida como:

$$\Theta(1) = \sum_{i=0}^{\infty} \Theta_i = B(1)^{-1} \quad (5.33)$$

Dado que queremos que el PIB real vuelva a su tendencia tras un choque de demanda -recordemos que no tiene efectos de largo plazo- implica imponer un cero en la posición θ_{12} de la matriz $\Theta(1)$:

$$\Theta(1) = \begin{bmatrix} \theta_{11}(1) & 0 \\ \theta_{21}(1) & \theta_{22}(1) \end{bmatrix} \quad (5.34)$$

el hecho de que $\theta_{12}(1) = 0$ implica que el choque de demanda no afecta permanentemente al PIB real y, por otro lado, que $\theta_{11}(1)$ permanezca libre, implica que los choques de oferta sí pueden tener efectos de largo plazo.

Para la identificación, recordemos que:

$$\Theta(1) = B(1)^{-1} = A(1)^{-1}B_0^{-1} \rightarrow B_0^{-1} = A(1)\Theta(1) \quad (5.35)$$

De este modo, una vez que conocemos $A(1)$ y $\Theta(1)$ podemos recuperar B_0^{-1} , recordando que la restricción en $\Theta(1)$ es equivalente a imponer una restricción sobre B_0 .

Para estimar $\Theta(1)$, sabemos que:

$$\Sigma_u = B_0^{-1}(B_0^{-1})' = \Theta(1)\Theta(1)' \quad (5.36)$$

y, entonces, si usamos $A(1)$:

$$A(1)^{-1}\Sigma_u A(1)^{-1'} = \Theta(1)\Theta(1)' \quad (5.37)$$

lo cual implica que podemos calcular el lado izquierdo usando solo parámetros estimados del VAR reducido y luego identificar $\Theta(1)$ imponiendo que tenga forma triangular inferior, y aplicar la **descomposición de Cholesky** para tener finalmente que:

$$B_0^{-1} = A(1)\Theta(1) \quad (5.38)$$

5.4.3 VAR Estructural - Imposición de restricciones de signo

Una alternativa a la identificación de las innovaciones estructurales además de la *Identificación recursiva* es la **Imposición de restricciones de signo**. Sigue un poco la misma lógica: queremos imponer restricciones de signo en vez de proponer una matriz triangular inferior en la matriz B_0^{-1} . Este método ofrece una alternativa menos restrictiva que imponer ceros exactos. Para explicar este método, tomemos el siguiente modelo ejemplo.

Consideremos un modelo biviado de un mercado de bienes con:

- Un choque de demanda $w_t^{demanda}$

- Un choque de oferta w_t^{oferta}

en donde nuestras variables observables son:

- Precio p_t
- Cantidad q_t

Y los errores en forma reducida están dados por:

$$u_t = \begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = B_0^{-1} w_t \quad \text{donde} \quad w_t = \begin{bmatrix} w_t^{oferta} \\ w_t^{demanda} \end{bmatrix} \quad (5.39)$$

Es sencillo: la interpretación económica es que los efectos de los choques u_t dependen de la pendiente de las curvas de demanda y oferta.

En un método tradicional -restricción de exclusión-, se asume que la oferta es vertical en el corto plazo, lo cual implica que los choques de demanda no afectan la cantidad contemporáneamente:

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \begin{bmatrix} * & 0 \\ * & * \end{bmatrix} \begin{bmatrix} w_t^{oferta} \\ w_t^{demanda} \end{bmatrix} \quad (5.40)$$

En donde los asteriscos representan *coeficientes libres* y el cero es una *restricción de exclusión*.

En cambio, si tomamos la teoría económica básica en el enfoque alternativo de **Imposición de restricción de signo** esperaríamos que, ante un choque de oferta positivo -la curva de oferta se desplaza a la derecha-, la cantidad aumente y el precio caiga, mientras que, ante un choque de demanda positiva -la curva de demanda se desplaza a la derecha-, la cantidad aumenta y el precio aumenta. De este modo, tenemos que:

$$\begin{bmatrix} u_t^q \\ u_t^p \end{bmatrix} = \underbrace{\begin{bmatrix} + & + \\ - & + \end{bmatrix}}_{B_0^{-1}} \begin{bmatrix} w_t^{oferta} \\ w_t^{demanda} \end{bmatrix} \quad (5.41)$$

Donde $+$ indica signo positivo estricto y $-$ negativo estricto.

La diferencia clave con el método de *Identificación recursiva mediante la descomposición de Cholesky* es que, en estos últimos modelos, los parámetros están puntualmente identificados mientras que con **restricciones de signo** los parámetros no se identifican exactamente, sino que quedan dentro de un conjunto compatible con las restricciones. De este modo, no hay una única solución, pues hay muchas matrices B_0^{-1} que cumplen con los signos impuestos. También, se pueden imponer **restricciones de signo** de la siguiente forma:

$$\begin{bmatrix} + & + \\ 0 & + \end{bmatrix} \text{ o } \begin{bmatrix} + & + \\ - & 0 \end{bmatrix} \quad (5.42)$$

Estas formas también pueden ser admisibles pero algunas podrían llegar a ser problemáticas si llevan a choques no identificables entre sí.

Se dice que los modelos identificados por signos son más generales que los recursivos pero no son modelos “anidados”, o sea, uno no puede validar o rechazar el otro con datos. Así, las **restricciones de signo** permiten más flexibilidad, pero imponen otras limitaciones como perder identificación puntual.

Ahora, de manera más puntual, para imponer las **restricciones de signo estáticas** consideremos un modelo VAR en su forma estructural:

$$B_0 y_t = B_1 y_{t-1} + \dots + B_p y_{t-p} + w_t \quad (5.43)$$

y normalizamos la matriz de varianzas-covarianzas del término de error estructural w_t tal que:

$$\mathbb{E}(w_t w_t') \equiv \Sigma_w = I_K. \quad (5.44)$$

Ahora, sea $u_t = P\eta_t$, donde u_t es la innovación del VAR en forma reducida y P es la descomposición de Cholesky de Σ_u . Por construcción, los choques η_t son mutuamente no correlacionados y tienen varianza unitaria. Por supuesto, no hay razón para que estos choques correspondan a choques estructurales económicamente interpretables, como los choques de oferta y demanda en el modelo bivariado mencionado anteriormente. Sin embargo, podemos buscar soluciones candidatas w_t^* para los choques estructurales desconocidos w_t construyendo un gran número de combinaciones de los choques η_t de la forma:

$$w_t^* = Q' \eta_t, \quad (5.45)$$

donde Q' es una matriz ortogonal cuadrada tal que $Q'Q = QQ' = I_K$ y $u_t = PQ\eta_t = PQw_t^*$. Por lo tanto, cada solución candidata w_t^* consiste en choques no correlacionados con varianza unitaria. Que una de estas soluciones candidatas w_t^* sea una solución admisible para el choque estructural desconocido w_t , dado el vector de parámetros en forma reducida, depende de si la matriz de impacto estructural implicado PQ satisface las restricciones de signo mantenidas sobre B_0^{-1} . De este modo, solo conservamos las soluciones que satisfagan las restricciones de signo y se descartan las demás. El hecho de repetir este procedimiento nos permite caracterizar el conjunto de todos los modelos estructurales que son consistentes con las restricciones de signo mantenidas y los parámetros en forma reducida. Conocer PQ permite construir todos los coeficientes estructurales de respuesta al impulso que se derivan de las estimaciones de los parámetros en forma reducida.

Los dos enfoques comunes para construir las matrices ortogonales Q están basados en:

1. Matrices de rotación de Givens.
2. Transformación de Householder.

La paquetería del código de MatLab, VAR Toolbox, ocupa el enfoque de **Matrices de rotación de Givens**.

5.5 Cointegración

Hasta ahora en el curso hemos usado el supuesto de que las series son estacionarias para el conjunto de técnicas $ARMA(p, q)$ y $VAR(p)$. No obstante, dado que relajamos el supuesto de estacionariedad (incluyendo la estacionariedad en varianza) y que establecimos una serie de pruebas para determinar cuándo una serie es estadísticamente estacionaria, ahora podemos plantear una técnica llamada Cointegración. Para esta técnica consideraremos sólo series que son $I(1)$ y reconoceremos que se originó con los trabajos de Engle y Granger (1987), Stock (1987) y Johansen (1988).

5.5.1 Definición y propiedades del proceso de cointegración

Cointegración puede ser caracterizada o definida en palabras sencillas como que dos o más variables tienen una relación común estable en el largo plazo. Es decir, estas no suelen tomar caminos o trayectorias diferentes, excepto por períodos de tiempo transitorios y eventuales. A continuación, utilizaremos la definición de Engle y Granger (1984) de cointegración.

Sea \mathbf{Y} un vector de k -series de tiempo, decimos que los elementos en \mathbf{Y} están cointegrados en un orden (d, c) , es decir, $\mathbf{Y} \sim CI(d, c)$, si todos los elementos de \mathbf{Y} son series integradas de orden d , $I(d)$, y si existe al menos una combinación lineal no trivial \mathbf{Z} de esas variables que es de orden $I(d - c)$, donde $d \geq c > 0$, si y sólo si:

$$\beta'_i \mathbf{Y}_t = \mathbf{Y}_{it} \sim I(d - c) \quad (5.46)$$

Donde $i = 1, 2, \dots, r$ y $r < k$.

A los diferentes vectores β_i se les denomina como vectores de cointegración. El rango de la matriz de vectores de cointegración r es el número de vectores de cointegración linealmente independientes. En general diremos que los vectores de la matriz de cointegración β tendrán la forma de:

$$\beta' \mathbf{Y}_t = \mathbf{Z}_t \quad (5.47)$$

Antes de continuar hagamos algunas observaciones. Si todas las variables de \mathbf{Y} son $I(1)$ y $0 \leq r < k$, diremos que las series no cointegran si $r = 0$. Si esto pasa, entonces, como demostraremos más adelante, la mejor opción será estimar un modelo VAR(p) en diferencias. Adicionalmente, asumiremos que $c = d = 1$, por lo que la relación de cointegración, en su caso, generará combinaciones lineales \mathbf{Z} estacionarias.

5.5.2 Cointegración para modelos de más de una ecuación o para modelos basados en Vectores Autoregresivos

Sean Y_1, Y_2, \dots, Y_k son series que forman \mathbf{Y} y que todas son $I(1)$, entonces los siguientes casos son posibles:

1. Si $r = 1$ entonces se trata de un caso de cointegración de Granger.
2. Si $r \geq 1$ entonces se trata de un caso de cointegración múltiple de Johansen.

Por lo anterior, en este curso analizaremos el caso de Cointegración de Johansen. Ahora plantearemos la forma de estimar el proceso de cointegración. El primer paso para ello es determinar un modelo VAR(p) con las k -series no estacionarias (series en niveles)—en este punto se vuelve fundamental caracterizar las series a través de pruebas de raíces unitarias—. Elegimos el valor de p mediante el uso de los criterios de información. De esta forma tendremos una especificación similar a:

$$\mathbf{Y}_t = \sum_{j=1}^p \mathbf{A}_j \mathbf{Y}_{t-j} + \mathbf{D}_t + \mathbf{U}_t \quad (5.48)$$

Donde \mathbf{U}_t es un término de error k -dimensional puramente aleatorio; \mathbf{D}_t contiene los componentes determinísticos de constante y tendencia, y \mathbf{A}_i , $i = 1, 2, \dots, p$, son matrices de $k \times k$ coeficientes. Notemos que el VAR(p) involucrado en este caso, a diferencia del VAR anteriormente estudiado, puede incluir un término de tendencia. Esto en razón de que hemos relajado el concepto de estacionariedad.

Si reescribimos la ecuación (5.48) en su forma de Vector Corrector de Errores

(VEC, por sus siglas en inglés) tenemos:

$$\begin{aligned}
\mathbf{Y}_t - \mathbf{Y}_{t-1} &= \Delta \mathbf{Y}_t \\
&= \sum_{j=1}^p \mathbf{A}_j \mathbf{Y}_{t-j} + \mathbf{D}_t - \mathbf{Y}_{t-1} + \mathbf{U}_t \\
&= (\mathbf{A}_1 - \mathbf{I}) \mathbf{Y}_{t-1} + \mathbf{A}_2 \mathbf{Y}_{t-2} + \dots + \mathbf{A}_p \mathbf{Y}_{t-p} + \mathbf{D}_t + \mathbf{U}_t \\
&= \left(\sum_{j=1}^p \mathbf{A}_j - \mathbf{I} \right) \mathbf{Y}_{t-1} + \sum_{j=1}^{p-1} \mathbf{A}_j^* \Delta \mathbf{Y}_{t-j} + \mathbf{D}_t + \mathbf{U}_t \\
&= - \left(\mathbf{I} - \sum_{j=1}^p \mathbf{A}_j \right) \mathbf{Y}_{t-1} + \sum_{j=1}^{p-1} \mathbf{A}_j^* \Delta \mathbf{Y}_{t-j} + \mathbf{D}_t + \mathbf{U}_t \\
\Delta \mathbf{Y}_t &= -\Pi \mathbf{Y}_{t-1} + \sum_{j=1}^{p-1} \mathbf{A}_j^* \Delta \mathbf{Y}_{t-j} + \mathbf{D}_t + \mathbf{U}_t \tag{5.49}
\end{aligned}$$

Donde $\mathbf{A}_j^* = -\sum_{i=j+1}^p \mathbf{A}_i$, $i = 1, 2, \dots, p-1$, y la matriz Π representa todas las relaciones de largo plazo entre las variables, por lo que la matriz es de rango completo $k \times k$. Por lo tanto, tenemos que dicha matriz en la ecuación (5.49) se puede factorizar como:

$$\Pi_{(k \times k)} = \Gamma_{(k \times r)} \beta'_{(r \times k)} (\#eq : Pi_Matrix) \tag{5.50}$$

Donde $\beta'_{(r \times k)} \mathbf{Y}_{t-1}$ son r combinaciones linealmente independientes que son estacionarias.

Dada la ecuación (5.49) podemos establecer la aproximación de Johansen (1988) que se realiza mediante una estimación por Máxima Verosimilitud de la ecuación:

$$\Delta \mathbf{Y}_t + \Gamma \beta' \mathbf{Y}_{t-1} = \sum_{j=1}^{p-1} \mathbf{A}_j^* \Delta \mathbf{Y}_{t-j} + \mathbf{D}_t + \mathbf{U}_t \tag{5.51}$$

Donde una vez estimado el sistema:

$$\beta = [v_1, v_2, \dots, v_r] \tag{5.52}$$

Cada v_i , $i = 1, 2, \dots, r$ es un vector propio que está asociado con los r valores propios positivos, mismos que están asociados con la prueba de hipótesis de cointegración. Dicha hipótesis está basada en dos estadísticas con las que se determina el rango r de Π :

1. Prueba de Traza: H_0 : Existen al menos r valores propios positivos o Existen al menos r relaciones de largo plazo estacionarias.

2. Prueba del valor propio máximo o λ_{max} : H_0 : Existen r valores propios positivos o Existen r relaciones de largo plazo estacionarias.

Ejemplo. Para ejemplificar el procedimiento de cointegración utilizaremos las series de INPC, Tipo de Cambio, rendimiento de los Cetes a 28 días, IGAE e Índice de Producción Industrial de Estados Unidos. Quizá el marco teórico de la relación entre las variables no sea del todo correcto, pero dejando de lado ese problema, estimaremos si las 5 series cointegran.

Por principio, probaremos que todas las series son $I(1)$, lo cual es cierto (ver Scrip para mayores detalles). En las Figuras 5.9, 5.10 y 5.11 se muestran las series en niveles y en diferencias, con lo cual ilustramos como es viable que las series sean $I(1)$.

```
library(ggplot2)
library(dplyr)
library(stats)
library(MASS)
library(strucchange)
library(zoo)
library(sandwich)
library(urca)
library(lmtest)
library(vars)

#
load("BD/Datos_Ad.RData")

#
## Conversion a series de tiempo:
Datos <- ts(Datos_Ad[7: 11],
            start = c(2000, 1),
            freq = 12)

LDatos <- log(Datos)

DLDatos <- diff(log(Datos, base = exp(1)),
               lag = 1,
               differences = 1)

plot(LDatos,
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     #main = "Series en Logaritmos",
     xlab = "Tiempo")
```

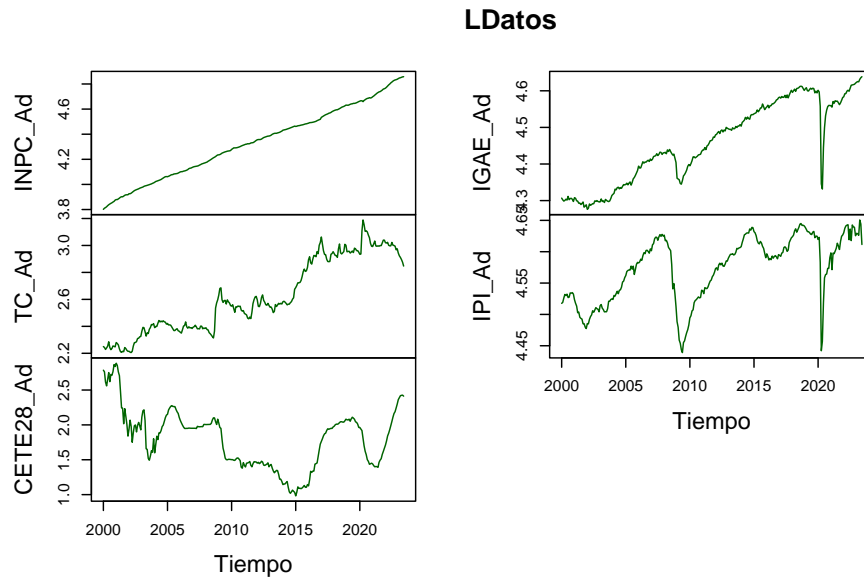


Figure 5.9: Series en niveles (logaritmos) para la prueba de Cointegración

```
plot(DLDatos,
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     #main = "Series en Diferencias Logaritmicas",
     xlab = "Tiempo")
```

```
plot(cbind(LDatos, DLDatos),
     plot.type = "m", nc = 2,
     col = c("darkgreen", "darkblue", "darkred", "orange", "purple"),
     #main = "Comparacion de Series en Diferencias",
     xlab = "Tiempo")
```

Posteriormente, determinamos cuál es el orden adecuado de un VAR(p) en niveles. En el Cuadro 5.8 mostramos los resultados de los criterios de información para determinar el número de rezagos óptimos, el cual resultó en $p = 3$ para los criterios AIC y FPE, $p = 2$ para el criterio HQ y $p = 1$ para el criterio SC. Por lo tanto, decidiremos utilizar un VAR(3) con tendencia y constante. Note que es posible elegir otros modelos de VAR que incluyan: solo tendencia, solo constante o ninguno de estos elementos.

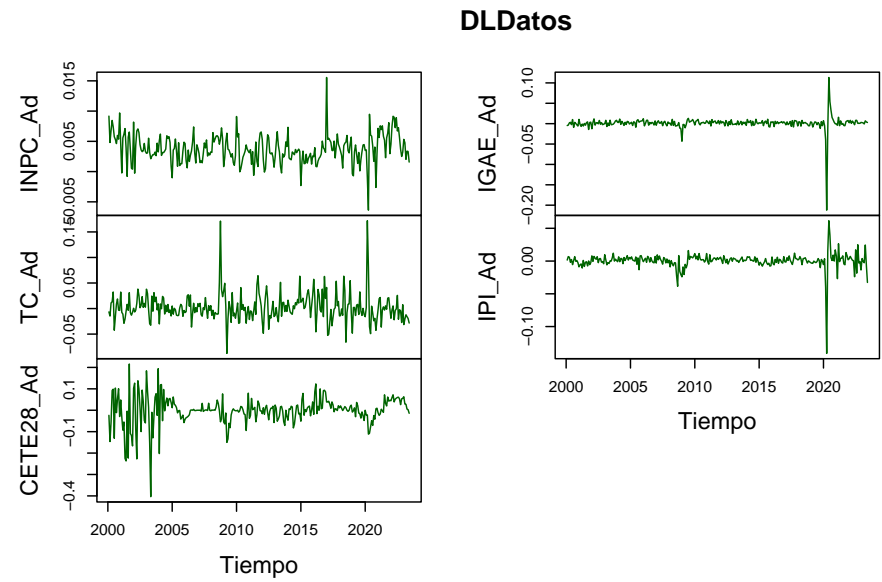


Figure 5.10: Series en Diferencias Logarítmicas para la prueba de Cointegración

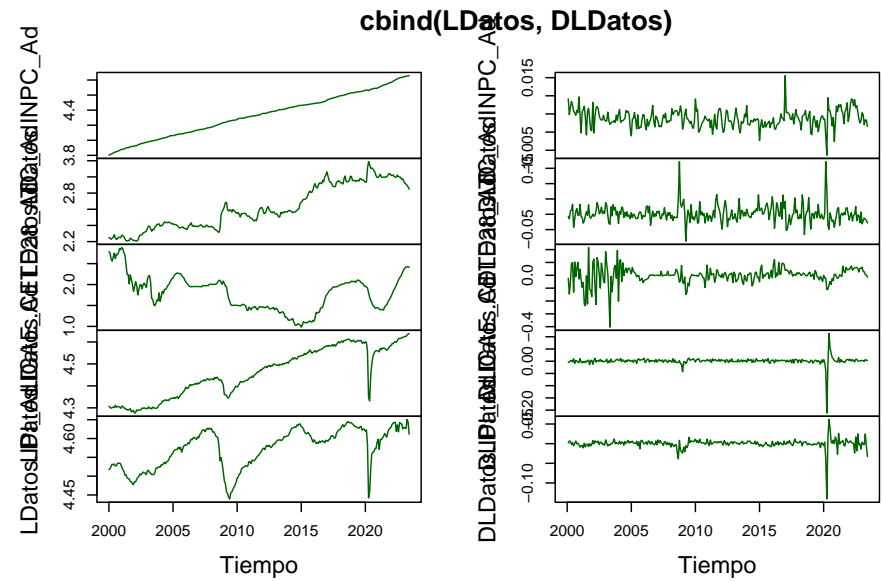


Figure 5.11: Comparacion de Series en Diferencias para la prueba de Cointegración

Table 5.8: Criterios de información para diferentes especificaciones de modelos VAR(p) con término constante y tendencia de las series $LINPC_t$, LTC_t , $LCETE28_t$, $LIGAE_t$ y $LIPi_t$.

Rezagos	AIC	HQ	SC	FPE
1	-4.606707e+01	-4.585260e+01	-4.553568e+01	9.848467e-21
2	-4.643287e+01	-4.606521e+01	-4.552191e+01	6.834064e-21
3	-4.647783e+01	-4.595697e+01	-4.518730e+01	6.539757e-21
4	-4.645834e+01	-4.578428e+01	-4.478824e+01	6.679778e-21

```
## VAR(p) Seleccion:
```

```
VARselect(LDatos, lag.max = 10, type = "both")
```

```
## $selection
## AIC(n) HQ(n) SC(n) FPE(n)
##      3      2      2      3
##
## $criteria
##                                1                                2
## AIC(n) -43.1308318646810135987834656 -43.5668661854371990216350241
## HQ(n)  -42.9445592539269327403417265 -43.2475417098587726627556549
## SC(n)  -42.6668499811502783813921269 -42.7714686708130784609238617
## FPE(n)  0.0000000000000000001855848  0.0000000000000000001200259
##                                3                                4
## AIC(n) -43.5797269785975984746073664 -43.5035550099222803055454278
## HQ(n)  -43.1273506381948337207177246 -42.9181268046951771566455136
## SC(n)  -42.4529138328800996760037378 -42.0453262331114032690493332
## FPE(n)  0.0000000000000000001185548  0.0000000000000000001280605
##                                5                                6
## AIC(n) -43.4665644031667497415583057 -43.3662074624718840709647338
## HQ(n)  -42.7480843331153010922207613 -42.5146755275960970266169170
## SC(n)  -41.6769199952624873617423873 -41.2451474234742434532563493
## FPE(n)  0.0000000000000000001330869  0.0000000000000000001474594
##                                7                                8
## AIC(n) -43.3576732552204475723556243 -43.4475783817780367712657608
## HQ(n)  -42.3730894555203079221428197 -42.3299427172535587260426837
## SC(n)  -40.9051975851294216113274160 -40.6636870805936325723450864
## FPE(n)  0.0000000000000000001491718  0.0000000000000000001368875
##                                9                               10
## AIC(n) -43.3742641642367701138027769 -43.3660356138277762738653109
## HQ(n)  -42.1235766348879465681420697 -41.9822962196546143331943313
## SC(n)  -40.2589572319589734661349212 -39.9193130504566013883049891
```

```
## FPE(n)    0.0000000000000000001480478    0.0000000000000000001502148
```

```
VARselect(LDatos, lag.max = 10, type = "trend")
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      3      2      2      3
##
## $criteria
##                                1                                2
## AIC(n) -43.0218334558386885646541486 -43.4733608627676417768270767
## HQ(n)  -42.8621712180494753852144640 -43.1806467601540902023771196
## SC(n)  -42.6241346985266318370122463 -42.7442464743621997058653506
## FPE(n)  0.0000000000000000002069525  0.0000000000000000001317817
##                                3                                4
## AIC(n) -43.4926570116813877575623337 -43.4403901004477006608794909
## HQ(n)  -43.0668910442434906826747465 -42.8815722681854651909816312
## SC(n)  -42.4321269921825603432807839 -42.0484444498554950087054749
## FPE(n)  0.0000000000000000001293226  0.0000000000000000001363788
##                                5                                6
## AIC(n) -43.4055918510038054591859691 -43.3210122680509144288407697
## HQ(n)  -42.7137221539172173834231216 -42.4960907061399879580676497
## SC(n)  -41.6822305693182144636921294 -41.2662353552719451954544638
## FPE(n)  0.0000000000000000001414042  0.0000000000000000001541996
##                                7                                8
## AIC(n) -43.3018930975745490741246613 -43.3631837990786621617189667
## HQ(n)  -42.3439196708392771029139112 -42.2721585075190517954979441
## SC(n)  -40.9157005537021944974185317 -40.6455756241129293471203709
## FPE(n)  0.0000000000000000001576227  0.0000000000000000001488111
##                                9                                10
## AIC(n) -43.3123078988938914335449226 -43.289311818269602838427090
## HQ(n)  -42.0882307425099355668862700 -41.932182797061308576758165
## SC(n)  -40.2632840928347732756265032 -39.908872381117106442616205
## FPE(n)  0.0000000000000000001573361  0.000000000000000000161972
```

```
VARselect(LDatos, lag.max = 10, type = "const")
```

```
## $selection
## AIC(n)  HQ(n)  SC(n) FPE(n)
##      3      2      2      3
##
## $criteria
##                                1                                2
## AIC(n) -43.0529318587713945021278050 -43.495191294123372927060700
```

```

## HQ(n) -42.8932696209821813226881204 -43.202477191509821352610743
## SC(n) -42.6552331014593377744859026 -42.766076905717930856098974
## FPE(n) 0.000000000000000002006156 0.00000000000000000128936
##                                     3 4
## AIC(n) -43.5086218228659333817631705 -43.4510965812944505159975961
## HQ(n) -43.0828558554280363068755833 -42.8922787490322150460997364
## SC(n) -42.4480918033671059674816206 -42.0591509307022448638235801
## FPE(n) 0.000000000000000001272744 0.000000000000000001349265
##                                     5 6
## AIC(n) -43.4144752420798099024068506 -43.325553474358613925687678
## HQ(n) -42.7226055449932289320713608 -42.500631912447687454914558
## SC(n) -41.6911139603942189069130109 -41.270776561579644692301372
## FPE(n) 0.000000000000000001401536 0.00000000000000000153501
##                                     7 8
## AIC(n) -43.3069657328556019137977273 -43.3724989978216228792007314
## HQ(n) -42.3489923061203299425869773 -42.2814737062620125129797088
## SC(n) -40.9207731889832473370915977 -40.6548908228558900646021357
## FPE(n) 0.000000000000000001568252 0.000000000000000001474313
##                                     9 10
## AIC(n) -43.316467649333738165751129 -43.2911053192979622394886974
## HQ(n) -42.092390492949782299092476 -41.9339762980896679778197722
## SC(n) -40.267443843274620007832709 -39.9106658821454658436778118
## FPE(n) 0.00000000000000000156683 0.000000000000000001616817

```

```
VARselect(LDatos, lag.max = 10, type = "none")
```

```

## $selection
## AIC(n) HQ(n) SC(n) FPE(n)
##      3      2      2      3
##
## $criteria
##                                     1 2
## AIC(n) -43.009650989538556586921914 -43.4653473630038362784944184
## HQ(n) -42.876599124714218191911641 -43.1992436333551523830465158
## SC(n) -42.678235358445178349029447 -42.8025161008170726972821285
## FPE(n) 0.00000000000000000209486 0.000000000000000001328347
##                                     3 4
## AIC(n) -43.4961615576300175689539174 -43.4416013221261536614292709
## HQ(n) -43.0970059631569952784957422 -42.9093938628287787651061080
## SC(n) -42.5019146643498757498491614 -42.1159387977526264990046911
## FPE(n) 0.000000000000000001288547 0.000000000000000001361851
##                                     5 6
## AIC(n) -43.3978234579566617412638152 -43.31937496312242075191534
## HQ(n) -42.7325641338349484499303799 -42.52106377417636196014428
## SC(n) -41.7407453024897563409467693 -41.33088117656213000827847

```

```
## FPE(n)    0.0000000000000000001424605    0.00000000000000000015438
##                                     7                                     8
## AIC(n) -43.2856068136311264993310033 -43.342694393587237300380366
## HQ(n)  -42.3542437598607293125496653 -42.278279474992494613161398
## SC(n)  -40.9656973959774575178016676 -40.691369344840175870103849
## FPE(n)    0.0000000000000000001601088    0.000000000000000000151765
##                                     9                                     10
## AIC(n) -43.2981268022445746623816376 -43.2689202279021429831118439
## HQ(n)  -42.1006600188254935801523970 -41.9384015796587164004449733
## SC(n)  -40.3153861224041349942126544 -39.9547639169683179716230370
## FPE(n)    0.0000000000000000001594144    0.0000000000000000001650919
```

El mismo número de rezagos los utilizaremos para probar la Cointegración, ya sea por una estadística de la Traza o por una del máximo valor propio. Dado que los resultados se sostienen, sólo mostraremos uno de los casos en que las series cointegran y únicamente para el caso de la prueba de la traza (el otro caso está disponible en el código de R disponible abajo). En el Cuadro 5.5.2 reportamos los resultados del Test de Cointegración para un modelo con 3 rezagos.

Cuadro: Prueba de la traza para cointegración considerando un VAR(p) con término constante y tendencia de las series $LINPC_t$, LTC_t , $LCETE28_t$, $LIGAE_t$ y $LIPt$.

$r \leq$	Estadística	10%	5%	1%
4	4.79	10.49	12.25	16.26
3	13.97	22.76	25.32	30.45
2	27.45	39.06	42.44	48.45
1	48.14	59.14	62.99	70.05
0	118.98	83.20	87.31	96.58

Los resultados del Cuadro 5.5.2 indican que aceptamos la hipótesis nula para el caso de $r \leq 1$ al 5%, por lo que podemos concluir que existe evidencia estadística para probar que existe al menos 1 vector de cointegración. Por lo que dicho vector normalizado a la primera entrada es:

$$\beta = \begin{bmatrix} 1.00000000 \\ 0.2100057 \\ 0.4812626 \\ -2.8386112 \\ -1.2576912 \\ 14.2887887 \end{bmatrix} \quad (5.53)$$

Donde el vector está normalizado para la serie $LINPC_t$, por lo que concluimos

que la relación de largo plazo que encontramos cointegra estará dada por:

$$\begin{aligned} LINPC_t = & -0.2100057LTC_t - 0.4812626LCETE28_t \\ & + 2.8386112LIGAE_t + 1.2576912LIPI_t \\ & - 14.2887887 \end{aligned}$$

```
## VAR Estimacion:

VAR_1 <- VAR(LDatos, p = 3, type = "both")

#summary(VAR_1)

#plot(VAR_1, names = "INPC_Ad")
#plot(VAR_1, names = "TC_Ad")
#plot(VAR_1, names = "CETE28_Ad")
#plot(VAR_1, names = "IGAE_Ad")
#plot(VAR_1, names = "IPI_Ad")

# Cointegration Test:
#ca.jo = function(x, type = c("eigen", "trace"), ecdet = c("none", "const",
#"trend"), K = 2, spec = c("longrun", "transitory"), season = NULL,
#dumvar = NULL)

#summary(ca.jo(LDatos, type = "trace", ecdet = "trend", K = 3, spec = "longrun"))

#summary(ca.jo(LDatos, type = "trace", ecdet = "const", K = 3, spec = "longrun"))

#summary(ca.jo(LDatos, type = "trace", ecdet = "none", K = 3, spec = "longrun"))

CA_1 <- ca.jo(LDatos, type = "trace", ecdet = "const", K = 3, spec = "longrun")

summary(CA_1)

##
## #####
## # Johansen-Procedure #
## #####
##
## Test type: trace statistic , without linear trend and constant in cointegration
##
## Eigenvalues (lambda):
## [1] 0.2242278157790942305638 0.0714829742929301154009 0.0471697381127354625763
## [4] 0.0323558252755368550013 0.0170373866386099859227 0.0000000000000001925707
##
## Values of teststatistic and critical values of test:
```

```

##
##          test 10pct  5pct  1pct
## r <= 4 |    4.79  7.52  9.24 12.97
## r <= 3 |   13.97 17.85 19.96 24.60
## r <= 2 |   27.45 32.00 34.91 41.07
## r <= 1 |   48.14 49.65 53.12 60.16
## r = 0  |  118.98 71.86 76.07 84.45
##
## Eigenvectors, normalised to first column:
## (These are the cointegration relations)
##
##          INPC_Ad.13  TC_Ad.13  CETE28_Ad.13  IGAE_Ad.13  IPI_Ad.13
## INPC_Ad.13    1.0000000  1.0000000    1.0000000    1.000000    1.00000000
## TC_Ad.13      0.2100057  0.3958412    0.6038714   -3.946863   -0.66160415
## CETE28_Ad.13  0.4812626 -0.3505729    0.2932620    1.025381    0.02134859
## IGAE_Ad.13    -2.8386112 -4.9448228   -4.1431506   10.685831   -1.62038123
## IPI_Ad.13     -1.2576912  0.8682351    2.2789480  -19.080015    1.60589077
## constant      14.2887887 13.2882018    1.6072247   43.673551   -2.75645487
##          constant
## INPC_Ad.13      1.00000000
## TC_Ad.13       -0.84381863
## CETE28_Ad.13   0.06866727
## IGAE_Ad.13      1.32792512
## IPI_Ad.13      -3.05544132
## constant        5.81151022
##
## Weights W:
## (This is the loading matrix)
##
##          INPC_Ad.13  TC_Ad.13  CETE28_Ad.13  IGAE_Ad.13
## INPC_Ad.d  0.001551917  0.00004896886  0.0002254138 -0.0001685412
## TC_Ad.d    0.002200808 -0.00575637923 -0.0293306746 -0.0003866727
## CETE28_Ad.d -0.009399815 -0.02250446234 -0.0166361225 -0.0160384399
## IGAE_Ad.d   0.001105776  0.01441007456  0.0023574644 -0.0015712694
## IPI_Ad.d    0.001442047  0.01286401841 -0.0014986278 -0.0011379265
##          IPI_Ad.13  constant
## INPC_Ad.d -0.0002093347  0.0000000000000001290255
## TC_Ad.d    0.0139760105 -0.00000000000000017253317
## CETE28_Ad.d -0.0111636433 -0.000000000000000295711342
## IGAE_Ad.d   0.0077903982  0.00000000000000030602278
## IPI_Ad.d   -0.0046051140  0.00000000000000027991698

```

Considerando lo anterior, podemos determinar \hat{U}_t para esta ecuación de cointegración. En la Figura 5.12 mostramos los residuales estimados. Derivado de la inspección visual, parecería que estos no son estacionarios, condición que debería ser cierta. De esta forma, una prueba deseable es aplicar todas las pruebas

de raíces unitarias a esta serie para mostrar que es $I(0)$. En el Scrip llamado Clase 18 en la carpeta de GoogleDrive se muestran algunas pruebas sobre esta serie y se encuentra que es posible que no sea estacionaria.

```
TT <- ts(c(1:282),
        start = c(2000, 1),
        freq = 12)

U <- LDatos[, 1] + 0.2100057 * LDatos[, 2] + 0.4812626 * LDatos[, 3] - 2.8386112 * LDatos[, 4]

#

plot(U,
     main = "Residuales de la Ecuación de Cointegración",
     type = "l",
     col = "darkred")
```

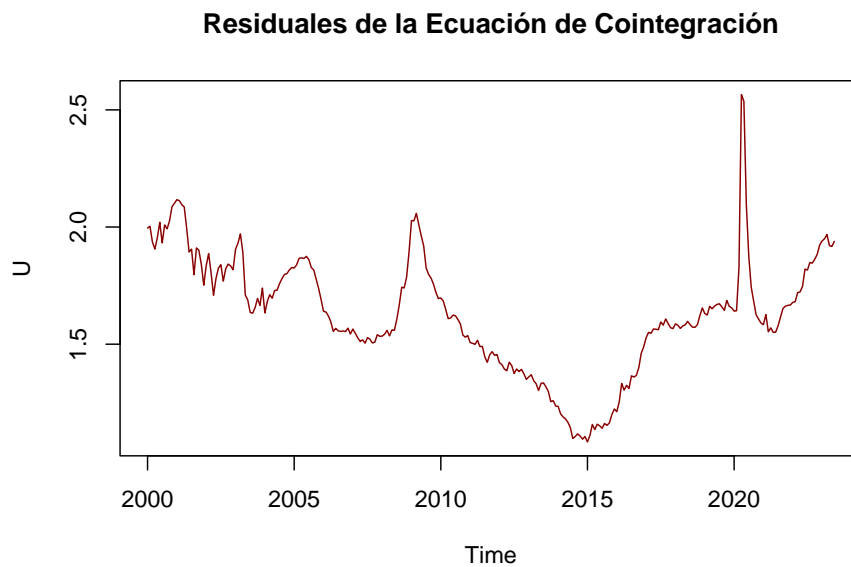


Figure 5.12: Residuales estimados de la ecuación de cointegración

5.6 Modelos ADRL

5.6.1 Teoría

Una vez que hemos analizado diversas técnicas de series de tiempo, el problema consiste en seleccionar el modelo correcto. La Figura 5.13 muestra un esquema o diagrama de cómo podríamos proceder para seleccionar el modelo correcto.

```
knitr::include_graphics("Plots/TimeSeries_Models.png")
```

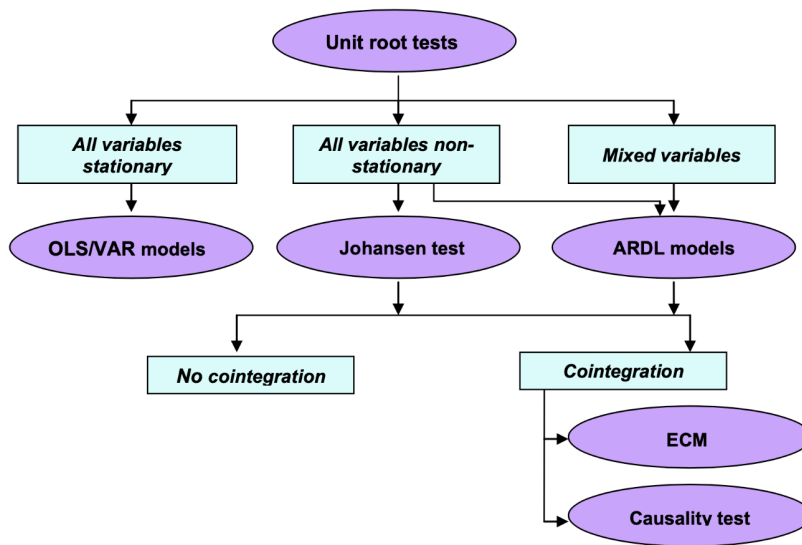


Figure 5.13: Method selection for time series data. OLS: Ordinary least squares; VAR: Vector autoregressive; ARDL: Autoregressive distributed lags; ECM: Error correction models, retomado de: Shrestha y Bhatta (2018)

En este caso incorporaremos a los modelos autogregresive distributed lag models (ARDL, por sus siglas en inglés). En estos casos el procedimiento de Johansen no podría aplicarse directamente cuando las variables incluidas son de un orden mixto o cuando simplemente todas no son estacionarias. Un modelo ARDL está basado en procedimientos de MCO.

Este tipo de modelos toma suficientes rezagos para capturar el mecanismo generador de datos. También es posible llegar a una especificación del mecanismo corrector de errores a partir de una transformación lineal del ARDL.

Consideremos la siguiente ecuación:

$$Y_t = \alpha + \delta X_t + \gamma Z_t + U_t \quad (5.54)$$

Dada la ecuación (5.54) podemos establecer su forma de mecanismo corrector de errores en forma ARDL dada por:

$$\begin{aligned}\Delta Y_t = & \alpha + \sum_{i=1}^p \beta_i \Delta Y_{t-i} + \sum_{i=1}^p \delta_i \Delta X_{t-i} + \sum_{i=1}^p \gamma_i \Delta Z_{t-i} \\ & + \lambda_1 Y_{t-1} + \lambda_2 X_{t-1} + \lambda_3 Z_{t-1} + U_t\end{aligned}$$

Donde los coeficientes β_i , δ_i , γ_i representan la dinámica de corto plazo y las λ 's la dinámica de largo plazo.

La hipótesis nula (H_0) es que las $\lambda_1 + \lambda_2 + \lambda_3 = 0$, es decir, que no existe relación de largo plazo.

En la práctica estimamos una especificación con rezagos distribuidos:

$$Y_t = \alpha + \sum_{i=1}^p \beta_i Y_{t-i} + \sum_{i=1}^p \delta_i X_{t-i} + \sum_{i=1}^p \gamma_i Z_{t-i} + U_t \quad (5.55)$$

Además de verificar si las series involucradas son estacionarias y decidir el número de rezagos p mediante criterios de información.

5.6.2 Ejemplo

5.6.2.1 DESCRIPCIÓN DEL PROBLEMA

Supongamos que queremos modelar el logaritmo de dinero (M2) como una función de LRY (logarithm of real income), IBO (bond rate) e IDE (bank deposit rate).

- El problema es que la aplicación de una regresión de MCO en datos no estacionarios daría lugar a una regresión espúrea.
- Los parámetros estimados serían consistentes solo si las series estuvieran cointegradas.

5.6.2.2 Importamos Datos desde un dataset de R:

Utilizaremos un dataset integrado en la biblioteca ARDL de R. Se trata de un dataframe con 55 renglones y 5 variables en el período de 1974:Q1 a 1987:Q3 de las siguientes variables:

LRM: logarithm of real money, M2

LRY: logarithm of real income

LPY: logarithm of price deflator

IBO: bond rate

IDE: bank deposit rate

```
library(zoo)
library(xts)
library(ARDL)

#
data(denmark)

names(denmark)
```

```
## [1] "LRM" "LRY" "LPY" "IBO" "IDE"
```

5.6.2.3 Procedimiento:

1. Calculamos un auto ADRL para determinar la combinación óptima de rezagos.

```
models <- auto_ardl(LRM ~ LRY + IBO + IDE, data = denmark, max_order = 5)

names(models)
```

```
## [1] "best_model" "best_order" "top_orders"
```

```
#
models$top_orders
```

```
##      LRM LRY IBO IDE      AIC
## 1      3   1   3   2 -251.0259
## 2      3   1   3   3 -250.1144
## 3      2   2   0   0 -249.6266
## 4      3   2   3   2 -249.1087
## 5      3   2   3   3 -248.1858
## 6      2   2   0   1 -247.7786
## 7      2   1   0   0 -247.5643
## 8      2   2   1   1 -246.6885
## 9      3   3   3   3 -246.3061
## 10     2   2   1   2 -246.2709
## 11     2   1   1   1 -245.8736
## 12     2   2   2   2 -245.7722
## 13     1   1   0   0 -245.6620
```

```
## 14  2  1  2  2 -245.1712
## 15  3  1  2  2 -245.0996
## 16  1  0  0  0 -244.4317
## 17  1  1  0  1 -243.7702
## 18  5  5  5  5 -243.3120
## 19  4  1  3  2 -243.0728
## 20  4  1  3  3 -242.4378
```

```
#
models$best_order
```

```
## LRM LRY IBO IDE
##   3   1   3   2
```

```
#
models$best_model
```

```
##
## Time series regression with "zooreg" data:
## Start = 1974 Q4, End = 1987 Q3
##
## Call:
## dynlm::dynlm(formula = full_formula, data = data, start = start,
##               end = end)
##
## Coefficients:
## (Intercept)    L(LRM, 1)    L(LRM, 2)    L(LRM, 3)          LRY    L(LRY, 1)
##         2.6202         0.3192         0.5326        -0.2687         0.6728        -0.2574
##          IBO    L(IBO, 1)    L(IBO, 2)    L(IBO, 3)          IDE    L(IDE, 1)
##        -1.0785        -0.1062         0.2877        -0.9947         0.1255        -0.3280
##    L(IDE, 2)
##         1.4079
```

```
#
BestMod <- models$best_model
```

```
summary(BestMod)
```

```
##
## Time series regression with "zooreg" data:
## Start = 1974 Q4, End = 1987 Q3
##
## Call:
## dynlm::dynlm(formula = full_formula, data = data, start = start,
```

```
##      end = end)
##
## Residuals:
##      Min          1Q      Median          3Q      Max
## -0.029939 -0.008856 -0.002562  0.008190  0.072577
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)   2.6202      0.5678   4.615 0.00004187 ***
## L(LRM, 1)      0.3192      0.1367   2.336  0.024735 *
## L(LRM, 2)      0.5326      0.1324   4.024  0.000255 ***
## L(LRM, 3)     -0.2687      0.1021  -2.631  0.012143 *
## LRY            0.6728      0.1312   5.129 0.00000832 ***
## L(LRY, 1)     -0.2574      0.1472  -1.749  0.088146 .
## IB0           -1.0785      0.3217  -3.353  0.001790 **
## L(IB0, 1)     -0.1062      0.5858  -0.181  0.857081
## L(IB0, 2)      0.2877      0.5691   0.505  0.616067
## L(IB0, 3)     -0.9947      0.3925  -2.534  0.015401 *
## IDE            0.1255      0.5545   0.226  0.822161
## L(IDE, 1)     -0.3280      0.7213  -0.455  0.651847
## L(IDE, 2)      1.4079      0.5520   2.550  0.014803 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0191 on 39 degrees of freedom
## Multiple R-squared:  0.988, Adjusted R-squared:  0.9843
## F-statistic: 266.8 on 12 and 39 DF,  p-value: < 0.00000000000000022
```

2. UECM (Unrestricted Error Correction Model) of the underlying ARDL.

```
UECM_BestMod <- uecm(BestMod)
```

```
summary(UECM_BestMod)
```

```
##
## Time series regression with "zooreg" data:
## Start = 1974 Q4, End = 1987 Q3
##
## Call:
## dynlm::dynlm(formula = full_formula, data = data, start = start,
##      end = end)
##
## Residuals:
##      Min          1Q      Median          3Q      Max
## -0.029939 -0.008856 -0.002562  0.008190  0.072577
```

```
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
## (Intercept)   2.62019    0.56777    4.615 0.00004187 ***
## L(LRM, 1)     -0.41685    0.09166   -4.548 0.00005154 ***
## L(LRY, 1)      0.41538    0.11761    3.532   0.00108 **
## L(IBO, 1)     -1.89172    0.39111   -4.837 0.00002093 ***
## L(IDE, 1)      1.20534    0.44690    2.697   0.01028 *
## d(L(LRM, 1)) -0.26394    0.10192   -2.590   0.01343 *
## d(L(LRM, 2))  0.26867    0.10213    2.631   0.01214 *
## d(LRY)         0.67280    0.13116    5.129 0.00000832 ***
## d(IBO)        -1.07852    0.32170   -3.353   0.00179 **
## d(L(IBO, 1))  0.70701    0.46874    1.508   0.13953
## d(L(IBO, 2))  0.99468    0.39251    2.534   0.01540 *
## d(IDE)         0.12546    0.55445    0.226   0.82216
## d(L(IDE, 1)) -1.40786    0.55204   -2.550   0.01480 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0191 on 39 degrees of freedom
## Multiple R-squared:  0.7458, Adjusted R-squared:  0.6676
## F-statistic: 9.537 on 12 and 39 DF,  p-value: 0.00000003001
```

3. RECM (Restricted Error Correction Model) of the underlying ARDL Obs: allowing the constant to join the short-run relationship (case 2), instead of the long-run (case 3)

```
RECM_BestMod <- recm(UECM_BestMod, case = 2)
```

```
summary(RECM_BestMod)
```

```
##
## Time series regression with "zooreg" data:
## Start = 1974 Q4, End = 1987 Q3
##
## Call:
## dynlm::dynlm(formula = full_formula, data = data, start = start,
##              end = end)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.029939 -0.008856 -0.002562  0.008190  0.072577
##
## Coefficients:
##              Estimate Std. Error t value    Pr(>|t|)
```

```
## d(L(LRM, 1)) -0.26394    0.09008   -2.930    0.005405 **
## d(L(LRM, 2))  0.26867    0.09127    2.944    0.005214 **
## d(LRY)        0.67280    0.11591    5.805 0.000000703 ***
## d(IBO)       -1.07852    0.30025   -3.592    0.000837 ***
## d(L(IBO, 1))  0.70701    0.44359    1.594    0.118300
## d(L(IBO, 2))  0.99468    0.36491    2.726    0.009242 **
## d(IDE)        0.12546    0.48290    0.260    0.796248
## d(L(IDE, 1)) -1.40786    0.48867   -2.881    0.006160 **
## ect          -0.41685    0.07849   -5.311 0.000003633 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01819 on 43 degrees of freedom
## (0 observations deleted due to missingness)
## Multiple R-squared:  0.7613, Adjusted R-squared:  0.7113
## F-statistic: 15.24 on 9 and 43 DF,  p-value: 0.0000000009545
```

4. long-run levels relationship (cointegration)

```
bounds_f_test(BestMod, case = 2)
```

```
##
## Bounds F-test (Wald) for no cointegration
##
## data: d(LRM) ~ L(LRM, 1) + L(LRY, 1) + L(IBO, 1) + L(IDE, 1) + d(L(LRM, 1)) + d(L(LRM, 2))
## F = 5.1168, p-value = 0.004418
## alternative hypothesis: Possible cointegration
## null values:
##      k      T
##      3 1000
```

5. Long-run multipliers (with standard errors, t-statistics and p-values)

```
multipliers(BestMod)
```

```
##      Term      Estimate Std. Error   t value      Pr(>|t|)
## 1 (Intercept)  6.2856579   0.7719160  8.142930 0.0000000006107445
## 2          LRY   0.9964676   0.1239310  8.040503 0.0000000008358472
## 3          IBO  -4.5381160   0.5202961 -8.722180 0.0000000001058619
## 4          IDE   2.8915201   0.9950853  2.905801 0.0060092393605960
```

```
#
Result <- coint_eq(BestMod, case = 2)
```

5.6.2.4 Make the plot

```
Datos <- cbind.zoo(LRM = denmark[, "LRM"], Result)
Datos <- xts(Datos)
plot(Datos, legend.loc = "right")
```

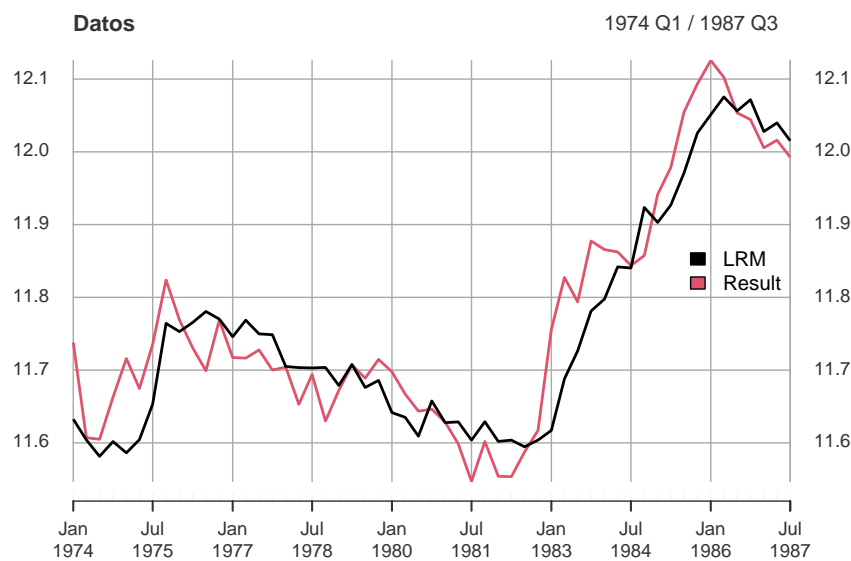


Figure 5.14: Gráfica de la ecuación de cointegración

5.7 Filtro Kalman

En secciones pasadas, entendimos la naturaleza del filtrado univariado como, por ejemplo, el *Filtro Hodrick-Prescott*. Este ha sido utilizado ampliamente en la literatura macroeconómica, sobre todo para la estimación del Producto Potencial y, de este modo, la Brecha del Producto. Sin embargo, este método presenta dos limitaciones importantes: 1) es univariado, es decir, únicamente se explica por sí mismo y, 2) el llamado “problema de colas”, que sí bien entendimos que se puede solucionar aplicando la metodología de “colas pesadas”, la estimación del λ es un tanto ambigua.

De este modo, el **Filtro Kalman** puede ser una alternativa viable y metodológicamente robusta, sobre todo si queremos trabajar con variables no observables

como el *Producto Potencial* o, incluso, la *Tasa de interés neutral*, variable clave en la decisión de Política Monetaria de los Bancos Centrales.

En este sentido, el **Filtro Kalman** nació a partir del artículo seminal de Rudolf E. Kálmán (1960) *A New Approach to Linear Filtering and Prediction Problems*. La idea principal de Kálmán era describir los procesos estocásticos con un modelo de “estado” y resolver, de alguna manera, el problema de estimación con un ciclo *predicción-corrección* ejecutado en tiempo real. Así, el **Filtro Kalman** es un algoritmo de estimación recursivo a partir de mediciones ruidosas. De este modo, también podemos entender al **Filtro Kalman** como una especie de filtro multivariado. Sus aplicaciones no se han limitado únicamente a la Economía, pues ha sido utilizado para trabajos de la *NASA* como el pronóstico de la posición y velocidad del *Apollo 11* en 1969.

5.7.1 Modelos Estado-Espacio y Teoría Económica

Para hacer uso del **Filtro Kalman**, necesitamos de un modelo estructural que vincule a nuestras variables (observables y no observables), representado en la *Forma Estado-Espacio*. En general, casi todos los modelos de series de tiempo pueden ser representados como *Modelos Estado-Espacio*. El modelo estructural puede basarse en la teoría económica como, por ejemplo, un modelo DSGE. Para efectos prácticos, tomaremos como ejemplo el *Modelo de Proyección Trimestral (QPM)* del IMF (2020). No nos enfocaremos en el por qué y cómo del *Modelo QPM*, pues el objetivo principal es entender el **Filtro Kalman** y su aplicación.

5.7.1.1 Modelo de Proyección Trimestral (QPM) del IMF (2020)

Supongamos una economía pequeña y abierta, en donde:

- **Demanda Agregada (Curva IS)**

$$\hat{y}_t = b_1 \hat{y}_{t-1} - b_2 mci_t + b_3 \hat{y}_t^* + \varepsilon_t^y \quad (5.56)$$

$$mci_t = b_4 \hat{r}_t + (1 - b_4)(-\hat{z}_t) \quad (5.57)$$

donde:

- \hat{y}_t = Brecha del Producto Doméstico,
- \hat{y}_{t-1} = Brecha del Producto Doméstico rezagada,
- mci_t = Índice de Condiciones Monetarias Real,
- \hat{y}_t^* = Brecha del Producto Foráneo,
- ε_t^y = Choque de Demanda,
- \hat{r}_t = Brecha de la tasa de interés real y

- \hat{z}_t = Brecha del tipo de cambio real.

- **Inflación (Curva de Phillips)**

$$\pi_t = a_1\pi_{t-1} + (1 - a_1)E\{\pi_{t+1}\} + a_2rmc_t + \varepsilon_t^\pi \quad (5.58)$$

$$rmc_t = a_3\hat{y}_t + (1 - a_3)\hat{z}_t \quad (5.59)$$

donde:

- π_t = Inflación,
- π_{t-1} = Inflación rezagada,
- $E\{\pi_{t+1}\}$ = Expectativa de inflación,
- rmc_t = Costo Marginal Real,
- ε_t^π = Choque de Oferta (cost-push shock),
- \hat{y}_t = Brecha del Producto Doméstico y
- \hat{z}_t = Brecha del tipo de cambio real.

- **Paridad Descubierta de Tasas de Interés (UIP)**

$$S_t = (1 - e_1)E_t\{S_{t+1}\} + e_1 \left[S_{t-1} + \frac{2(\pi_t - \pi_t^* + \Delta\bar{z}_t)}{4} + \frac{i_t^* - i_t + prem_t}{4} \right] + \varepsilon_s \quad (5.60)$$

donde:

- $E_t\{S_{t+1}\}$ = Expectativas sobre el tipo de cambio nominal futuro,
- S_{t-1} = Tipo de cambio nominal rezagado,
- π_t = Inflación,
- π_t^* = Inflación foránea,
- $\Delta\bar{z}_t$ = Depreciación del tipo de cambio tendencial,
- i_t^* = Tasa de interés real foránea,
- i_t = Tasa de interés real doméstica,
- $prem_t$ = Premio Riesgo-País y
- ε_s = Choque del Tipo de cambio.

- **Función de Reacción de Política Monetaria (Regla de Taylor)**

$$i_t = g_1i_{t-1} + (1 - g_1) \{i_t^n + g_2(E_t[\pi_{t+N}^4] - \pi_{t+N}^T) + g_3\hat{y}_t\} + \varepsilon_t \quad (5.61)$$

donde:

- i_t = Tasa de interés (o política),
- i_{t-1} = Tasa de interés rezagada,

- i_t^n = Tasa de interés natural o neutral,
- $\mathbb{E}_t[\pi_{t+N}^A]$ = Inflación esperada (YoY) en el periodo $t+N$,
- π_{t+N}^T = Meta de inflación,
- \hat{y}_t = Brecha del Producto Doméstica y
- ε_t = Choque de política monetaria no sistémico.

• **Bloque externo**

En este caso, todas las variables siguen procesos autorregresivos simples

Brecha del producto foráneo	Tasa natural (tendencia) de interés real foránea
$\hat{y}_t^* = \rho_y \hat{y}_{t-1}^* + \varepsilon_t^*$	$\tilde{r}_t^* = \rho_{\tilde{r}} \tilde{r}_{t-1}^* + (1 - \rho_{\tilde{r}}) \tilde{r}^{*SS} + \varepsilon_t^*$
Tasa de interés nominal foránea	Brecha de la tasa de interés real foránea
$i_t^* = \rho_i i_{t-1}^* + (1 - \rho_i)(r_t^* + \pi_t^*) + \varepsilon_t^*$	$\hat{r}_t^* = r_t^* - \tilde{r}_t^*$
Tasa de interés real foránea	Inflación foránea
$r_t^* = i_t^* - \pi_t^*$	$\pi_t^* = \rho_{\pi} \pi_{t-1}^* + (1 - \rho_{\pi}) \pi^{*SS} + \varepsilon_t^*$

5.7.1.2 Modelos Estado-Espacio

La idea general detrás de los *Modelos Estado-Espacio* es que una (múltiples) variable observable Y_1, \dots, Y_T depende de un posible estado no observable X_t que sigue un proceso estocástico. La relación entre Y_t y X_t es descrita por la **Ecuación de Observación**

$$Y_t = CX_T + v_t \quad (5.62)$$

donde C es una matriz de parámetros conocidos y v_t es un vector de errores observados (si los hay), los cuales son independientes e idénticamente distribuidos con media cero y matriz de covarianza R , es decir, $v_t \sim iid(0, R)$. De este modo, la ecuación (5.62) describe la relación estática entre las variables observables y no observables.

La ecuación que describe la dinámica de las variables de estado, también conocida como **Ecuación de Transición**, es definida como

$$X_t = AX_{t-1} + \varepsilon_t, \quad (5.63)$$

en donde A es una matriz de parámetros conocidos, ε_t es un vector de errores, los cuales son independientes e idénticamente distribuidos con media cero y matriz de covarianza Q , es decir, $\varepsilon_t \sim iid(0, Q)$. Adicionalmente, se asume que v_t y ε_t son independientes entre sí, es decir, $Cov(v_t, \varepsilon_t) = 0$.

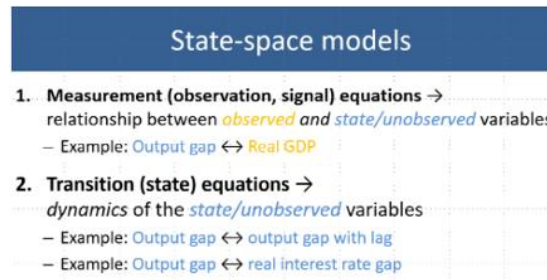


Figure 5.15: State-space models. Recuperado de IMF (2020).

En la *Figura 5.15* se puede observar un ejemplo de la relación de las variables en los *Modelos Estado-Espacio* en el *QPM* básico del IMF (2020).

Afortunadamente, nuestro *Modelo QPM* está construido de tal manera que tenemos **ecuaciones de transición** y **observación** que vinculan las variables observables con las no observables:

```
% =====
% Basic Quarterly Projection Model (QPM)
% =====

!transition_variables
'Real GDP (100*log)'                                L_GDP
'Trend in Real GDP (100*log)'                        L_GDP_BAR
'Output Gap (in %)'                                  L_GDP_GAP
'Quarterly Growth in Real GDP(in % pa)'              DLA_GDP
'Real GDP Growth YoY (in % pa)'                      D4L_GDP
'Real GDP Trend Growth QoQ annualized (in % pa)'     DLA_GDP_BAR

'Real Monetary Condition Index (in % pa)'             MCI

'CPI (level, 100*log)'                                L_CPI
'CPI Inflation QoQ annualized (in % pa)'              DLA_CPI
'Expected CPI Inflation QoQ annualized (in % pa)'     E_DLA_CPI
'CPI Inflation YoY (in % pa)'                        D4L_CPI
'Inflation Target (in % pa)'                         D4L_CPI_TAR

'Real Marginal Cost (in %)'                           RMC

'Nominal Exchange Rate (LCY/FCY, 100*log)'            L_S
'Nominal Exch. Rate Depreciation QoQ annualized (in % pa)' DLA_S
'Nominal Exch. Rate Depreciation YoY (in % pa)'       D4L_S
'Country Risk Premium (in % pa)'                     PREM
```

```

'Nominal Policy Interest Rate (in % pa)'          RS
'Real Interest Rate (in % pa)'                    RR
'Trend Real Interest Rate (in % pa)'              RR_BAR
'Real Interest Rate Gap (in %)'                   RR_GAP
'Nominal Policy Neutral Interest Rate (in % pa)'   RSNEUTRAL

'Real Exchange Rate (level, 100*log)'              L_Z
'Trend Real Exchange Rate (level, 100*log)'        L_Z_BAR
'Real Exchange Rate Gap (in %)'                   L_Z_GAP
'Real Exchange Rate Depreciation QoQ annualized (in % pa)' DLA_Z
'Trend Real Exchange Rate Depreciation QoQ annualized(in % pa)' DLA_Z_BAR

'Foreign Output Gap (in %)'                       L_GDP_RW_GAP
'Foreign Nominal Interest Rate (in % pa)'          RS_RW
'Foreign Real Interest Rate (in % pa)'             RR_RW
'Foreign Real Interest Rate Trend (in % pa)'       RR_RW_BAR
'Foreign Real Interest Rate Gap (in %)'            RR_RW_GAP
'Foreign CPI (level, 100*log)'                    L_CPI_RW
'Foreign Inflation QoQ annualized (in % pa)'       DLA_CPI_RW

% ----- %
!transition_shocks
'Shock: Output gap (demand)'                      SHK_L_GDP_GAP
'Shock: CPI inflation (cost-push)'                SHK_DLA_CPI
'Shock: Exchange rate (UIP)'                     SHK_L_S
'Shock: Interest rate (monetary policy)'          SHK_RS
'Shock: Inflation target'                        SHK_D4L_CPI_TAR

'Shock: Real interest rate'                      SHK_RR_BAR
'Shock: Real exchange rate depreciation'          SHK_DLA_Z_BAR
'Shock: Potential GDP growth'                   SHK_DLA_GDP_BAR

'Shock: Foreign output gap'                      SHK_L_GDP_RW_GAP
'Shock: Foreign nominal interest rate'           SHK_RS_RW
'Shock: Foreign inflation'                      SHK_DLA_CPI_RW
'Shock: Foreign real interest rate'              SHK_RR_RW_BAR

% ----- %
!parameters
b1 b2 b3 b4
a1 a2 a3
e1
g1 g2 g3

rho_D4L_CPI_TAR
rho_DLA_Z_BAR

```

```

rho_RR_BAR
rho_DLA_GDP_BAR

rho_L_GDP_RW_GAP
rho_RS_RW
rho_DLA_CPI_RW
rho_RR_RW_BAR

ss_D4L_CPI_TAR
ss_DLA_Z_BAR
ss_RR_BAR
ss_DLA_GDP_BAR

ss_DLA_CPI_RW
ss_RR_RW_BAR

%% ----- %
!transition_equations
%% === Aggregate demand (the IS curve) ===
L_GDP_GAP = b1*L_GDP_GAP{-1} - b2*MCI + b3*L_GDP_RW_GAP + SHK_L_GDP_GAP;

%-- Real monetary conditions index
MCI = b4*RR_GAP + (1-b4)*(- L_Z_GAP);

%% === Inflation (the Phillips curve) ===
DLA_CPI = a1*DLA_CPI{-1} + (1-a1)*DLA_CPI{+1} + a2*RMC + SHK_DLA_CPI;

%-- Real marginal cost
RMC = a3*L_GDP_GAP + (1-a3)*L_Z_GAP;

%- expected inflation
E_DLA_CPI = DLA_CPI{+1};

%% === Monetary policy reaction function (a forward-looking Taylor-type Rule) ===
RS = g1*RS{-1} + (1-g1)*(RSNEUTRAL + g2*(D4L_CPI{+4} - D4L_CPI_TAR{+4}) + g3*L_GDP_GAP);

%- Neutral nominal policy interest rate
RSNEUTRAL = RR_BAR + D4L_CPI{+1};

%% === Modified Uncovered Interest Rate Parity (UIP) condition ===
L_S = (1-e1)*L_S{+1} + e1*(L_S{-1} + 2/4*(D4L_CPI_TAR - ss_DLA_CPI_RW + DLA_Z_BAR)) +

%% === Definitions ===

%- Fisher equation (RIR)
RR = RS - D4L_CPI{+1};

```

```

%- Real exchange rate (RER)
L_Z = L_S + L_CPI_RW - L_CPI;

%- Long-term version of UIP (consistency of trends)
DLA_Z_BAR{+1} = RR_BAR - RR_RW_BAR - PREM;

%% === Identities ===
DLA_GDP_BAR = 4*(L_GDP_BAR - L_GDP_BAR{-1});
DLA_Z_BAR   = 4*(L_Z_BAR - L_Z_BAR{-1});
DLA_Z       = 4*(L_Z - L_Z{-1});
DLA_GDP     = 4*(L_GDP - L_GDP{-1});
DLA_CPI     = 4*(L_CPI - L_CPI{-1});
DLA_S       = 4*(L_S - L_S{-1});

D4L_GDP     = L_GDP - L_GDP{-4};
D4L_CPI     = L_CPI - L_CPI{-4};
D4L_S       = L_S - L_S{-4};

%% === Gaps ===
RR_GAP      = RR - RR_BAR;
L_Z_GAP     = L_Z - L_Z_BAR;
L_GDP_GAP   = L_GDP - L_GDP_BAR;

%% === Trends ===
D4L_CPI_TAR = rho_D4L_CPI_TAR*D4L_CPI_TAR{-1} + (1-rho_D4L_CPI_TAR)*ss_D4L_CPI_TAR + SHK_D4L_CPI_TAR;
DLA_Z_BAR   = rho_DLA_Z_BAR*DLA_Z_BAR{-1} + (1-rho_DLA_Z_BAR)*ss_DLA_Z_BAR + SHK_DLA_Z_BAR;
RR_BAR      = rho_RR_BAR*RR_BAR{-1} + (1-rho_RR_BAR)*ss_RR_BAR + SHK_RR_BAR;
DLA_GDP_BAR = rho_DLA_GDP_BAR*DLA_GDP_BAR{-1} + (1-rho_DLA_GDP_BAR)*ss_DLA_GDP_BAR + SHK_DLA_GDP_BAR;

%% === Foreign Sector Equations ===
L_GDP_RW_GAP = rho_L_GDP_RW_GAP*L_GDP_RW_GAP{-1} + SHK_L_GDP_RW_GAP;
DLA_CPI_RW   = rho_DLA_CPI_RW*DLA_CPI_RW{-1} + (1-rho_DLA_CPI_RW)*ss_DLA_CPI_RW + SHK_DLA_CPI_RW;
RS_RW       = rho_RS_RW*RS_RW{-1} + (1-rho_RS_RW)*(RR_RW_BAR + DLA_CPI_RW) + SHK_RS_RW;
RR_RW_BAR   = rho_RR_RW_BAR*RR_RW_BAR{-1} + (1-rho_RR_RW_BAR)*ss_RR_RW_BAR + SHK_RR_RW_BAR;
RR_RW       = RS_RW - DLA_CPI_RW;
RR_RW_GAP   = RR_RW - RR_RW_BAR;

DLA_CPI_RW   = 4*(L_CPI_RW - L_CPI_RW{-1});

%% ----- %
!measurement_variables
OBS_L_GDP
OBS_L_CPI
OBS_RS
OBS_L_S

```

```

OBS_D4L_CPI_TAR

OBS_L_GDP_RW_GAP
OBS_DLA_CPI_RW
OBS_RS_RW

!measurement_equations
OBS_L_GDP = L_GDP;
OBS_L_CPI = L_CPI;
OBS_RS    = RS;
OBS_L_S   = L_S;
OBS_D4L_CPI_TAR = D4L_CPI_TAR;

OBS_L_GDP_RW_GAP = L_GDP_RW_GAP;
OBS_DLA_CPI_RW   = DLA_CPI_RW;
OBS_RS_RW        = RS_RW;

%% ----- %
Legend
_GAP      cyclical deviation from a trend
_BAR      trend (equilibrium)
ss_       steady-state value
DLA_      q-o-q change
D4L_      y-o-y change
_RW       foreign variable
SHK_      equation residual

```

Sin embargo, la **Ecuación de Transición** de los estados presentes dependen de los estados pasados, mientras que el *Modelo QPM* es prospectivo, es decir, las variables dependen de sus expectativas así que la ecuación toma la forma de

$$F\mathbb{E}_t[X_{t+1}] + GX_t + HX_{t-1} + M\eta_t = 0 \quad (5.64)$$

donde F , G , H , M son matrices de parámetros estructurales conocidos, \mathbb{E}_t es el operador de expectativas y η_t es un vector de choques estructurales. Para pasar de este modelo a la *Forma Estado-Espacio*, tenemos que resolver para *Expectativas Racionales*. Para esto, proponemos una *hipótesis de solución*

$$X_t = AX_{t-1} + P\eta_t \quad \text{con} \quad \varepsilon_t \equiv P\eta_t \quad (5.65)$$

Ahora, suponemos que η_t sigue un proceso autorregresivo, es decir, tiene su propia dinámica:

$$\eta_{t+1} = S\eta_t + \gamma_{t+1} \quad \text{donde} \quad \gamma_{t+1} \sim iid(0, W) \quad (5.66)$$

Así, sabemos que

$$\mathbb{E}_t[X_{t+1}] = AX_t + P\mathbb{E}_t[\eta_{t+1}] = A(AX_{t-1} + P\eta_t) + PS\eta_t \quad (5.67)$$

De este modo, si sustituimos la ecuación (5.65) en la (5.64), tenemos que las matrices deben de satisfacer la identidad

$$F[A(AX_{t-1} + P\eta_t) + P(S\eta_t)] + G[AX_{t-1} + P\eta_t] + HX_{t-1} + M\eta_t = 0 \quad (5.68)$$

para todos los valores X_{t-1} y η_t , de tal modo que A y P son las soluciones de ecuaciones

$$FA^2 + GA + H = 0, \quad FAP + FPS + GP + M = 0 \quad (5.69)$$

De esta forma, obtenemos nuestro modelo en la *Forma Estado-Espacio*.

5.7.2 Recursión del Filtro Kalman

Ya que tenemos nuestro modelo en la *Forma Estado-Espacio* como se muestra en las ecuaciones (5.62) y (5.63), podemos hacer uso del **Filtro Kalman**.

Para comenzar, definamos $Y_{t/t-1}$ y $X_{t/t-1}$ como los valores predichos de Y_t y X_t , respectivamente, condicionado a la información del periodo $t-1$. Además, definamos $X_{t/t}$ como la estimación de X_t condicionada a la información del periodo t .

Como se mencionó en un inicio, el **Filtro Kalman** utiliza un algoritmo recursivo para cada periodo $1, 2, \dots, T$. Cada recursión está compuesta de dos pasos. Supongamos que tenemos los datos observados y las estimaciones del estado para el periodo $t-1$, es decir, conocemos Y_{t-1} y $X_{t-1/t-1}$. De este modo, para el periodo t :

- **Paso 1.** Predicción del estado y observación de las variables en el periodo t dada la información de $t-1$. Como no tenemos nueva información, el promedio de la predicción se basa en el *Modelo Estado-Espacio* (5.62) - (5.63):

$$X_{t/t-1} = AX_{t-1/t-1} \quad (5.70)$$

$$Y_{t/t-1} = CX_{t/t-1} \quad (5.71)$$

- **Paso 2.** Actualización de la estimación del estado dada la información en el periodo t . Una vez que conocemos la nueva información del periodo t , existirá una discrepancia entre la información actual (Y_t) y la predicción del modelo ($Y_{t/t-1}$). El **Filtro Kalman** utiliza dicha discrepancia para actualizar la estimación de los estados no observables utilizando la matriz K_t , que se llama **Ganancia de Kalman**:

$$X_{t/t} = X_{t/t-1} + K_t(Y_t - Y_{t/t-1}) \quad (5.72)$$

5.7.3 Ganancia de Kalman

Supongamos que los errores siguen una distribución normal, es decir, $v_t \sim N(0, R)$, $\varepsilon_t \sim N(0, Q)$.

Para la primera estimación (Paso 1) de la *Recursión de Kalman*, $P_{t/t-1}$ denota la varianza del error de predicción de las variables de estado (X_t) dada la información del periodo $t-1$, o sea, $P_{t/t-1} = \text{Var}_{t-1}[(X_t - X_{t/t-1})]$. De manera similar, $F_{t/t-1}$ es la varianza del error de predicción de las variables observables (Y_t) dada la información del periodo $t-1$, o bien, $F_{t/t-1} = \text{Var}_{t-1}[(Y_t - Y_{t/t-1})]$.

Para la actualización (Paso 2) de la *Recursión de Kalman*, $P_{t/t}$ denota la varianza del error de las variables de estado (X_t) dada la información del periodo t , es decir, $P_{t/t} = \text{Var}_t[(X_t - X_{t/t})]$.

Así, la distribución multivariada del vector (Y_t, X_t) toma la forma

$$\begin{pmatrix} Y_t \\ X_t \end{pmatrix} \bigg| Y_t, \dots, Y_{t-1} \sim N \left(\begin{pmatrix} Y_{t/t-1} \\ X_{t/t-1} \end{pmatrix}, \begin{pmatrix} F_{t/t-1} & CP_{t/t-1} \\ P_{t/t-1}C' & P_{t/t-1} \end{pmatrix} \right). \quad (5.73)$$

De (5.73), la distribución condicional de X_t dado Y_t toma la forma de

$$X_t | Y_t, Y_1, \dots, Y_{t-1} \sim N \begin{pmatrix} X_{t/t-1} + P_{t/t-1}C'F_{t/t-1}^{-1}[Y_t - Y_{t/t-1}], \\ P_{t/t-1} - P_{t/t-1}C'F_{t/t-1}^{-1}CP_{t/t-1} \end{pmatrix} \quad (5.74)$$

Y de la ecuación (5.74) obtenemos directamente que

$$X_{t/t} = X_{t/t-1} + P_{t/t-1}C'F_{t/t-1}^{-1}[Y_t - Y_{t/t-1}] = X_{t/t-1} + K_t[Y_t - Y_{t/t-1}], \quad (5.75)$$

$$P_{t/t} = P_{t/t-1} - P_{t/t-1}C'F_{t/t-1}^{-1}CP_{t/t-1} = (I - K_tC)P_{t/t-1}, \quad (5.76)$$

en donde la **Ganancia de Kalman** es

$$K_t = P_{t/t-1} C' F_{t/t-1}^{-1}. \quad (5.77)$$

Notese que $F_{t/t-1}$ sigue:

$$F_{t|t-1} = \text{Var}_{t-1} [(Y_t - Y_{t|t-1})] = \text{Var}_{t-1} [(CX_t + v_t - CX_{t|t-1})] = CP_{t|t-1}C' + R \quad (5.78)$$

Por el contrario, $P_{t/t-1}$ sigue:

$$P_{t|t-1} = \text{Var}_{t-1} [(X_t - X_{t|t-1})] \quad (5.79)$$

$$= \text{Var}_{t-1} [(AX_{t-1} + \varepsilon_t - AX_{t-1|t-1})] \quad (5.80)$$

$$= AP_{t-1|t-1}A' + Q \quad (5.81)$$

$$= A(I - K_{t-1}C)P_{t-1|t-2}A' + Q \quad (5.82)$$

Y si sustituimos $F_{t/t-1}$ y $P_{t/t-1}$ en las ecuaciones (5.75) - (5.77), obtenemos las siguientes ecuaciones para el algoritmo del **Filtro Kalman**

$$K_t = P_{t/t-1} C' (CP_{t/t-1}C' + R)^{-1} \quad (5.83)$$

$$X_{t/t} = X_{t/t-1} + K_t[Y_t - Y_{t|t-1}] \quad (5.84)$$

$$P_{t+1|t} = A(I - K_tC)P_{t/t-1}A' + Q \quad (5.85)$$

De este modo, la **Ganancia de Kalman** en (5.83) depende de manera positiva de la varianza del error de predicción del estado ($P_{t/t-1}$) y de manera negativa de la varianza de la ecuación de observación R . Esto es bastante intuitivo:

- Si existen grandes errores en las predicciones del estado $X_{t/t-1}$ en el Paso 1 de la *Recursión de Kalman*, es decir, $P_{t/t-1}$ es grande, se le da más peso a la nueva información observada Y_t , lo cual significa que la **Ganancia de Kalman** es grande. En particular, esto es cierto para las ecuaciones de transición que no se cumplen de manera estricta, es decir, Q es grande;
- Si la información es ruidosa, o sea, R es grande, se le da menos peso a la información nueva por lo que K_t es pequeño.

5.7.4 Algoritmo del Filtro Kalman y del Suavizador de Kalman

Ya que encontramos las expresiones más relevantes, podemos combinar todas para entender cómo funciona el **Filtro Kalman**.

Sabemos que el filtro es recursivo. De este modo, para poder correrlo necesitamos *condiciones iniciales* para todos los estados al principio de la muestra ($X_{0|0}$), al igual que para la varianza del error de predicción del estado ($P_{0|0}$). Para muestras pequeñas, las *condiciones iniciales* pueden afectar significativamente los resultados del filtro. Si no tenemos información previa, una opción es asumir que las *condiciones iniciales* de los estados y de la varianza del error de predicción no están lejos de sus valores en estado estacionario. En particular, para modelos estacionarios, se puede establecer que las *condiciones iniciales* son iguales a los valores en estado estacionario.

Ahora, ya que hemos asignado las *condiciones iniciales*, el **Filtro Kalman** corre el algoritmo recursivo para cada periodo de tiempo $1, 2, \dots, T$. Cada recursión está compuesta de dos pasos:

- **Paso 1.** Predicción utilizando el *Modelo Estado-Espacio*:

$$X_{t/t-1} = AX_{t-1/t-1}, \quad (5.86)$$

$$Y_{t/t-1} = CX_{t/t-1}. \quad (5.87)$$

- **Paso 2.** Actualización utilizando la información:

$$K_t = P_{t/t-1}C'(CP_{t/t-1}C' + R)^{-1}, \quad (5.88)$$

$$X_{t/t} = X_{t/t-1} + K_t[Y_t - Y_{t/t-1}], \quad (5.89)$$

$$P_{t+1/t} = A(I - K_tC)P_{t/t-1}A' + Q. \quad (5.90)$$

El **Filtro Kalman** nos proporciona el estimador lineal óptimo de mínimos cuadrados del estado, dada la información hasta el periodo t . Si los errores están distribuidos normalmente, es óptimo entre todos los estimadores de X_t condicionados a la información disponible hasta ese periodo.

Para cada periodo t , el **Filtro Kalman** utiliza únicamente la información disponible hasta ese periodo para estimar X_t . Dichas estimaciones se encuentran en $X_{t/t-1}$ (del Paso 1) y en $X_{t/t}$ (del Paso 2). Sin embargo, para estimar

X_t , comúnmente es deseable utilizar toda la información disponible, hasta el último periodo T . Es decir, podría ser deseable obtener estimaciones $X_{t/T}$, que se conocen como **estimaciones suavizadas** de los estados. Precisamente eso es lo que hace el **Suavizador de Kalman (Kalman smoother)**.

Para correr el **Suavizador de Kalman**, necesitamos primero correr el **Filtro Kalman** y obtener todas las estimaciones de los estados: $X_{1/1}, X_{2/2}, \dots, X_{T/T}$.

Después, el algoritmo del **Suavizador de Kalman** funciona recursivamente hacia atrás, desde $T - 1, T - 2, \dots, 1$. Para cada periodo $t+1$, observamos la diferencia entre las estimaciones suavizadas y las estimaciones filtradas de los estados: $X_{t+1/T} - X_{t+1/t}$. El **Suavizador de Kalman** utiliza la diferencia para actualizar las estimaciones filtradas de los estados en el periodo t , con el fin de obtener las estimaciones suavizadas:

$$X_{t/T} = X_{t/t} + J_t[X_{t+1/T} - X_{t+1/t}]. \quad (5.91)$$

En la *Figura 5.16*, podemos observar los resultados del código de Matlab. Otra de las virtudes del **Filtro Kalman** es que podemos hacer una especie de descomposición de la varianza para, así, observar de donde vinieron los choques:

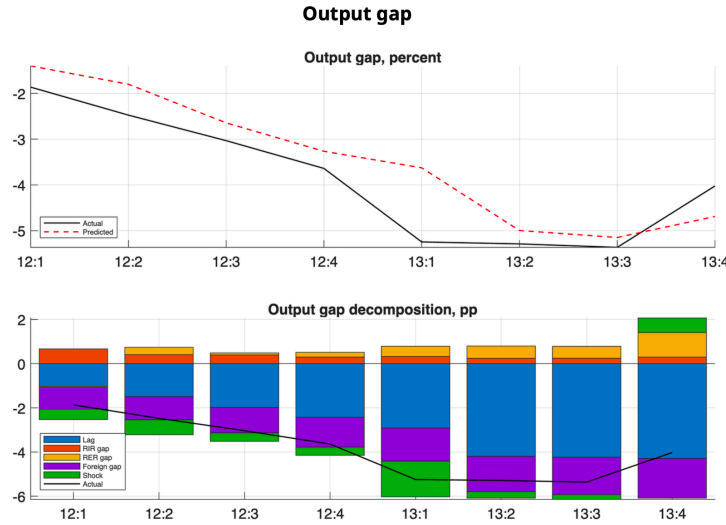


Figure 5.16: Simulación de la Descomposición de la Brecha del Producto. Elaboración propia con datos de IMF (2020).

Nota: Para este ejemplo, se ocupa el QPM básico presentado por IMF (2020) en su curso “Monetary Policy Analysis and Forecasting” que se puede tomar en la plataforma edX. Asimismo, el código de Matlab fue realizado por el IMF.

Chapter 6

Modelos Univariados y Multivariados de Volatilidad

6.1 Motivación

En este capítulo discutiremos los modelos de Heterocedasticidad Condicional Autoregresiva (ARCH, por sus siglas en inglés) y modelos de Heterocedasticidad Condicional Autoregresiva Generalizados (GARCH, por sus siglas en inglés) tienen la característica de modelar situaciones como las que ilustra la Figura 6.2. Es decir:

- 1) Existen zonas donde la variación de los datos es mayor y zonas donde la variación es más estable—a estas situaciones se les conoce como de variabilidad por clúster—, y
- 2) los datos corresponden a información de alta frecuencia.

Iniciemos por las bibliotecas necesarias:

```
library(expm)
library(Matrix)
library(ggplot2)
library(quantmod)
library(moments)
library(dynlm)
library(broom)
```

```

library(FinTS)
library(lubridate)
library(forecast)
library(readxl)
library(MASS)
library(rugarch)
library(tsbox)
library(MTS)
library(rmgarch)
library(Rcpp)

```

Para el análisis de temas financieros existe una librería de mucha utilidad llamada **quantmod**. En primer lugar esta librería permite acceder a datos financieros de un modo muy simple, es posible decargar series financieras desde *yahoo*, la *FRED* (*Federal Reserve Economic Data*), *google*, etc. Por otro lado, también es una librería que permite realizar gráficos altamente estéticos con unas cuantas líneas de código. Ahora, usaremos datos de Yahoo Finance respecto de la cotización del bitcoin (ticker: “BTC-USD”).

```

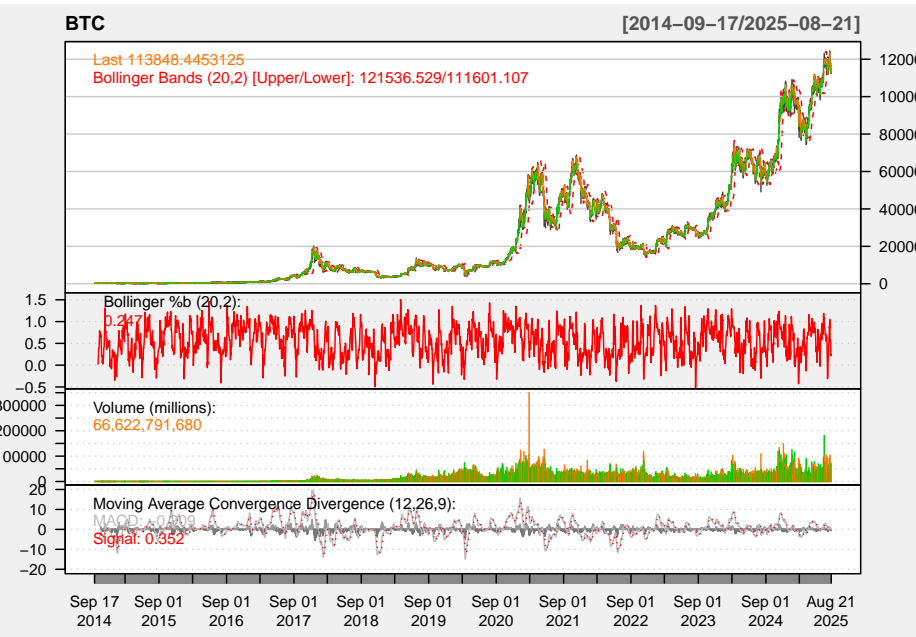
options("getSymbols.warning4.0"=FALSE)

BTC <- getSymbols("BTC-USD", src = "yahoo", auto.assign = FALSE)

BTC <- na.omit(BTC)

chartSeries(BTC, TA='addBBands();
              addBBands(draw="p");
              addVo();
              addMACD()', # subset='2021',
              theme="white")

```

head(BTC)

##	BTC-USD.Open	BTC-USD.High	BTC-USD.Low	BTC-USD.Close	BTC-USD.Volume
## 2014-09-17	465.864	468.174	452.422	457.334	21056800
## 2014-09-18	456.860	456.860	413.104	424.440	34483200
## 2014-09-19	424.103	427.835	384.532	394.796	37919700
## 2014-09-20	394.673	423.296	389.883	408.904	36863600
## 2014-09-21	408.085	412.426	393.181	398.821	26580100
## 2014-09-22	399.100	406.916	397.130	402.152	24127600
##	BTC-USD.Adjusted				
## 2014-09-17	457.334				
## 2014-09-18	424.440				
## 2014-09-19	394.796				
## 2014-09-20	408.904				
## 2014-09-21	398.821				
## 2014-09-22	402.152				

tail(BTC)

##	BTC-USD.Open	BTC-USD.High	BTC-USD.Low	BTC-USD.Close	BTC-USD.Volume
## 2025-08-16	117398.4	117996.1	117272.0	117491.4	48036922378
## 2025-08-17	117492.8	118595.8	117279.5	117453.1	45852169525
## 2025-08-18	117453.9	117614.2	114723.7	116252.3	72787808090
## 2025-08-19	116241.9	116764.5	112730.4	112831.2	71657600353

```
## 2025-08-20      112828.0      114625.8      112388.0      114274.7      67993811526
## 2025-08-21      114269.5      114691.3      113622.9      113848.4      66622791680
##
##          BTC-USD.Adjusted
## 2025-08-16      117491.4
## 2025-08-17      117453.1
## 2025-08-18      116252.3
## 2025-08-19      112831.2
## 2025-08-20      114274.7
## 2025-08-21      113848.4
```

Para fines del ejercicio de esta clase, usaremos el valor de la acción ajustado. Esto nos servirá para calcular el rendimiento diario, o puesto en lenguaje de series temporales podemos decir que usaremos la serie en diferencias logarítmicas.

```
plot(BTC$`BTC-USD.Adjusted`)
```

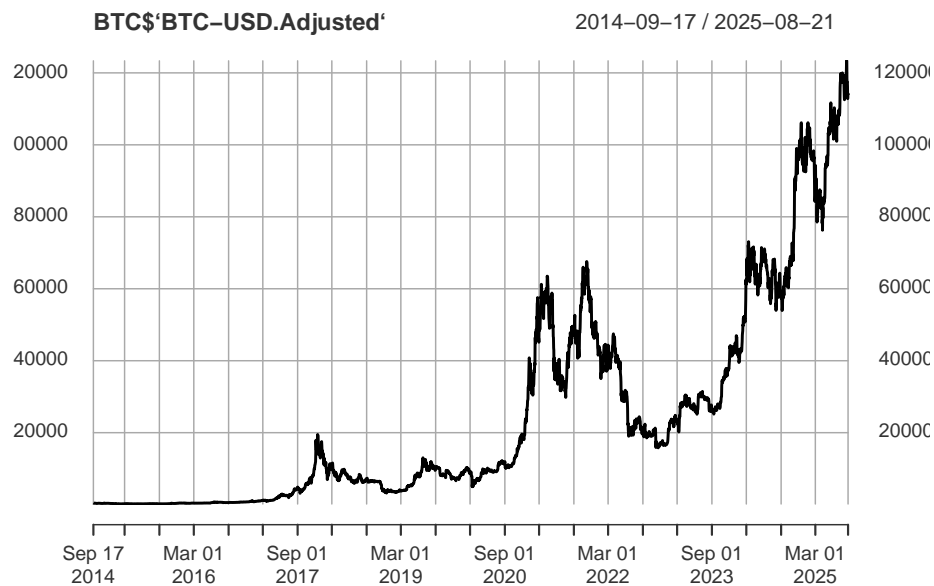


Figure 6.1: Evolución del precio del Bitcoin

Una de las preguntas relevantes al observar la serie en diferencias, es si podríamos afirmar que esta serie cumple con el supuesto de homoscedasticidad. Para ello, la Figura 6.2 muestra que las variaciones en el precio del Bitcoin muestran un escenario en el que no se cumple el supuesto de la homocedasticidad.

```
logret <- ts(diff(log(BTC$`BTC-USD.Adjusted`))[-1])
plot(logret)
```

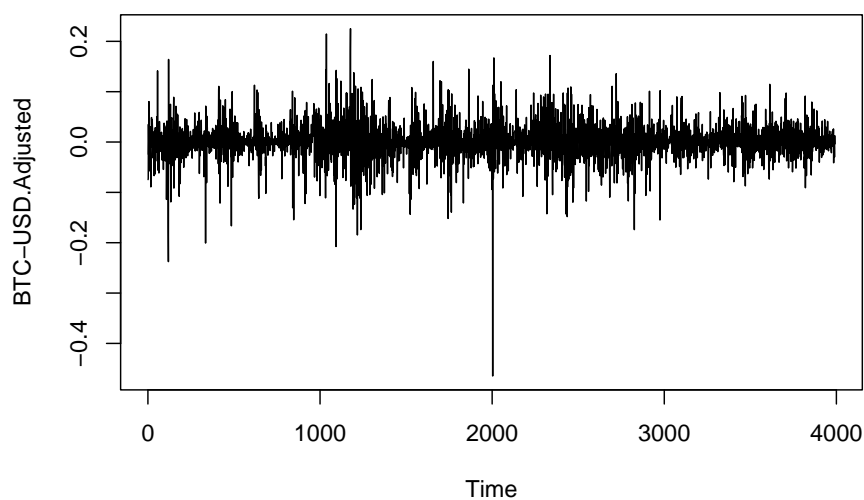


Figure 6.2: Evolución del rendimiento (diferencias logarítmicas) del Bitcoin

6.1.1 Value at Risk (VaR)

Utilicemos como ejemplo el Valor en Riesgo. Este es básicamente es un cálculo que nos permite estimar el monto que una acción o portafolio podría perder dada una probabilidad $(1 - \alpha)$. Supongamos un $\alpha = 0.05$, de esta forma, la Figura 6.3 ilustra la región de la distribución que consideraríamos como el VaR.

```
alpha <- 0.05
VaR <- quantile( logret, alpha )
round( VaR, 4 ) * 100
```

```
##      5%
## -5.57
```

```

qplot(logret , geom = 'histogram') +
  geom_histogram(fill = 'lightblue' , bins = 30) +
  geom_histogram( aes(logret[logret < quantile(logret , 0.05)]) ,
    fill = 'red' , bins = 30) +
  labs(x = 'Daily Returns')

```

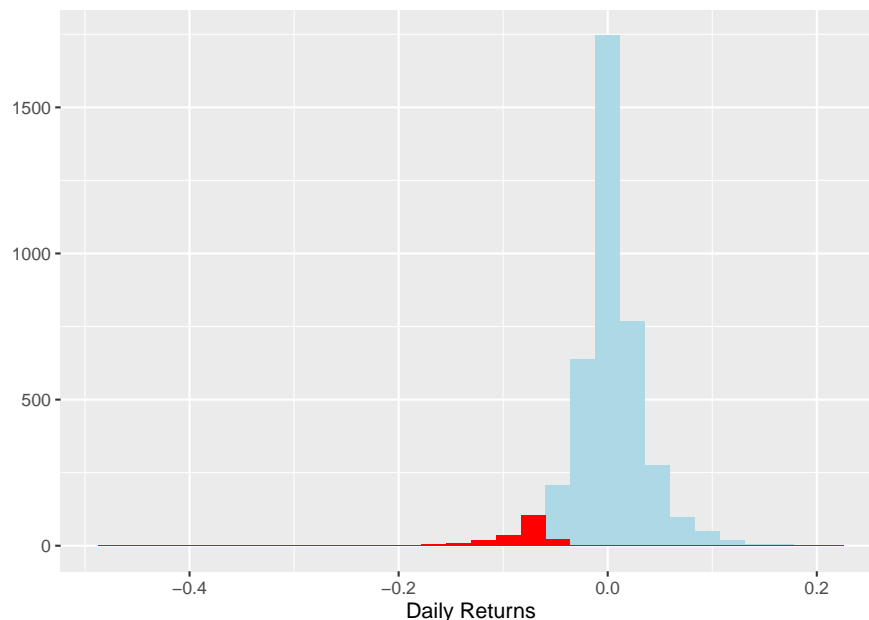


Figure 6.3: Histograma de rendimientos del Bitcoin

Ahora bien, una de las preguntas que nos podemos hacer es si los rendimientos del Bitcoin se aproximan a una distribución normal. Para ello, la Figura 6.4 ilustra esta comparación, de la cual podemos observar que una prueba de normalidad rechaza esa hipótesis—el estadístico Jarque-Bera indica que la serie de rendimientos no tiene una distribución normal—.

```

normal_dist <- rnorm(100000, mean(logret), sd(logret))
VaR_n <- quantile(normal_dist, 0.05)
ES_n <- mean(normal_dist[normal_dist < VaR_n])

ggplot()+
  geom_density(aes(logret, geom = 'density', col = 'returns'))+
  geom_density(aes(normal_dist, col = 'normal'))

```

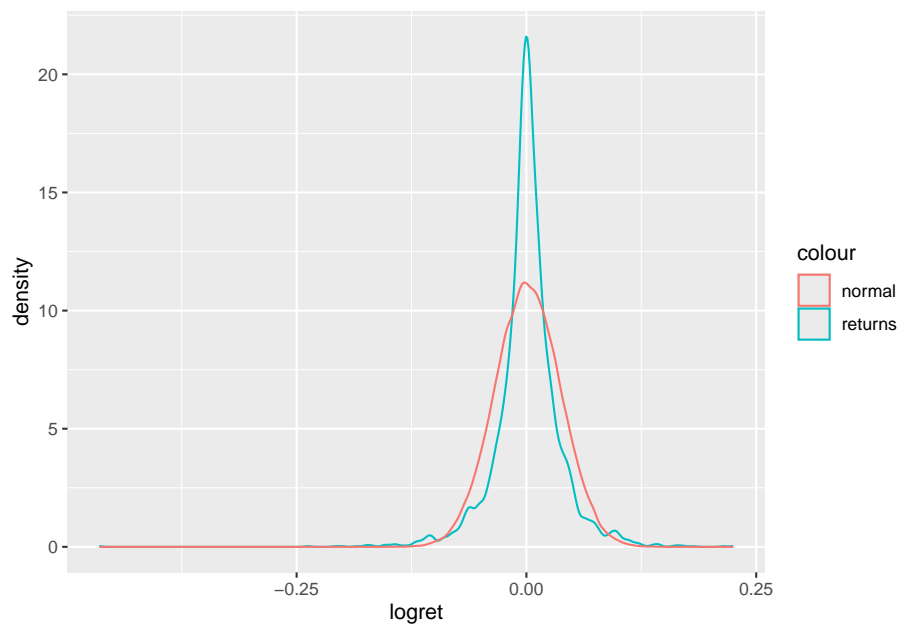


Figure 6.4: Densidad de rendimientos del Bitcoin Vs. una distribución normal

```
vector_ret <- as.vector(logret)

##Kurtosis
round(kurtosis(vector_ret),2)
```

```
## [1] 14.61
```

```
##Sesgo
round(skewness(vector_ret),2)
```

```
## [1] -0.72
```

6.2 Prueba de normalidad

$$H_o : K = S = 0$$

```
jarque.test(vector_ret)
```

```
##
##  Jarque-Bera Normality Test
##
## data:  vector_ret
## JB = 22753, p-value < 0.00000000000000022
## alternative hypothesis: greater
```

6.3 Modelos ARCH y GARCH Univariados

Para plantear el modelo, supongamos—por simplicidad—que hemos construido y estimado un modelo AR(1). Es decir, asumamos que el proceso subyacente para la media condicional está dada por:

$$X_t = a_0 + a_1 X_{t-1} + U_t \quad (6.1)$$

Donde $|a_1| < 1$ para garantizar la convergencia del proceso en el largo plazo, en el cual:

$$\begin{aligned} \mathbb{E}[X_t] &= \frac{a_0}{1 - a_1} = \mu \\ \text{Var}[X_t] &= \frac{\sigma^2}{1 - a_1^2} \end{aligned}$$

Ahora, supongamos que este tipo de modelos pueden ser extendidos y generalizados a un modelo ARMA(p, q), que incluya otras variables exógenas. Denotemos a \mathbf{Z}_t como el conjunto que incluye los componentes AR, MA y variables exógenas que pueden explicar a X_t de forma que el proceso estará dado por:

$$X_t = \mathbf{Z}_t \beta + U_t \quad (6.2)$$

Donde U_t es un proceso estacionario que representa el error asociado a un proceso ARMA(p, q) y donde siguen diendo válidos los supuestos:

$$\begin{aligned} \mathbb{E}[U_t] &= 0 \\ \text{Var}[U_t^2] &= \sigma^2 \end{aligned}$$

No obstante, en este caso podemos suponer que existe autocorrelación en el término de error al cuadrado que puede ser capturada por un proceso similar a uno de medias móviles (MA) dado por:

$$U_t^2 = \gamma_0 + \gamma_1 U_{t-1}^2 + \gamma_2 U_{t-2}^2 + \dots + \gamma_q U_{t-q}^2 + \nu_t \quad (6.3)$$

Donde ν_t es un ruido blanco y $U_{t-i} = X_{t-i} - \mathbf{Z}_{t-i} \beta$, $i = 1, 2, \dots$. Si bien los procesos son estacionarios por los supuestos antes enunciados, la varianza

condicional estará dada por:

$$\begin{aligned}\sigma_{t|t-1}^2 &= \text{Var}[U_t|\Omega_{t-1}] \\ &= \mathbb{E}[U_t^2|\Omega_{t-1}]\end{aligned}$$

Donde $\Omega_{t-1} = \{U_{t-1}, U_{t-2}, \dots\}$ es el conjunto de toda la información pasada de U_t y observada hasta el momento $t-1$, por lo que:

$$U_t|\Omega_{t-1} \sim \mathbb{D}(0, \sigma_{t|t-1}^2)$$

Así, de forma similar a un proceso MA(q) podemos decir que la varianza condicional tendrá efectos ARCH de orden q (ARCH(q)) cuando:

$$\sigma_{t|t-1}^2 = \gamma_0 + \gamma_1 U_{t-1}^2 + \gamma_2 U_{t-2}^2 + \dots + \gamma_q U_{t-q}^2 \quad (6.4)$$

Donde $\mathbb{E}[\nu_t] = 0$ y γ_0 y $\gamma_i \geq 0$, para $i = 1, 2, \dots, q-1$ y $\gamma_q > 0$. Estas condiciones son necesarias para garantizar que la varianza sea positiva. En general, la varianza condicional se expresa de la forma $\sigma_{t|t-1}^2$, no obstante, para facilitar la notación, nos referiremos en cada caso a esta simplemente como σ_t^2 .

Podemos generalizar esta situación si asumimos a la varianza condicional como dependiente de los valores de la varianza rezagados, es decir, como si fuera un proceso AR de orden p para la varianza y juntándolo con la ecuación (6.4). Bollerslev (1986) y Taylor (1986) generalizaron el problema de heterocedasticidad condicional. El modelo se conoce como GARCH(p, q), el cual se especifica como:

$$\begin{aligned}\sigma_t^2 &= \gamma_0 + \gamma_1 U_{t-1}^2 + \gamma_2 U_{t-2}^2 + \dots + \gamma_q U_{t-q}^2 \\ &\quad + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2 + \dots + \beta_p \sigma_{t-p}^2\end{aligned} \quad (6.5)$$

Donde las condiciones de no negatividad son que $\gamma_0 > 0$, $\gamma_i \geq 0$, $i = 1, 2, \dots, q-1$, $\beta_j \geq 0$, $j = 1, 2, \dots, p-1$, $\gamma_q > 0$ y $\beta_p > 0$. Además, otra condición de convergencia es que:

$$\gamma_1 + \dots + \gamma_q + \beta_1 + \dots + \beta_p < 1$$

Usando el operador rezago L en la ecuación (6.6) podemos obtener:

$$\sigma_t^2 = \gamma_0 + \alpha(L)U_t^2 + \beta(L)\sigma_t^2 \quad (6.6)$$

De donde podemos establecer:

$$\sigma_t^2 = \frac{\gamma_0}{1 - \beta(L)} + \frac{\alpha(L)}{1 - \beta(L)} U_t^2 \quad (6.7)$$

Por lo que la ecuación (6.6) del GARCH(p, q) representa un ARCH(∞):

$$\sigma_t^2 = \frac{a_0}{1 - b_1 - b_2 - \dots - b_p} + \sum_{i=1}^{\infty} U_{t-i}^2 \quad (6.8)$$

6.3.1 Ejemplo ARCH(1)

Hasta ahora, las distribuciones utilizadas para medir el Valor en Riesgo de un activo (Bitcoin, en este caso) asumen que no existe correlación serial en los retornos diarios del activo. Observemos un par de gráficas de la función de autocorrelación para corroborar este hecho—ver la Figura 6.5—.

```
acf(logret)
```

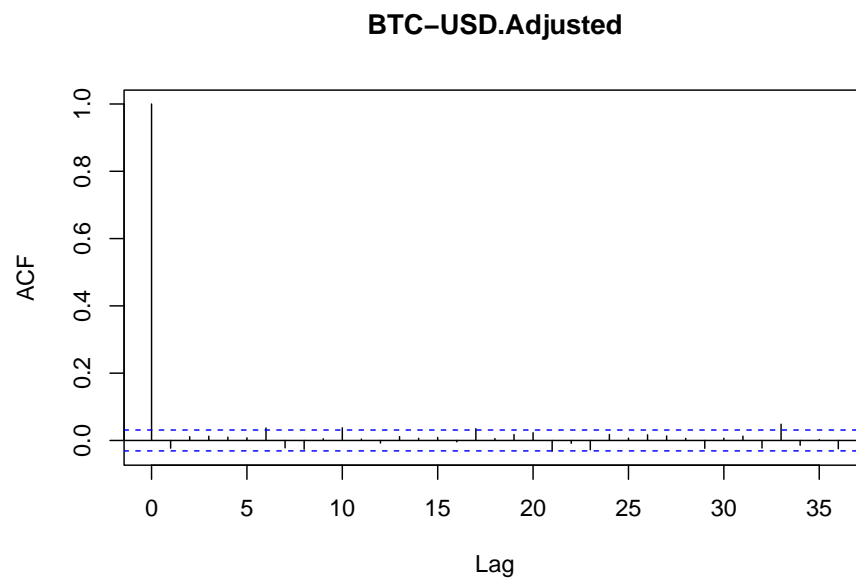


Figure 6.5: Función de autocorrelación parcial de lo rendimientos del Bitcoin

La idea de clusterización de volatilidad asume que períodos de alta volatilidad serán seguidos por alta volatilidad y viceversa. Por esta razón, la función de autocorrelación útil para saber si existen clústers de volatilidad es utilizando el valor absoluto, ya que lo que importa es saber si la serie está autocorrelacionada en la magnitud de los movimientos. La Figura 6.6 muestra el comportamiento de los rendimientos en valor absoluto y la Figura 6.7 la función de autocorrelación parcial.

```
plot(abs(logret))
```

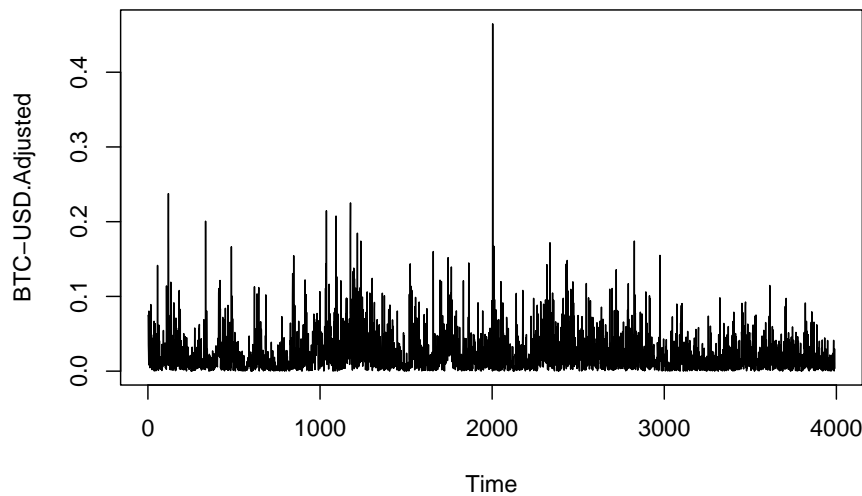



Figure 6.6: Evolución de los rendimientos del Bitcoin en valor absoluto

```
acf(abs(logret))
```

Otra manera de corroborar esta idea es volviendo IID nuestra serie de datos y observar que de este modo se pierde la autocorrelación serial, lo que refuerza la idea de que en esta serie existen clusters de volatilidad.

```
logret_random <- sample(as.vector(logret), size = length(logret), replace = FALSE)
acf(abs(logret_random))
```

```
par(mfrow = c(1,2))
plot(logret)
plot(logret_random, type = 'l')
```

Ahora busquemos una prueba formal. El primer enfoque para comprobar aceptar o rechazar la hipótesis de que necesitamos estimar un ARCH(q) es realizar una prueba de efectos ARCH. De acuerdo con esta, nuestros datos de Bitcoin muestran que se rechaza la hipótesis de no efectos ARCH. Esta prueba resulta del siguiente código de R:

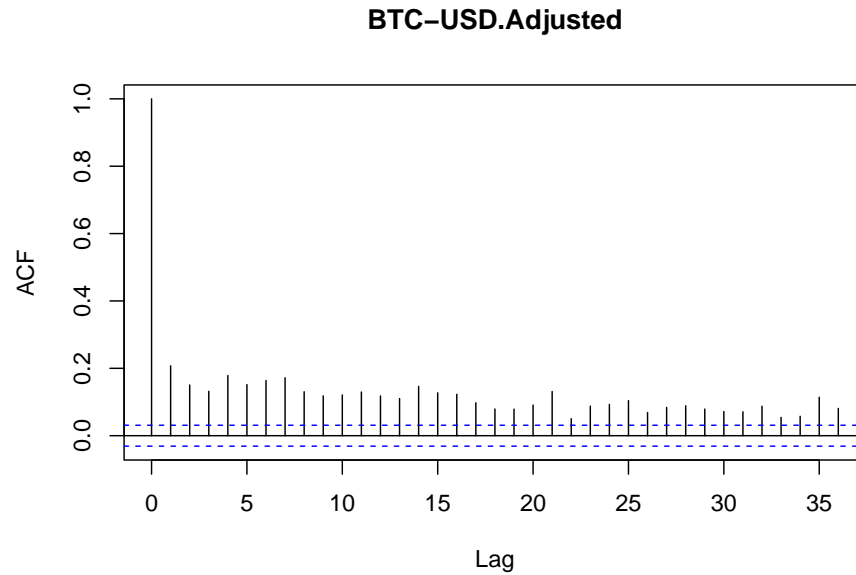


Figure 6.7: Función de autocorrelación parcial de los rendimientos del Bitcoin en valor absoluto

```
logret_mean = dynlm(logret~1)
summary(logret_mean)
```

```
##
## Time series regression with "ts" data:
## Start = 1, End = 3991
##
## Call:
## dynlm(formula = logret ~ 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.46611 -0.01386 -0.00012  0.01483  0.22374
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0013824  0.0005667   2.439  0.0148 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

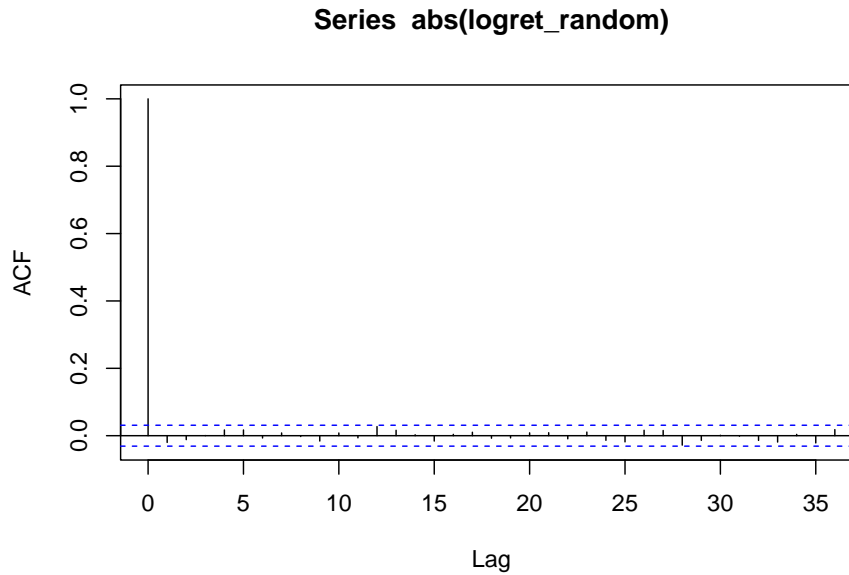


Figure 6.8: Función de autocorrelación parcial de los rendimientos del Bitcoin en valor absoluto

```
##
## Residual standard error: 0.0358 on 3990 degrees of freedom
```

```
ehatsq = ts(resid(logret_mean)^2)
ARCH_m = dynlm(ehatsq~L(ehatsq))
summary(ARCH_m)
```

```
##
## Time series regression with "ts" data:
## Start = 2, End = 3991
##
## Call:
## dynlm(formula = ehatsq ~ L(ehatsq))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.017239 -0.001113 -0.000985 -0.000307  0.216149
##
## Coefficients:
```

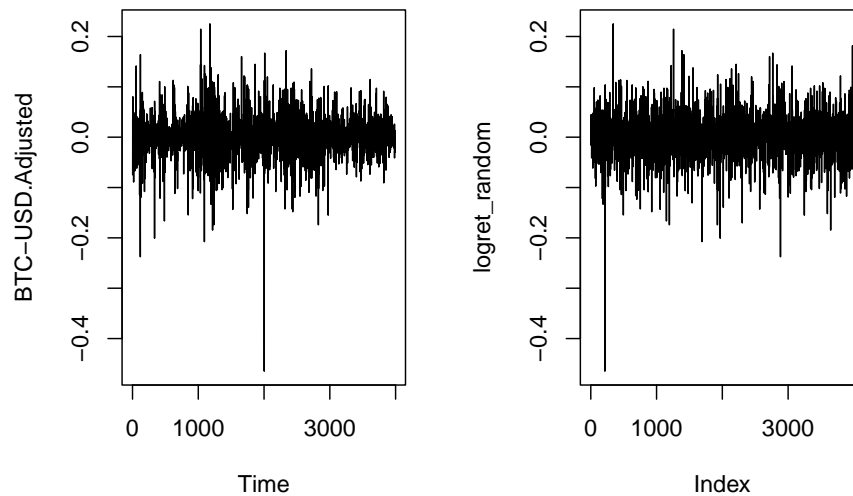
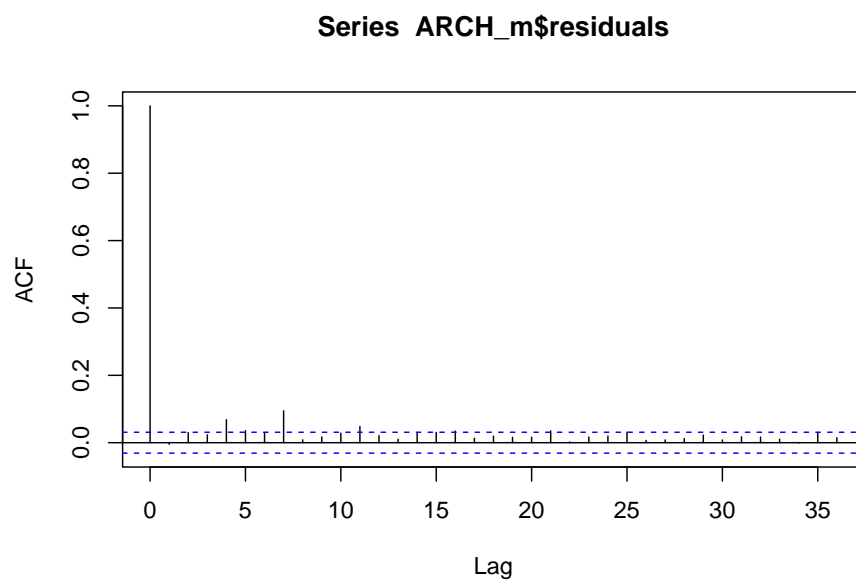


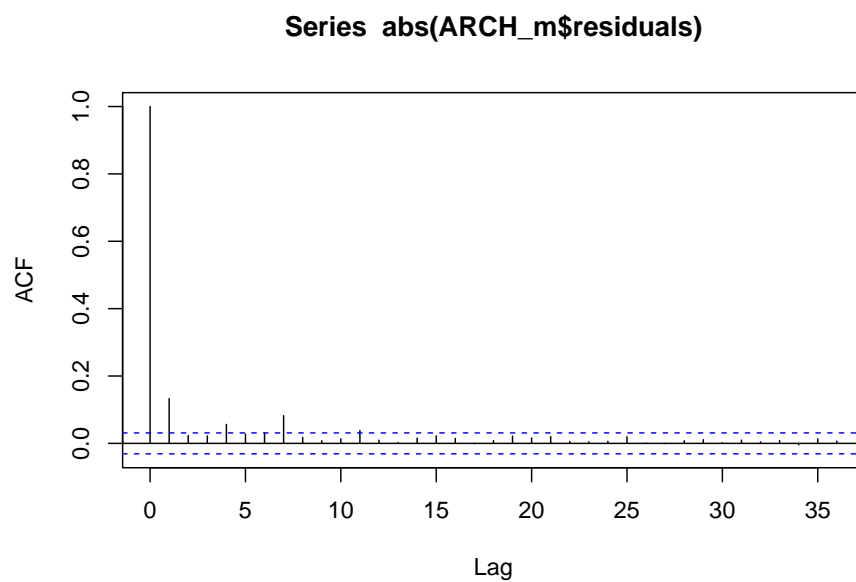
Figure 6.9: Función de autocorrelación parcial de los rendimientos del Bitcoin en valor absoluto

```
##               Estimate Std. Error t value      Pr(>|t|)
## (Intercept)  0.00111209 0.00007689  14.463 <0.0000000000000002 ***
## L(ehatsq)    0.13124874 0.01569650   8.362 <0.0000000000000002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.004688 on 3988 degrees of freedom
## Multiple R-squared:  0.01723,    Adjusted R-squared:  0.01698
## F-statistic: 69.92 on 1 and 3988 DF,  p-value: < 0.00000000000000022
```

```
acf(ARCH_m$residuals)
```



```
acf(abs(ARCH_m$residuals))
```



```
ArchTest(logret, lags = 1, demean = TRUE)
```

```
##
## ARCH LM-test; Null hypothesis: no ARCH effects
##
## data: logret
## Chi-squared = 68.747, df = 1, p-value < 0.00000000000000022
```

La prueba puede repetirse para diferentes especificaciones de ARCH, sin embargo, para efectos ilustrativos usaremos un ARCH(1). Estimemos un ARCH(1), considerando la siguiente especificación:

$$\begin{aligned} Y_t &= \mu + \sqrt{h_t} \varepsilon_t \\ h_t &= \omega + \alpha_1 h_{t-1} + u_t \\ \varepsilon &\sim N(0, 1) \end{aligned}$$

Considerando que la media es una AR(2), los resultados de esta estimación son los siguientes:

Table 6.1: Resumen de la estimación ARIMA(2,0,0) con media distinta de cero (y posible componente GARCH) $\sigma^2 = 0.001324$, $\log\text{-likelihood} = 7170.99$, $AIC = -14333.99$, $AICc = -14333.97$, $BIC = -14309.03$

Parameter	Estimate	Std. Error	t value	Pr(> t)
mu	0.001642394	0.0003553101	4.622424	3.792818e-06
ar1	-0.051859993	0.0160262284	-3.235945	1.212408e-03
ar2	-0.001225748	0.0132050756	-0.092824	9.260434e-01
omega	0.001956396	0.0004839458	4.042593	5.286335e-05
alpha1	0.730450078	0.2140120509	3.413126	6.422226e-04
shape	2.347586356	0.1135148037	20.680883	0.000000e+00

Notas:

1. **Primera línea en negrita** como título de la tabla (Markdown no tiene un `\caption{}` como LaTeX).
2. Bajo el título, se añade un párrafo en *itálicas* con los valores adicionales (σ^2 , log-likelihood, AIC, etc.).

3. La fila de encabezados va seguida por una fila de separadores (|:---|---:| etc.) para indicar la **alineación** a la izquierda, derecha o centro.
4. Las celdas con datos numéricos se alinean a la **derecha** (:---: o ---:) para mantener la estética de los números.
5. Si prefieres, puedes omitir la columna de “**Notas**” y añadir la información directamente en el título de la tabla o en un párrafo posterior.

```
library(rugarch)
```

```
auto.arima(logret)
```

```
## Series: logret
## ARIMA(2,0,0) with non-zero mean
##
## Coefficients:
##          ar1      ar2      mean
##      -0.0228  0.0106  0.0014
## s.e.    0.0158  0.0158  0.0006
##
## sigma^2 = 0.001282: log likelihood = 7627.86
## AIC=-15247.72   AICc=-15247.71   BIC=-15222.56
```

```
##?ugarchspec
```

```
model.spec = ugarchspec( variance.model = list(model = 'sGARCH' ,
                                                garchOrder = c(1, 0)),
                        mean.model = list(armaOrder = c(2,0)),
                        distribution.model = "std")
```

```
##?ugarchfit
```

```
arch.fit = ugarchfit(spec = model.spec , data = logret,
                    solver = 'solnp')
```

```
arch.fit@fit$matcoef
```

##	Estimate	Std. Error	t value	Pr(> t)
## mu	0.001533153	0.0003372857	4.5455629	0.000005478863
## ar1	-0.053488931	0.0156702298	-3.4134108	0.000641551571
## ar2	-0.003574044	0.0128610322	-0.2778971	0.781091318092
## omega	0.001899027	0.0004676134	4.0611051	0.000048840972

```
## alpha1 0.768548082 0.2218148450 3.4648181 0.000530590302
## shape 2.339763015 0.1097843607 21.3123527 0.000000000000
```

```
boot.garch <- ugarchboot(arch.fit,
                        method = "Partial",
                        sampling = "raw", #bootstrap from fitted varepsilon
                        n.ahead = 1,      #simulation horizon
                        n.bootpred = 100000, #number of simulations
                        solver = "solnp")
```

```
boot.garch
```

```
##
## *-----*
## *      GARCH Bootstrap Forecast      *
## *-----*
## Model : sGARCH
## n.ahead : 1
## Bootstrap method: partial
## Date (T[0]): 3991-01-01
##
## Series (summary):
##      min      q.25      mean      q.75      max forecast[analytic]
## t+1 -0.46649 -0.010675 0.001515 0.014601 0.20108                0.001775
## .....
##
## Sigma (summary):
##      min      q0.25      mean      q0.75      max forecast[analytic]
## t+1 0.043779 0.043779 0.043779 0.043779 0.043779                0.043779
## .....
```

6.4 Ejemplo GARCH(0,1)

Estimemos un GARCH(0,1) usando la siguiente especificación:

$$\begin{aligned} Y_t &= \mu + \sqrt{h_t} \varepsilon_t \\ h_t &= \omega + \beta_i \sigma_{t-i}^2 + u_t \\ \varepsilon &\sim N(0,1) \end{aligned}$$

Considerando que la media es una AR(2), los resultados de esta estimación son los siguientes:

Table 6.2: Resumen de estimaciones del modelo GARCH(0,1)

Parameter	Estimate	Std. Error	t value	Pr(> t)
mu	1.743510e-03	3.728562e-04	4.6760931	2.923919e-06
ar1	-6.104183e-02	1.366530e-02	-4.4669219	7.935308e-06
ar2	-4.231427e-03	1.311300e-02	-0.3226895	7.469304e-01
omega	3.843436e-06	6.414905e-08	59.9141495	0.000000e+00
beta1	9.981016e-01	3.215075e-05	31044.4303	0.000000e+00
shape	2.515980e+00	2.796609e-02	89.9653552	0.000000e+00

```

model.spec = ugarchspec(variance.model = list(model = 'sGARCH' ,
                                              garchOrder = c(0,1)),
                        mean.model = list(armaOrder = c(2,0)),
                        distribution.model = "std")

fit.garch.n = ugarchfit(spec = model.spec, data = logret,
                        solver = "solnp")

fit.garch.n@fit$matcoef

```

```

##           Estimate      Std. Error      t value      Pr(>|t|)
## mu      0.001633928242 0.00035471634895      4.6062953 0.000004099062
## ar1     -0.062287298275 0.01330160588629     -4.6826901 0.000002831344
## ar2     -0.004627082887 0.01279336245566     -0.3616784 0.717592368774
## omega    0.000004364163 0.00000006731874     64.8283429 0.000000000000
## beta1    0.997708584743 0.00004207935059    23710.1706839 0.000000000000
## shape    2.523521563642 0.02710634646569     93.0970748 0.000000000000

```

```

boot.garch <- ugarchboot(fit.garch.n,
                        method = "Partial",
                        sampling = "raw", #bootstrap from fitted varepsilon
                        n.ahead = 1,      #simulation horizon
                        n.bootpred = 100000, #number of simulations
                        solver = "solnp")

boot.garch

```

```

##
## *-----*
## *      GARCH Bootstrap Forecast      *
## *-----*
## Model : sGARCH

```

```

## n.ahead : 1
## Bootstrap method: partial
## Date (T[0]): 3991-01-01
##
## Series (summary):
##      min      q.25    mean      q.75    max forecast[analytic]
## t+1 -0.46532 -0.012576 0.0014 0.016353 0.23934          0.001917
## .....
##
## Sigma (summary):
##      min      q0.25    mean      q0.75    max forecast[analytic]
## t+1 0.043641 0.043641 0.043641 0.043641 0.043641          0.043641
## .....

```

6.4.1 Selección GARCH(p,q) óptimo

¿Cómo seleccionamos el orden adecuado para un GARCH(p,q)? Aquí una respuesta.

$$Y_t = \mu + \sqrt{h_t} \varepsilon_t$$

$$h_t = \omega + \beta_i h_{t-i} + \alpha_i \varepsilon_{t-i}^2 + u_t$$

$$\varepsilon \sim N(0, 1)$$

6.4.1.1 Criterios de información

Table 6.3: Comparación de criterios de información.

Criterio	Valor
Akaike	-4.083927
Bayes	-4.074035
Shibata	-4.083932
Hannan-Quinn	-4.080410

```
infocriteria(fit.garch.n)
```

```

##
## Akaike      -4.120182
## Bayes      -4.110723
## Shibata    -4.120186
## Hannan-Quinn -4.116828

```

6.4.1.2 Selección del modelo óptimo

Podemos hacer una búsqueda del mejor modelo de entre varios que probemos en un espectro de hasta un GARCH(4,4). Los resultados son los reportados en el siguiente Cuadro.

Table 6.4: Resumen de criterios de información en función de p y q de un GARCH(p,q).

q	p	AIC	Óptimo
1	1	-10.93647	0
1	2	-10.93657	0
1	3	-10.93799	0
1	4	-10.93727	0
2	1	-10.93579	0
2	2	-10.93607	0
2	3	-10.93743	0
2	4	-10.93510	0
3	1	-10.93524	0
3	2	-10.93510	0
3	3	-10.93765	0
3	4	-10.93645	0
4	1	-10.93840	0
4	2	-10.93935	1
4	3	-10.93924	0
4	4	-10.93816	0

```
source("Lag_Opt_GARCH.R")
```

```
Lag_Opt_GARCH(ehatsq,4,4)
```

```
##      q p      AIC Optimo
## [1,] 1 1 -11.01907      0
## [2,] 1 2 -11.01941      0
## [3,] 1 3 -11.01907      0
## [4,] 1 4 -11.02019      0
## [5,] 2 1 -11.01857      0
## [6,] 2 2 -11.01893      0
## [7,] 2 3 -11.01868      0
## [8,] 2 4 -11.02078      1
## [9,] 3 1 -11.01777      0
## [10,] 3 2 -11.01800      0
## [11,] 3 3 -11.01793      0
## [12,] 3 4 -11.02028      0
```

```
## [13,] 4 1 -11.01991      0
## [14,] 4 2 -11.02071      0
## [15,] 4 3 -11.01885      0
## [16,] 4 4 -11.02020      0
```

6.4.1.3 Estimación de modelo óptimo

De esta forma, el modelo óptimo es un GARCH(2,4). Los resultados de la estimación son los del Cuadro

Table 6.5: Resumen de estimaciones del modelo (parámetros, errores estándar, estadísticos t y p-values).

Parameter	Estimate	Std. Error	t value	Pr(> t)
mu	1.353612e-03	3.233488e-04	4.186229e+00	2.836277e-05
ar1	-4.916331e-02	1.495855e-02	-	1.013919e-03
			3.286636e+00	
ar2	7.273432e-04	1.385161e-02	5.250964e-02	9.581226e-01
omega	2.759072e-05	5.151381e-06	5.355984e+00	8.509188e-08
alpha1	1.290996e-01	2.165606e-02	5.961363e+00	2.501419e-09
alpha2	1.768468e-07	3.676320e-02	4.810430e-06	9.999962e-01
alpha3	6.424436e-08	2.663630e-02	2.411910e-06	9.999981e-01
alpha4	3.333343e-02	2.238762e-02	1.488923e+00	1.365077e-01
beta1	4.260058e-01	2.263614e-01	1.881972e+00	5.983983e-02
beta2	4.105609e-01	1.819274e-01	2.256730e+00	2.402498e-02
shape	3.183045e+00	1.305307e-01	2.438541e+01	0.000000e+00

```
model.spec = ugarchspec(variance.model = list(model = 'sGARCH',
                                              garchOrder = c(4,2)),
                        mean.model = list(armaOrder = c(2,0)),
                        distribution.model = "std")

model.fit = ugarchfit(spec = model.spec , data = logret,
                      solver = 'solnp')

model.fit@fit$matcoef
```

```
##              Estimate      Std. Error      t value      Pr(>|t|)
## mu      0.00127431846326 0.000310016025  4.110492238865 0.0000394816551619
## ar1     -0.05043940906507 0.014578268430 -3.459903987027 0.0005403680260374
## ar2     -0.00118565480491 0.013554426468 -0.087473624035 0.9302950500228699
## omega    0.00002712017695 0.000004539222  5.974631686506 0.0000000023061060
```

```
## alpha1 0.13002379592178 0.021263901527 6.114766650699 0.0000000009669836
## alpha2 0.00000034065677 0.040085714409 0.000008498209 0.9999932194104486
## alpha3 0.00000008107872 0.024881041011 0.000003258655 0.9999973999697527
## alpha4 0.02872099361214 0.022852869499 1.256778437118 0.2088338761998922
## beta1 0.43272444998772 0.235611728606 1.836599784521 0.0662689991645959
## beta2 0.40753030009967 0.193897651433 2.101780486187 0.0355725092842449
## shape 3.17568857589049 0.130728859377 24.292176884399 0.0000000000000000
```

```
boot.garch <- ugarchboot(model.fit,
  method = "Partial",
  sampling = "raw", #bootstrap from fitted varepsilon
  n.ahead = 1,      #simulation horizon
  n.bootpred = 100000, #number of simulations
  solver = "solnp")
```

```
boot.garch
```

```
##
## *-----*
## *      GARCH Bootstrap Forecast      *
## *-----*
## Model : sGARCH
## n.ahead : 1
## Bootstrap method: partial
## Date (T[0]): 3991-01-01
##
## Series (summary):
##      min      q.25      mean      q.75      max forecast[analytic]
## t+1 -0.25981 -0.007714 0.001473 0.011172 0.19728                0.001514
## .....
##
## Sigma (summary):
##      min      q0.25      mean      q0.75      max forecast[analytic]
## t+1 0.021108 0.021108 0.021108 0.021108 0.021108                0.021108
## .....
```

6.4.1.4 Forecasting with GARCH(1,1)

Para realizar pronósticos con la estimación de un GARCH, utilizando la librería *rugarch*, es necesario utilizar la función `ugarchforecast()`.

```
model.spec = ugarchspec(variance.model = list(model = 'sGARCH',
  garchOrder = c(1,1)),
  mean.model = list(armaOrder = c(4,2)),
```

```

distribution.model = "std")

model.fit = ugarchfit(spec = model.spec , data = logret,
                     solver = 'solnp')

spec = getspec(model.fit)

setfixed(spec) <- as.list(coef(model.fit))

```

Esta función precisa como argumentos nuestra estimación del modelo GARCH, con una modificación en la manera en que se presentan los coeficientes, realizada en la última línea del código anterior y que llamamos *spec*. *n.ahead* es el número de periodos que vamos a pronosticar, *n.roll* señala el número de pronósticos móviles que utilizaremos, en caso de que haya más información para realizar el pronóstico. Finalmente damos como input nuestro set de datos y como producto obtendremos el pronóstico de Sigma tanto como de la serie.

```

forecast = ugarchforecast(spec, n.ahead = 12, n.roll = 0, logret)

sigma(forecast)

```

```

##      3991-01-01
## T+1  0.02121260
## T+2  0.02161771
## T+3  0.02201497
## T+4  0.02240480
## T+5  0.02278758
## T+6  0.02316366
## T+7  0.02353337
## T+8  0.02389700
## T+9  0.02425482
## T+10 0.02460709
## T+11 0.02495404
## T+12 0.02529590

```

```

fitted(forecast)

```

```

##      3991-01-01
## T+1  0.001465389
## T+2  0.001605238
## T+3  0.001116383
## T+4  0.001013733
## T+5  0.001474082
## T+6  0.001461704

```

```
## T+7  0.001025356
## T+8  0.001143766
## T+9  0.001531446
## T+10 0.001323668
## T+11 0.001004290
## T+12 0.001280706
```

```
forecast
```

```
##
## *-----*
## *          GARCH Model Forecast          *
## *-----*
## Model: sGARCH
## Horizon: 12
## Roll Steps: 0
## Out of Sample: 0
##
## 0-roll forecast [T0=]:
##      Series  Sigma
## T+1  0.001465 0.02121
## T+2  0.001605 0.02162
## T+3  0.001116 0.02201
## T+4  0.001014 0.02240
## T+5  0.001474 0.02279
## T+6  0.001462 0.02316
## T+7  0.001025 0.02353
## T+8  0.001144 0.02390
## T+9  0.001531 0.02425
## T+10 0.001324 0.02461
## T+11 0.001004 0.02495
## T+12 0.001281 0.02530
```

6.5 Modelos ARCH y GARCH Multivariados

De forma similar a los modelos univariados, los modelos multivariados de heterocedasticidad condicional asumen una estructura de la media condicional. En este caso, descrita por un VAR(p) cuyo proceso estocástico \mathbf{X} es estacionario de dimensión k . De esta forma, la expresión reducida del modelo o el proceso VAR(p) estará dado por:

$$\mathbf{X}_t = \delta + \mathbf{A}_1 \mathbf{X}_{t-1} + \mathbf{A}_2 \mathbf{X}_{t-2} + \dots + \mathbf{A}_p \mathbf{X}_{t-p} + \mathbf{U}_t \quad (6.9)$$

Donde cada uno de las \mathbf{A}_i , $i = 1, 2, \dots, p$, son matrices cuadradas de dimensión k y \mathbf{U}_t representa un vector de dimensión $k \times 1$ con los residuales en el momento

del tiempo t que son un proceso puramente aleatorio. También se incorpora un vector de términos constantes denominado como δ , el cual es de dimensión $k \times 1$ —en este caso también es posible incorporar procesos determinísticos adicionales—.

Así, suponemos que el término de error tendrá estructura de vector:

$$\mathbf{U}_t = \begin{bmatrix} U_{1t} \\ U_{2t} \\ \vdots \\ U_{Kt} \end{bmatrix}$$

De forma que diremos que:

$$\mathbf{U}_t | \Omega_{t-1} \sim (0, \Sigma_t |_{t-1})$$

Dicho lo anterior, entonces, el modelo ARCH(q) multivariado será descrito por:

$$Vech(\Sigma_t |_{t-1}) = \gamma_0 + \Gamma_1 Vech(\mathbf{U}_{t-1} \mathbf{U}_{t-1}') + \dots + \Gamma_q Vech(\mathbf{U}_{t-q} \mathbf{U}_{t-q}') \quad (\#eq : M_{ARCH}) \quad (6.10)$$

Donde $Vech$ es un operador que apila en un vector la parte superior de la matriz a la cual se le aplique, γ_0 es un vector de constantes, Γ_i , $i = 1, 2, \dots$ son matrices de coeficientes asociados a la estimación.

Para ilustrar la ecuación @ref(eq:M_ARCH), tomemos un ejemplo de $K = 2$, de esta forma tenemos que un M-ARCH(1) será:

$$\Sigma_t |_{t-1} = \begin{bmatrix} \sigma_{1,t|t-1}^2 & \sigma_{12,t|t-1} \\ \sigma_{21,t|t-1} & \sigma_{2,t|t-1}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \Sigma_t$$

Donde hemos simplificado la notación de las varianzas y la condición de que están en función de $t - 1$. Así,

$$Vech(\Sigma_t) = Vech \begin{bmatrix} \sigma_{11,t} & \sigma_{12,t} \\ \sigma_{21,t} & \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \sigma_{11,t} \\ \sigma_{12,t} \\ \sigma_{22,t} \end{bmatrix}$$

De esta forma, podemos establecer el modelo M-ARCH(1) con $K = 2$ será de la forma:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{12,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \gamma_{30} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \begin{bmatrix} U_{1,t-1}^2 \\ U_{1,t-1} U_{2,t-1} \\ U_{2,t-1}^2 \end{bmatrix}$$

Como notarán, este tipo de procedimientos implica la estimación de muchos parámetros. En esta circunstancia, se suelen estimar modelos restringidos para

reducir el número de coeficientes estimados. Por ejemplo, podríamos querer estimar un caso como:

$$\begin{bmatrix} \sigma_{11,t} \\ \sigma_{12,t} \\ \sigma_{22,t} \end{bmatrix} = \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \gamma_{30} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & 0 & 0 \\ 0 & \gamma_{22} & 0 \\ 0 & 0 & \gamma_{33} \end{bmatrix} \begin{bmatrix} U_{1,t-1}^2 \\ U_{1,t-1}U_{2,t-1} \\ U_{2,t-1}^2 \end{bmatrix}$$

Finalmente y de forma análoga al caso univariado, podemos plantear un modelo M-GARCH(p, q) como:

$$Vech(\Sigma_{t|t-1}) = \gamma_0 + \sum_{j=1}^q \Gamma_j Vech(\mathbf{U}_{t-j} \mathbf{U}_{t-j}') + \sum_{m=1}^p \mathbf{G}_m Vech(\Sigma_{t-m|t-m-1}) \quad (6.11)$$

Donde cada una de las \mathbf{G}_m es una matriz de coeficientes. Para ilustrar este caso, retomemos el ejemplo anterior, pero ahora para un modelo M-GARCH(1, 1) con $K = 2$ de forma que tendríamos:

$$\begin{aligned} \begin{bmatrix} \sigma_{11,t} \\ \sigma_{12,t} \\ \sigma_{22,t} \end{bmatrix} &= \begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \gamma_{30} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} & \gamma_{13} \\ \gamma_{21} & \gamma_{22} & \gamma_{23} \\ \gamma_{31} & \gamma_{32} & \gamma_{33} \end{bmatrix} \begin{bmatrix} U_{1,t-1}^2 \\ U_{1,t-1}U_{2,t-1} \\ U_{2,t-1}^2 \end{bmatrix} \\ &+ \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{bmatrix} \begin{bmatrix} \sigma_{11,t-1} \\ \sigma_{12,t-1} \\ \sigma_{22,t-1} \end{bmatrix} \end{aligned}$$

6.6 Pruebas para detectar efectos ARCH

La prueba que mostraremos es conocida como una ARCH-LM, la cual está basada en una regresión de los residuales estimados de un modelo VAR(p) o cualquier otra estimación que deseemos probar, con el objeto de determinar si existen efectos ARCH –esta prueba se puede simplificar para el caso univariado–.

Partamos de plantear:

$$\begin{aligned} Vech(\hat{\mathbf{U}}_t \hat{\mathbf{U}}_t') &= \mathbf{B}_0 + \mathbf{B}_1 Vech(\hat{\mathbf{U}}_{t-1} \hat{\mathbf{U}}_{t-1}') + \dots \\ &+ \mathbf{B}_q Vech(\hat{\mathbf{U}}_{t-q} \hat{\mathbf{U}}_{t-q}') + \varepsilon_t \end{aligned} \quad (6.12)$$

Dada la estimación en la ecuación (6.13), plantearemos la estructura de hipótesis dada por:

$$\begin{aligned} H_0 &: \mathbf{B}_1 = \mathbf{B}_2 = \dots = \mathbf{B}_q = 0 \\ H_a &: \text{No } H_0 \end{aligned}$$

La estadística de prueba será determinada por:

$$LM_{M-ARCH} = \frac{1}{2} TK(K+1) - \text{Traza}(\hat{\Sigma}_{ARCH} \hat{\Sigma}_0^{-1}) \sim \chi_{[qK^2(K+1)^2/4]}^2 \quad (6.13)$$

Donde la matriz $\hat{\Sigma}_{ARCH}$ se calcula de acuerdo con la ecuación (6.13) y la matriz $\hat{\Sigma}_0$ sin considerar una estructura dada para los errores.

Chapter 7

Modelos de Datos Panel, Panel VAR y otros Modelos No Lineales

7.1 Modelos de Datos Panel y Panel VAR

7.1.1 Motivación

Los datos panel son una extensión natural del análisis de series de tiempo en el que observamos una misma variable a lo largo del tiempo, pero para múltiples individuos. Denotemos con Y_{it} el componente o individuo i , $i = 1, 2, \dots, N$ en el tiempo t , $t = 1, 2, \dots, T$. Típicamente, las series de tiempo en forma de panel se caracterizan por ser de dimensión de individuos corta y de dimensión temporal larga.

En general, escribiremos:

$$Y_{it} = \alpha_i + \beta_i X_{it} + U_{it} \quad (7.1)$$

Donde α_i es un efecto fijo especificado como:

$$\alpha_i = \beta_0 + \gamma_i Z_i \quad (7.2)$$

7.1.2 Pruebas de Raíces Unitarias en Panel

Las pruebas de raíces unitarias para panel suelen ser usadas principalmente en casos macroeconómicos. De forma similar al caso de series univariadas, asumire-

mos una forma de AR(1) para una serie en panel:

$$\Delta Y_{it} = \mu_i + \rho_i Y_{it-1} + \sum_{j=1}^{k_i} \varphi_{ij} \Delta Y_{it-j} + \varepsilon_{it} (\#eq : eq_A R_{panel}) \quad (7.3)$$

Donde $i = 1, \dots, N$, $t = 1, \dots, T$ y ε_{it} es una v.a. iid que cumple con:

$$\begin{aligned} \mathbb{E}[\varepsilon_{it}] &= 0 \\ \mathbb{E}[\varepsilon_{it}^2] &= \sigma_i^2 < \infty \\ \mathbb{E}[\varepsilon_{it}^4] &< \infty \end{aligned}$$

Al igual que en el caso univariado, en este tipo de pruebas buscamos identificar cuándo las series son I(1) y cuándo I(0). Para tal efecto, la prueba de raíz unitaria que utilizaremos está basada en una prueba Dickey-Fuller Aumentada en la cual la hipótesis nula (H_0) es que todas las series en el panel son no estacionarias, es decir, son I(1). Es decir,

$$H_0 : \rho_i = 0 \quad (7.4)$$

Por su parte, en el caso de la hipótesis nula tendremos dos:

- H_1^A : Todas las series son I(0) – caso homogéneo, o
- H_1^B : Al menos una de las series es I(0) – caso heterogéneo

7.1.3 Ejemplo: Pruebas de Raíces Unitarias en Panel

Dependencies and Setup

Download data from library PLM

```
data("EmplUK", package="plm")

data("Produc", package="plm")

data("Grunfeld", package="plm")

data("Wages", package="plm")
```

Describe data: Grunfeld data (Grunfeld 1958) comprising 20 annual observations on the three variables real gross investment (invest), real value of the firm (value), and real value of the capital stock (capital) for 10 large US firms for the years 1935–1954

```
names(Grunfeld)
```

```
## [1] "firm"    "year"    "inv"     "value"   "capital"
```

```
head(Grunfeld)
```

```
##   firm year  inv  value capital
## 1    1 1935 317.6 3078.5      2.8
## 2    1 1936 391.8 4661.7     52.6
## 3    1 1937 410.6 5387.1    156.9
## 4    1 1938 257.7 2792.2    209.2
## 5    1 1939 330.8 4313.2    203.4
## 6    1 1940 461.2 4643.9    207.2
```

```
Invest <- data.frame(split( Grunfeld$inv, Grunfeld$firm )) # individuals in columns
```

```
names(Invest)
```

```
## [1] "X1" "X2" "X3" "X4" "X5" "X6" "X7" "X8" "X9" "X10"
```

```
names(Invest) <- c( "Firm_1", "Firm_2", "Firm_3", "Firm_4",
                    "Firm_5", "Firm_6", "Firm_7", "Firm_8",
                    "Firm_9", "Firm_10" )
```

```
names(Invest)
```

```
## [1] "Firm_1" "Firm_2" "Firm_3" "Firm_4" "Firm_5" "Firm_6" "Firm_7"
## [8] "Firm_8" "Firm_9" "Firm_10"
```

Plot:

```
plot(Invest$Firm_1, type = "l", col = 1, ylim = c(0, 1500), lty = 1,
      xlab = "Tiempo", ylab = "Real gross investment")
lines(Invest$Firm_2, type = "l", col = 2, lty = 2)
lines(Invest$Firm_3, type = "l", col = 3, lty = 1)
lines(Invest$Firm_4, type = "l", col = 4, lty = 2)
lines(Invest$Firm_5, type = "l", col = 5, lty = 1)
lines(Invest$Firm_6, type = "l", col = 6, lty = 2)
lines(Invest$Firm_7, type = "l", col = 7, lty = 1)
lines(Invest$Firm_8, type = "l", col = 8, lty = 2)
lines(Invest$Firm_9, type = "l", col = 9, lty = 1)
lines(Invest$Firm_10, type = "l", col = 10, lty = 2)
legend("topleft", legend=c("Firm_1", "Firm_2", "Firm_3", "Firm_4", "Firm_5",
                           "Firm_6", "Firm_7", "Firm_8", "Firm_9", "Firm_10"),
      col = c(1, 2, 3, 4, 5, 6, 7, 8, 9, 10), lty = 1:2)
```

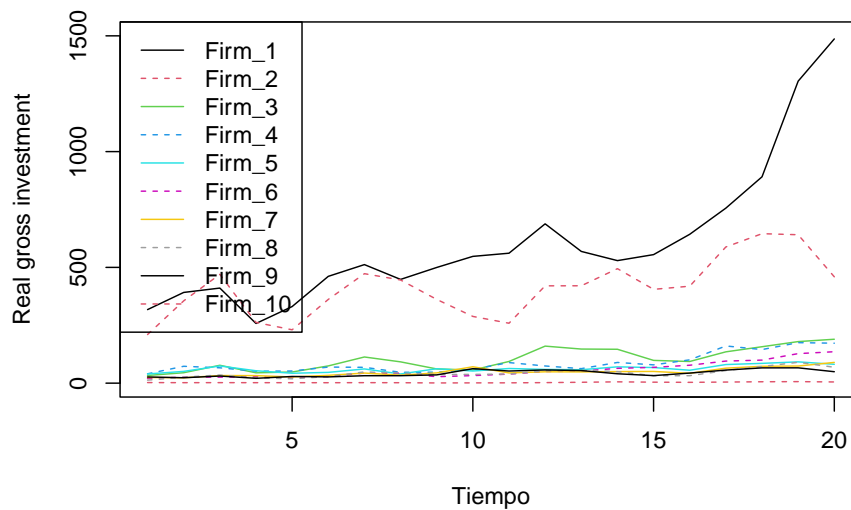


Figure 7.1: Evolución de la inversión por empresa

Unit Root Test: Test specifies the type of test to be performed among Levin et al. (2002), Im et al. (2003), Maddala and Wu (1999) and Hadri (2000)

Consider Levin et al. (2002)

```
purtest(log(Invest), test = "levinlin", exo = "intercept",
        lags = "AIC", pmax = 4)
```

```
## Warning in selectT(1, theTs): the time series is short
```

```
##
## Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)
##
## data: log(Invest)
## z = -0.63214, p-value = 0.2636
## alternative hypothesis: stationarity
```

Same via:

```
ts_LInvest <- ts(log(Invest), start = 1935, end = 1954, freq = 1)
```

```
ts_DLInvest <- diff(ts(log(Invest), start = 1935, end = 1954, freq = 1),
                    lag = 1, differences = 1)
```

```
purtest(ts_LInvest, test = "levinlin", exo = "intercept",
        lags = "AIC", pmax = 4)
```

```
## Warning in selectT(l, theTs): the time series is short
```

```
##
## Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)
##
## data: ts_LInvest
## z = -0.63214, p-value = 0.2636
## alternative hypothesis: stationarity
```

```
summary(purtest(ts_LInvest, test = "levinlin", exo = "intercept",
                lags = "AIC", pmax = 4))
```

```
## Warning in selectT(l, theTs): the time series is short
```

```
## Levin-Lin-Chu Unit-Root Test
## Exogenous variables: Individual Intercepts
## Automatic selection of lags using AIC: 0 - 3 lags (max: 4)
## statistic: -0.632
## p-value: 0.264
##
##      lags obs      rho      trho      p.trho      sigma2ST      sigma2LT
## Firm_1    0  19  0.006034802  0.05068864  0.96193110  0.03592530  0.014769876
## Firm_2    1  18 -0.597193025 -2.95851997  0.03895745  0.05020448  0.014565004
## Firm_3    3  16 -0.257291999 -1.97135237  0.29979349  0.03073205  0.019141095
## Firm_4    0  19 -0.172070863 -1.21388533  0.67084314  0.06196079  0.018454390
## Firm_5    1  18 -0.444726602 -1.69693864  0.43286539  0.04165516  0.006949122
## Firm_6    2  17  0.062922587  0.72626225  0.99275186  0.02442139  0.009212851
## Firm_7    2  17 -0.013830411 -0.08442715  0.94937802  0.03230724  0.007274365
## Firm_8    1  18 -0.365024064 -2.26588313  0.18328720  0.06270208  0.026960667
## Firm_9    0  19 -0.247940802 -1.72877417  0.41668489  0.05073398  0.013291332
## Firm_10   1  18 -0.197003144 -1.52323519  0.52199891  0.07672404  0.062496477
```

```
purtest(ts_DLInvest, test = "levinlin", exo = "intercept",
        lags = "AIC", pmax = 4)
```

```
## Warning in selectT(l, theTs): the time series is short
```

```
##
## Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts)
##
## data: ts_DLInvest
## z = -8.5076, p-value < 0.00000000000000022
## alternative hypothesis: stationarity
```

```
summary(purtest(ts_DLInvest, test = "levinlin", exo = "intercept",
               lags = "AIC", pmax = 4))
```

```
## Warning in selectT(1, theTs): the time series is short
```

```
## Levin-Lin-Chu Unit-Root Test
## Exogenous variables: Individual Intercepts
## Automatic selection of lags using AIC: 0 - 3 lags (max: 4)
## statistic: -8.508
## p-value: 0
##
##      lags obs      rho      trho      p.trho      sigma2ST      sigma2LT
## Firm_1    0  18 -0.9778879 -4.196070 0.000665206460867 0.03693616 0.007853674
## Firm_2    3  15 -2.7288380 -4.327510 0.000392850467043 0.02832436 0.031954702
## Firm_3    2  16 -1.7848678 -5.650917 0.000000790572465 0.03819653 0.022971185
## Firm_4    1  17 -1.5737470 -5.320458 0.000004309793729 0.04231849 0.030313940
## Firm_5    2  16 -2.8118586 -5.181259 0.000008564306919 0.02737020 0.015226407
## Firm_6    1  17 -1.4850311 -4.776618 0.000057171776145 0.02517911 0.008588652
## Firm_7    1  17 -2.1076263 -6.541451 0.000000005364403 0.03232079 0.012780800
## Firm_8    2  16 -1.7790739 -4.976338 0.000022818242011 0.03277746 0.034905699
## Firm_9    0  18 -1.2600946 -5.304469 0.000004667479100 0.05666016 0.019502840
## Firm_10   0  18 -0.7889550 -3.421651 0.010275559745253 0.08661396 0.030216984
```

Consider Im-Pesaran-Shin Unit-Root Test (2003)

```
purtest(ts_LInvest, test = "ips", exo = "intercept",
       lags = "AIC", pmax = 4)
```

```
##
## Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)
##
## data: ts_LInvest
## Wtbar = 0.6833, p-value = 0.7528
## alternative hypothesis: stationarity
```



```
summary(purtest(ts_LInvest, test = "ips", exo = "intercept",
               lags = "AIC", pmax = 4))
```

```
## Im-Pesaran-Shin Unit-Root Test
## Exogenous variables: Individual Intercepts
## Automatic selection of lags using AIC: 0 - 3 lags (max: 4)
## statistic (Wtbar): 0.683
## p-value: 0.753
##
##      lags obs      rho      trho      p.trho      mean      var
## Firm_1    0  19  0.006034802  0.05068864  0.96193110 -1.5204  0.8654
## Firm_2    1  18 -0.597193025 -2.95851997  0.03895745 -1.5108  0.9534
## Firm_3    3  16 -0.257291999 -1.97135237  0.29979349 -1.3754  1.1522
## Firm_4    0  19 -0.172070863 -1.21388533  0.67084314 -1.5204  0.8654
## Firm_5    1  18 -0.444726602 -1.69693864  0.43286539 -1.5108  0.9534
## Firm_6    2  17  0.062922587  0.72626225  0.99275186 -1.4034  1.0344
## Firm_7    2  17 -0.013830411 -0.08442715  0.94937802 -1.4034  1.0344
## Firm_8    1  18 -0.365024064 -2.26588313  0.18328720 -1.5108  0.9534
## Firm_9    0  19 -0.247940802 -1.72877417  0.41668489 -1.5204  0.8654
## Firm_10   1  18 -0.197003144 -1.52323519  0.52199891 -1.5108  0.9534
```

```
purtest(ts_DLInvest, test = "ips", exo = "intercept",
       lags = "AIC", pmax = 4)
```

```
##
## Im-Pesaran-Shin Unit-Root Test (ex. var.: Individual Intercepts)
##
## data: ts_DLInvest
## Wtbar = -11.14, p-value < 0.000000000000000022
## alternative hypothesis: stationarity
```

```
summary(purtest(ts_DLInvest, test = "ips", exo = "intercept",
               lags = "AIC", pmax = 4))
```

```
## Im-Pesaran-Shin Unit-Root Test
## Exogenous variables: Individual Intercepts
## Automatic selection of lags using AIC: 0 - 3 lags (max: 4)
## statistic (Wtbar): -11.14
## p-value: 0
##
##      lags obs      rho      trho      p.trho      mean      var
## Firm_1    0  18 -0.9778879 -4.196070  0.000665206460867 -1.5188  0.8798
## Firm_2    3  15 -2.7288380 -4.327510  0.000392850467043 -1.3660  1.1810
```

```
## Firm_3      2  16 -1.7848678 -5.650917 0.000000790572465 -1.3952 1.0562
## Firm_4      1  17 -1.5737470 -5.320458 0.000004309793729 -1.5082 0.9726
## Firm_5      2  16 -2.8118586 -5.181259 0.000008564306919 -1.3952 1.0562
## Firm_6      1  17 -1.4850311 -4.776618 0.000057171776145 -1.5082 0.9726
## Firm_7      1  17 -2.1076263 -6.541451 0.000000005364403 -1.5082 0.9726
## Firm_8      2  16 -1.7790739 -4.976338 0.000022818242011 -1.3952 1.0562
## Firm_9      0  18 -1.2600946 -5.304469 0.000004667479100 -1.5188 0.8798
## Firm_10     0  18 -0.7889550 -3.421651 0.010275559745253 -1.5188 0.8798
```

7.1.4 Ejemplo: Modelo Dif-in-Dif para el INPC de análisis clínicos (update nov-2022)

Este ejemplo modela el efecto que tiene que una empresa de laboratorios de análisis clínicos compre a sus competidores en diferentes ciudades del país. Para lo cual usaremos información del Índice Nacional de Precios al Consumidor de los servicios de análisis clínicos.

Dependencies and Setup

Importamos Datos desde un archivo de Excel. Los datos están en formato panel e incluyen:

- INPC (Indice Nacional de Precios al Consumidor (2QJul2018 = 100)) para un tipo de servicio. Source: <https://www.inegi.org.mx/programas/inpc/2018/#Tabulados>
- Dummy variable to indicate the time when the treatment started
- Dummy variable to identify the group exposed to the treatment
- Pend: Create an interaction between time and treated and add trend variable
- OTRAS variables de control

Partamos de la siguiente ecuación:

$$y_{it} = \beta_1 + \beta_2 T_t + \beta_3 D_t + \theta T_t \times D_t + \gamma \mathbf{X}_{it} + \varepsilon_{it}(\#eq : eq_D i D) \quad (7.5)$$

Donde y_{it} es la variable sobre la cual queremos evaluar el efecto de un tratamiento, \mathbf{X}_{it} es un conjunto de variables de control, que incluye a las variables de efectos fijos (temporales, de individuos o del tipo que se decida) y, en su caso, una tendencia, entre otras. Finalmente, las variables T_t y D_i son variables dicotómicas que identifican con el valor de 1 el momento a partir del cual se observa el tratamiento y con 0 en cualquier otro caso, y con 1 los individuos que fueron tratados y con 0 en cualquier otro caso, respectivamente.

De esta forma, el producto $T_t \times D_t$ sería una variable dicotómica que indicaría con un 1 a aquellos individuos que fueron tratados a partir del momento en que se implementó el tratamiento y con 0 en cualquier otro caso. Bajo este escenario, θ es un coeficiente que captura el efecto del tratamiento, condicional a que se ha controlado por una serie de factores.

Otra forma de ver lo que se indica en la ecuación @ref(eq:eq_DiD) es como la comparación de la diferencia en diferencia descrita por:

$$\begin{aligned} E &= [(\bar{y}_{exit}|treatment) - (\bar{y}_{baseline}|treatment)] \\ &\quad - [(\bar{y}_{exit}|placebo) - (\bar{y}_{baseline}|placebo)](\#eq : eq_D i D_E) \end{aligned} \quad (7.6)$$

La expresión @ref(eq:eq_DiD_E) cuantifica el efecto que tiene el tratamiento como la diferencia de medias de dos grupos: un grupo tratado y otro no tratado. Dentro de los cuales se han comparado las medias respecto de una línea base. Podemos plantear esto desde esta otra perspectiva. Pensemos que solo tenemos dos periodos $t = 1, 2$, en el cual en 1 no ha ocurrido el tratamiento y en 2 ya ha ocurrido. Así, el cambio en la variable de respuesta para los individuos tratados será (suponiendo que omitimos la matriz \mathbf{X}_{it} en la ecuación @ref(eq:eq_DiD)):

$$\begin{aligned} \mathbb{E}[y_{i2}|D_i = 1] - \mathbb{E}[y_{i1}|D_i = 1] &= (\beta_1 + \beta_2 + \beta_3 + \theta) - (\beta_1 + \beta_3 + \theta) \\ &= \beta_2 + \theta(\#eq : eq_D i D_{Treat}) \end{aligned} \quad (7.7)$$

Ahora, sobre los no tratados:

$$\begin{aligned} \mathbb{E}[y_{i2}|D_i = 0] - \mathbb{E}[y_{i1}|D_i = 0] &= (\beta_1 + \beta_2) - (\beta_1) \\ &= \beta_2(\#eq : eq_D i D_{NoTreat}) \end{aligned} \quad (7.8)$$

Así, la diferencia en diferencia resulta de restar la ecuación @ref(eq:eq_DiD_NoTreat) a la ecuación @ref(eq:eq_DiD_Treat):

$$\begin{aligned} &(\mathbb{E}[y_{i2}|D_i = 1] - \mathbb{E}[y_{i1}|D_i = 1]) - (\mathbb{E}[y_{i2}|D_i = 0] - \mathbb{E}[y_{i1}|D_i = 0]) \\ &= \theta(\#eq : eq_{Aplicada}) \end{aligned} \quad (7.9)$$

En nuestro caso, utilizaremos como tratamiento la entrada al mercado de un laboratorio de análisis clínicos en una ciudad determinada.

```
## [1] 13420      17
```

```
## # A tibble: 6 x 17
```

##	Fecha	Mes	Anio	Ciudad	INPC	Dummy_FIRMA	Marca_01	Marca_02	Marca_03	Marca_04
##	<chr>	<dbl>	<dbl>	<chr>	<chr>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>
## 1	Jul ~	7	2002	Área ~	68.3~	0	0	0	0	0
## 2	Ago ~	8	2002	Área ~	68.3~	0	0	0	0	0

```
## 3 Sep ~      9  2002 Área ~ 68.3~      0      0      0      0      0
## 4 Oct ~     10  2002 Área ~ 68.3~      0      0      0      0      0
## 5 Nov ~     11  2002 Área ~ 68.3~      0      0      0      0      0
## 6 Dic ~     12  2002 Área ~ 68.3~      0      0      0      0      0
## # i 7 more variables: Marca_05 <dbl>, Marca_06 <dbl>, Marca_07 <dbl>,
## #   Marca_08 <dbl>, Marca_09 <dbl>, Marca_10 <dbl>, Marca_11 <dbl>
```

Selección de ciudades:

```
Data <- Data[ which( Data$Ciudad != 'Atacomulco, Edo. de Méx.' &
  Data$Ciudad != 'Cancún, Q.R.' &
  Data$Ciudad != 'Coatzacoalcas, Ver.' &
  Data$Ciudad != 'Esperanza, Son.' &
  Data$Ciudad != 'Izúcar de Matamoros, Pue.' &
  Data$Ciudad != 'Pachuca, Hgo.' &
  Data$Ciudad != 'Tuxtla Gutiérrez, Chis.' &
  Data$Ciudad != 'Zacatecas, Zac.' ), ]
```

Transformaciones:

```
# Log de INPC
Data$LINPC <- log( as.numeric(Data$INPC) , base = exp(1))

# Volvemos factor a la fecha
Data$Periodo <- factor(Data$Fecha, order = TRUE,
  levels = c( "Jul 2002", "Ago 2002", "Sep 2002",
    "Oct 2002", "Nov 2002", "Dic 2002",
    "Ene 2003", "Feb 2003", "Mar 2003",
    "Abr 2003", "May 2003", "Jun 2003",
    "Jul 2003", "Ago 2003", "Sep 2003",
    "Oct 2003", "Nov 2003", "Dic 2003",
    "Ene 2004", "Feb 2004", "Mar 2004",
    "Abr 2004", "May 2004", "Jun 2004",
    "Jul 2004", "Ago 2004", "Sep 2004",
    "Oct 2004", "Nov 2004", "Dic 2004",
    "Ene 2005", "Feb 2005", "Mar 2005",
    "Abr 2005", "May 2005", "Jun 2005",
    "Jul 2005", "Ago 2005", "Sep 2005",
    "Oct 2005", "Nov 2005", "Dic 2005",
    "Ene 2006", "Feb 2006", "Mar 2006",
    "Abr 2006", "May 2006", "Jun 2006",
    "Jul 2006", "Ago 2006", "Sep 2006",
    "Oct 2006", "Nov 2006", "Dic 2006",
    "Ene 2007", "Feb 2007", "Mar 2007",
    "Abr 2007", "May 2007", "Jun 2007",
```

```

"Jul 2007", "Ago 2007", "Sep 2007",
"Oct 2007", "Nov 2007", "Dic 2007",
"Ene 2008", "Feb 2008", "Mar 2008",
"Abr 2008", "May 2008", "Jun 2008",
"Jul 2008", "Ago 2008", "Sep 2008",
"Oct 2008", "Nov 2008", "Dic 2008",
"Ene 2009", "Feb 2009", "Mar 2009",
"Abr 2009", "May 2009", "Jun 2009",
"Jul 2009", "Ago 2009", "Sep 2009",
"Oct 2009", "Nov 2009", "Dic 2009",
"Ene 2010", "Feb 2010", "Mar 2010",
"Abr 2010", "May 2010", "Jun 2010",
"Jul 2010", "Ago 2010", "Sep 2010",
"Oct 2010", "Nov 2010", "Dic 2010",
"Ene 2011", "Feb 2011", "Mar 2011",
"Abr 2011", "May 2011", "Jun 2011",
"Jul 2011", "Ago 2011", "Sep 2011",
"Oct 2011", "Nov 2011", "Dic 2011",
"Ene 2012", "Feb 2012", "Mar 2012",
"Abr 2012", "May 2012", "Jun 2012",
"Jul 2012", "Ago 2012", "Sep 2012",
"Oct 2012", "Nov 2012", "Dic 2012",
"Ene 2013", "Feb 2013", "Mar 2013",
"Abr 2013", "May 2013", "Jun 2013",
"Jul 2013", "Ago 2013", "Sep 2013",
"Oct 2013", "Nov 2013", "Dic 2013",
"Ene 2014", "Feb 2014", "Mar 2014",
"Abr 2014", "May 2014", "Jun 2014",
"Jul 2014", "Ago 2014", "Sep 2014",
"Oct 2014", "Nov 2014", "Dic 2014",
"Ene 2015", "Feb 2015", "Mar 2015",
"Abr 2015", "May 2015", "Jun 2015",
"Jul 2015", "Ago 2015", "Sep 2015",
"Oct 2015", "Nov 2015", "Dic 2015",
"Ene 2016", "Feb 2016", "Mar 2016",
"Abr 2016", "May 2016", "Jun 2016",
"Jul 2016", "Ago 2016", "Sep 2016",
"Oct 2016", "Nov 2016", "Dic 2016",
"Ene 2017", "Feb 2017", "Mar 2017",
"Abr 2017", "May 2017", "Jun 2017",
"Jul 2017", "Ago 2017", "Sep 2017",
"Oct 2017", "Nov 2017", "Dic 2017",
"Ene 2018", "Feb 2018", "Mar 2018",
"Abr 2018", "May 2018", "Jun 2018",
"Jul 2018", "Ago 2018", "Sep 2018",

```

```

"Oct 2018", "Nov 2018", "Dic 2018",
"Ene 2019", "Feb 2019", "Mar 2019",
"Abr 2019", "May 2019", "Jun 2019",
"Jul 2019", "Ago 2019", "Sep 2019",
"Oct 2019", "Nov 2019", "Dic 2019",
"Ene 2020", "Feb 2020", "Mar 2020",
"Abr 2020", "May 2020", "Jun 2020",
"Jul 2020", "Ago 2020", "Sep 2020",
"Oct 2020", "Nov 2020", "Dic 2020",
"Ene 2021", "Feb 2021", "Mar 2021",
"Abr 2021", "May 2021", "Jun 2021",
"Jul 2021", "Ago 2021", "Sep 2021",
"Oct 2021", "Nov 2021", "Dic 2021",
"Ene 2022", "Feb 2022", "Mar 2022",
"Abr 2022", "May 2022", "Jun 2022",
"Jul 2022", "Ago 2022", "Sep 2022", "
Oct 2022", "Nov 2022", "Dic 2022" ))

```

Plot

```

Data %>%
  ggplot( aes(x = Periodo, y = LINPC, group = Ciudad, color = Ciudad )) +
  geom_line() +
  theme(legend.title = element_text(color = "black", size = 8),
        legend.text = element_text(color = "black", size = 5 )) +
  theme(legend.position="bottom") +
  theme(axis.text.x = element_text( size = 8, angle = 90))

```



Hagamos una pruebas de Raíces Unitarias.

Change to format wide:

```
## # A tibble: 6 x 48
##   Fecha   `Acapulco, Gro.` `Aguascalientes, Ags.` `Área Metropolitana de la Cd~1
##   <chr>         <dbl>         <dbl>         <dbl>
## 1 Abr 2003      4.39          4.35          4.26
## 2 Abr 2004      4.39          4.36          4.31
## 3 Abr 2005      4.44          4.36          4.36
## 4 Abr 2006      4.47          4.33          4.40
## 5 Abr 2007      4.54          4.38          4.42
## 6 Abr 2008      4.60          4.41          4.44
## # i abbreviated name: 1: `Área Metropolitana de la Cd. de México`
## # i 44 more variables: `Campeche, Camp.` <dbl>, `Cd. Acuña, Coah.` <dbl>,
## #   `Cd. Jiménez, Chih.` <dbl>, `Cd. Juárez, Chih.` <dbl>,
## #   `Chetumal, Q.R.` <dbl>, `Chihuahua, Chih.` <dbl>, `Colima, Col.` <dbl>,
## #   `Córdoba, Ver.` <dbl>, `Cortazar, Gto.` <dbl>, `Cuernavaca, Mor.` <dbl>,
## #   `Culiacán, Sin.` <dbl>, `Durango, Dgo.` <dbl>, `Fresnillo, Zac.` <dbl>,
## #   `Guadalajara, Jal.` <dbl>, `Hermosillo, Son.` <dbl>, ...
```

Time Series:

```
ts_LINPC <- ts( LINPC[c( 2:48 )],
               start = c(2002, 7),
```

```

        freq = 12)

ts_DLINPC <- diff(ts( LINPC[c( 2:48 )],
                    start = c(2002, 7),
                    freq = 12),
                lag = 1,
                differences = 1)

```

Unit Root Test: Test specifies the type of test to be performed among Levin et al. (2002), Im et al. (2003), Maddala and Wu (1999) and Hadri (2000)

Optimal lags:

$$\begin{aligned}
 p &= \text{Int}(4 * (T/100)^{(1/4)}) \\
 &= \text{Int}(4 * (244/100)^{(1/4)}) \\
 &= \text{Int}(4.99) \\
 &= 4
 \end{aligned}$$

Consider Levin et al. (2002):

```

purtest(ts_LINPC, test = "levinlin", exo = "trend", # exo = c("none", "intercept", "tr
lags = "AIC", pmax = 4)

```

```

##
## Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and
## Trend)
##
## data: ts_LINPC
## z = -56.022, p-value < 2.2e-16
## alternative hypothesis: stationarity

```

```

purtest(ts_DLINPC, test = "levinlin", exo = "trend", # exo = c("none", "intercept", "t
lags = "AIC", pmax = 4)

```

```

##
## Levin-Lin-Chu Unit-Root Test (ex. var.: Individual Intercepts and
## Trend)
##
## data: ts_DLINPC
## z = -107.84, p-value < 2.2e-16
## alternative hypothesis: stationarity

```

Consider Choi (2001):


```
purtest(ts_LINPC, test = "logit", exo = "trend", # exo = c("none", "intercept", "trend"),
        lags = "AIC", pmax = 4)
```

```
##
## Choi's Logit Unit-Root Test (ex. var.: Individual Intercepts and Trend)
##
## data: ts_LINPC
## L* = -85.506, df = 239, p-value < 2.2e-16
## alternative hypothesis: stationarity
```

```
purtest(ts_DLINPC, test = "logit", exo = "trend", # exo = c("none", "intercept", "trend"),
        lags = "AIC", pmax = 4)
```

```
##
## Choi's Logit Unit-Root Test (ex. var.: Individual Intercepts and Trend)
##
## data: ts_DLINPC
## L* = -278.56, df = 239, p-value < 2.2e-16
## alternative hypothesis: stationarity
```

Consider Hadri (2000):

```
purtest(ts_LINPC, test = "hadri", exo = "trend", # exo = c("none", "intercept", "trend"),
        lags = "AIC", pmax = 4)
```

```
##
## Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked.
## Consistent)
##
## data: ts_LINPC
## z = -2.7, p-value = 0.9965
## alternative hypothesis: at least one series has a unit root
```

```
purtest(ts_DLINPC, test = "hadri", exo = "trend", # exo = c("none", "intercept", "trend"),
        lags = "AIC", pmax = 4)
```

```
##
## Hadri Test (ex. var.: Individual Intercepts and Trend) (Heterosked.
## Consistent)
##
## data: ts_DLINPC
## z = -8.9351, p-value = 1
## alternative hypothesis: at least one series has a unit root
```

Estimación PANEL con TREND:

```
# Modelo Final
fixed <- plm( LINPC ~ Dummy_FIRMA + Marca_01 + Marca_02 + Marca_03 + Marca_04 + Marca_05 +
              Marca_06 + Marca_07 + Marca_08 + Marca_09 + Marca_10 + Marca_11 + as.numeric(Periodo),
              data = Data,
              index=c("Ciudad", "Periodo"),
              model = "within")
```

```
## Warning in pdata.frame(data, index): at least one NA in at least one index dimension
## to find out which, use, e.g., table(index(your_pdataframe), useNA = "ifany")
```

```
summary(fixed)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = LINPC ~ Dummy_FIRMA + Marca_01 + Marca_02 + Marca_03 +
##      Marca_04 + Marca_05 + Marca_06 + Marca_07 + Marca_08 + Marca_09 +
##      Marca_10 + Marca_11 + as.numeric(Periodo), data = Data, model = "within",
##      index = c("Ciudad", "Periodo"))
##
## Balanced Panel: n = 47, T = 243, N = 11421
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -0.4220510 -0.0416329  0.0026154  0.0471479  0.2275376
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## Dummy_FIRMA   -0.05972092  0.01892851  -3.1551 0.0016087 **
## Marca_01        0.04545555  0.02069006   2.1970 0.0280423 *
## Marca_02        0.05452288  0.02372909   2.2977 0.0215956 *
## Marca_03       -0.02925504  0.02041744  -1.4328 0.1519296
## Marca_04       -0.07440826  0.02418665  -3.0764 0.0021000 **
## Marca_05        0.02633713  0.01045219   2.5198 0.0117567 *
## Marca_06        0.12521494  0.01991655   6.2870 3.356e-10 ***
## Marca_07       -0.09242328  0.02514895  -3.6750 0.0002389 ***
## Marca_08        0.03362324  0.02086579   1.6114 0.1071193
## Marca_09        0.05630828  0.02121641   2.6540 0.0079655 **
## Marca_10       -0.04866105  0.02638856  -1.8440 0.0652062 .
## Marca_11       -0.03582620  0.03031018  -1.1820 0.2372361
## as.numeric(Periodo) 0.00220384  0.00001018 216.4961 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Total Sum of Squares:    330.79
## Residual Sum of Squares: 60.08
## R-Squared:              0.81838
## Adj. R-Squared: 0.81743
## F-statistic: 3937.79 on 13 and 11361 DF, p-value: < 2.22e-16
```

```
# Modelo Final
fixed <- plm( LINPC ~ Dummy_FIRMA + as.numeric(Periodo),
              data = Data,
              index=c("Ciudad", "Periodo"),
              model = "within")
```

```
## Warning in pdata.frame(data, index): at least one NA in at least one index dimension in result
## to find out which, use, e.g., table(index(your_pdataframe), useNA = "ifany")
```

```
summary(fixed)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = LINPC ~ Dummy_FIRMA + as.numeric(Periodo), data = Data,
##      model = "within", index = c("Ciudad", "Periodo"))
##
## Balanced Panel: n = 47, T = 243, N = 11421
##
## Residuals:
##      Min.      1st Qu.      Median      3rd Qu.      Max.
## -0.4337032 -0.0421466  0.0026796  0.0480377  0.2504866
##
## Coefficients:
##              Estimate Std. Error t-value Pr(>|t|)
## Dummy_FIRMA    -0.02239578  0.00417897  -5.3592 0.00000008525 ***
## as.numeric(Periodo)  0.00220318  0.00001032 213.4825 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    330.79
## Residual Sum of Squares: 61.94
## R-Squared:              0.81275
## Adj. R-Squared: 0.81196
## F-statistic: 24680.5 on 2 and 11372 DF, p-value: < 2.22e-16
```

7.1.5 Panel VAR

En esta sección extenderemos el caso del modelo VAR(p) a uno en forma panel. En este caso asumimos que las variables exógenas son los p rezagos de las k variables endógenas. Consideremos el siguiente caso de un modelo panel VAR con efectos fijos –ciertamente es posible hacer estimaciones con efectos aleatorios, no obstante, requiere de supuestos adicionales que no contemplamos en estas notas–, el cual es la forma más común de estimación:

$$\mathbf{Y}_{it} = \mu_i + \sum_{l=1}^p \mathbf{A}_l \mathbf{Y}_{it-l} + \mathbf{B} \mathbf{X}_{it} + \varepsilon_{it} (\#eq : eq_PVAR) \quad (7.10)$$

Donde \mathbf{Y}_{it} es un vector de variables endógenas estacionarias para la unidad de corte transversal i en el tiempo t , \mathbf{X}_{it} es una matriz que contiene variables exógenas, y ε_{it} es un término de error que cumple con:

$$\begin{aligned} \mathbb{E}[\varepsilon_{it}] &= 0 \\ Var[\varepsilon_{it}] &= \Sigma_\varepsilon \end{aligned}$$

Donde Σ_ε es una matriz definida positiva.

Para el proceso de estimación, la ecuación @ref(eq:eq_PVAR) se modifica en su versión en diferencias para quedar como:

$$\Delta \mathbf{Y}_{it} = \sum_{l=1}^p \mathbf{A}_l \Delta \mathbf{Y}_{it-l} + \mathbf{B} \Delta \mathbf{X}_{it} + \Delta \varepsilon_{it} (\#eq : eq_Dinamic_PVAR) \quad (7.11)$$

La ecuación @ref(eq:eq_Dinamic_PVAR) se estima por un GMM.

7.2 Ejemplos: Panel VAR(p)

Dependencies and Setup

7.2.1 Ejemplo 1

We used the dynamic panel literature by Arellano and Bond (1991), Blundell and Bond (1998) and Roodman (2009b). This data set describes employment, wages, capital and output of 140 firms in the United Kingdom from 1976 to 1984. We estimate: Employment is explained by past values of employment (“1” lags), current and first lag of wages and output and current value of capital.


```

system_instruments = FALSE,
max_instr_dependent_vars = 99,
min_instr_dependent_vars = 2L,
collapse = FALSE)

```

```
summary(Arellano_Bond_1991_table4b)
```

Dynamic Panel VAR estimation, two-step GMM

Transformation: First-differences
 Group variable: id
 Time variable: year
 Number of observations = 611
 Number of groups = 140
 Obs per group: min = 4
 Obs per group: avg = 4.36428571428571
 Obs per group: max = 6
 Number of instruments = 38

n

lag1_n

0.4742*

(0.1854)

lag2_n

-0.0530

(0.0517)

w

-0.5132***

(0.1456)

wL1

0.2246

(0.1419)

k

0.2927***

(0.0626)

ys

0.6098***

(0.1563)

ysL1

-0.4464*

(0.2173)

yr1979

0.0105

(0.0099)

yr1980

0.0247

(0.0158)

yr1981

-0.0158

(0.0267)

yr1982

-0.0374

(0.0300)

yr1983

-0.0393

(0.0347)

yr1984

-0.0495

(0.0349)

***p < 0.001; **p < 0.01; *p < 0.05

Instruments for equation

Standardw, wL1, k, ys, ysL1, yr1979, yr1980, yr1981, yr1982, yr1983, yr1984

GMM-typeDependent vars: L(2,6))Collapse = FALSE

Hansen test of overid. restrictions: chi2(25) = 30.11 Prob > chi2 = 0.22(Robust, but weakened by many instruments.)

7.2.2 Ejemplo 2:

We used the panel data set consists of 265 Swedish municipalities and covers 9 years (1979-1987). These variables include total expenditures (expenditures), total own-source revenues (revenues) and intergovernmental grants received by the municipality (grants). Source: Dahlberg and Johansson (2000)

Grants from the central to the local government are of three kinds: support to municipalities with small tax capacity, grants toward the running of certain local government activities and grants toward certain investments.

```
data("Dahlberg")
```

```
names(Dahlberg)
```

```
## [1] "id"          "year"        "expenditures" "revenues"    "grants"
```

```
head(Dahlberg)
```

```
##      id year expenditures revenues grants
## 1 114 1979    0.0229736 0.0181770 0.0054429
## 2 114 1980    0.0266307 0.0209142 0.0057304
## 3 114 1981    0.0273253 0.0210836 0.0056647
## 4 114 1982    0.0288704 0.0234310 0.0058859
## 5 114 1983    0.0226474 0.0179979 0.0055908
## 6 114 1984    0.0215601 0.0179949 0.0047536
```

Estimación:


```
ex1_dahlberg_data <- pvargmm(dependent_vars = c("expenditures", "revenues", "grants"),
  lags = 1,
  transformation = "fod",
  data = Dahlberg,
  panel_identifier=c("id", "year"),
  steps = c("twostep"),
  system_instruments = FALSE,
  max_instr_dependent_vars = 99,
  max_instr_predet_vars = 99,
  min_instr_dependent_vars = 2L,
  min_instr_predet_vars = 1L,
  collapse = FALSE
)
```

```
## Warning in pvargmm(dependent_vars = c("expenditures", "revenues", "grants"), :
## The matrix D_e is singular, therefore the general inverse is used
```

```
summary(ex1_dahlberg_data)
```

Dynamic Panel VAR estimation, two-step GMM

Transformation: Forward orthogonal deviations
 Group variable: id
 Time variable: year
 Number of observations = 1855
 Number of groups = 265
 Obs per group: min = 7
 Obs per group: avg = 7
 Obs per group: max = 7
 Number of instruments = 252

expenditures

revenues

grants

lag1__expenditures

0.2846***

0.2583**

0.0167

(0.0664)

(0.0795)

(0.0172)

lag1__revenues

-0.0470
 0.0588
 -0.0405**

(0.0637)
 (0.0726)
 (0.0151)
 lag1_grants
 -1.6746***
 -2.2367***
 0.3204***

(0.2818)
 (0.2846)
 (0.0521)

***p < 0.001; **p < 0.01; *p < 0.05

Instruments for equation

Standard

GMM-typeDependent vars: L(2,7))Collapse = FALSE

Hansen test of overid. restrictions: chi2(243) = 263.01 Prob > chi2 = 0.18(Robust, but weakened by many instruments.)

Model selection procedure of Andrews and Lu (2001) to select the optimal lag length for our example

```
Andrews_Lu_MMSC(ex1_dahlberg_data)
```

```
## $MMSC_BIC
## [1] -1610.877
##
## $MMSC_AIC
## [1] -234.9924
##
## $MMSC_HQIC
## [1] -792.3698
```

```
ex2_dahlberg_data <- pvargmm(dependent_vars = c("expenditures", "revenues", "grants"),
                             lags = 2,
                             transformation = "fod",
                             data = Dahlberg,
                             panel_identifier=c("id", "year"),
                             steps = c("twostep"),
                             system_instruments = FALSE,
                             max_instr_dependent_vars = 99,
                             max_instr_predet_vars = 99,
                             min_instr_dependent_vars = 2L,
                             min_instr_predet_vars = 1L,
                             collapse = FALSE)
```

```
## Warning in pvargmm(dependent_vars = c("expenditures", "revenues", "grants"), :
## The matrix D_e is singular, therefore the general inverse is used
```

```
summary(ex2_dahlberg_data)
```

Dynamic Panel VAR estimation, two-step GMM

Transformation: Forward orthogonal deviations
 Group variable: id
 Time variable: year
 Number of observations = 1590
 Number of groups = 265
 Obs per group: min = 6
 Obs per group: avg = 6
 Obs per group: max = 6
 Number of instruments = 243

expenditures

revenues

grants

lag1_expenditures

0.2572**

0.2486*

0.0135

(0.0893)

(0.1018)

(0.0178)

lag1_revenues

-0.1219

-0.0634

-0.0293

(0.0879)

(0.0997)

(0.0171)

lag1_grants

-3.0718***

-3.5849***

0.1581*

(0.5941)

(0.6168)

(0.0636)

lag2_expenditures

-0.0247

0.0252

0.0178

(0.0791)

(0.0834)

(0.0157)

lag2_revenues

-0.2584**

-0.2446**

-0.0237

(0.0785)

(0.0800)

(0.0157)

lag2_grants

-1.5265***

-1.6687***

0.0702

(0.1873)

(0.2113)

(0.0586)

***p < 0.001; **p < 0.01; *p < 0.05

Instruments for equation

Standard

GMM-typeDependent vars: L(2,6))Collapse = FALSE

Hansen test of overid. restrictions: chi2(225) = 260.11 Prob > chi2 = 0.054(Robust, but weakened by many instruments.)

```
Andrews_Lu_MMSC(ex2_dahlberg_data)
```

```
## $MMSC_BIC
```

```
## [1] -1486.935
```

```
##
```

```
## $MMSC_AIC
```

```
## [1] -213.8917
```

```
##
```

```
## $MMSC_HQIC
```

```
## [1] -734.1071
```

Stability of the autoregressive process:

```
stab_ex1_dahlberg_data <- stability(ex1_dahlberg_data)
```

```
print(stab_ex1_dahlberg_data)
```

```
## Eigenvalue stability condition:
```

```
##
```

```
##           Eigenvalue      Modulus
```

```
## 1 0.5370657+0.00000000i 0.53706574
```

```
## 2 0.0633651+0.06442426i 0.09036383
```

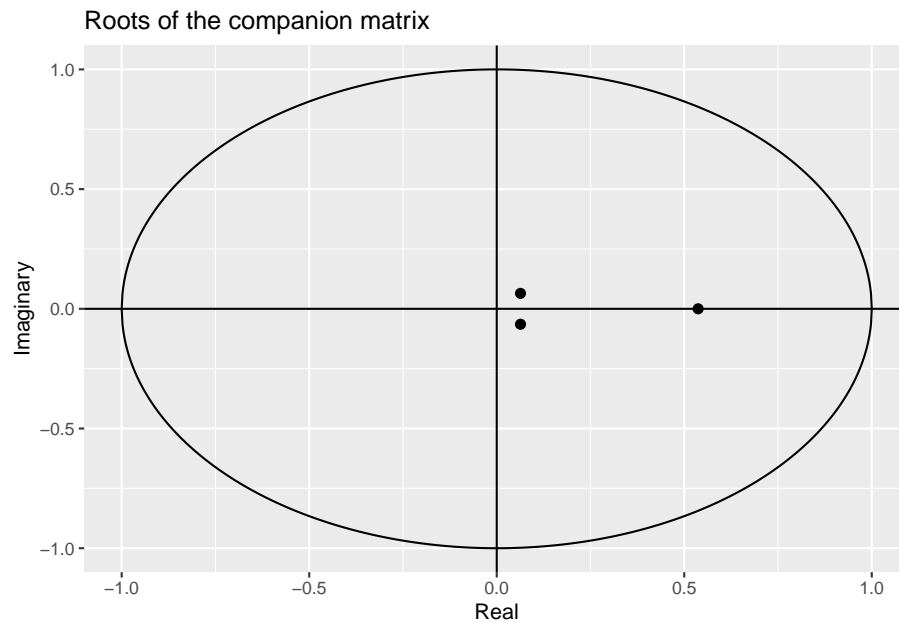
```
## 3 0.0633651-0.06442426i 0.09036383
```

```
##
```

```
## All the eigenvalues lie inside the unit circle.
```

```
## PVAR satisfies stability condition.
```

```
plot(stab_ex1_dahlberg_data)
```



Generate impulse response functions.

```
# ex1_dahlberg_data_oirf <- oirf(ex1_dahlberg_data, n.ahead = 8)
#
# ex1_dahlberg_data_girf <- girf(ex1_dahlberg_data, n.ahead = 8, ma_approx_steps= 8)
#
# ex1_dahlberg_data_bs <- bootstrap_irf(ex1_dahlberg_data, typeof_irf = c("GIRF"),
#                                     n.ahead = 8,
#                                     nof_Nstar_draws = 500,
#                                     confidence.band = 0.95)
#
# plot(ex1_dahlberg_data_girf, ex1_dahlberg_data_bs)
```

7.3 Cointegración

[PENDIENTE]

7.4 Modelos No Lineales

7.4.1 Modelos de cambio de régimen

En años recientes, los modelos de serie de tiempo han sido incorporados en análisis de la existencia de diferentes estados que son generados por procesos estocásticos subyacentes. En esta sección del curso revisaremos algunos modelos de cambio de régimen. Restringimos nuestra revisión a modelos que asuman que la dinámica de las series puede ser descrita por modelos del tipo AR y dejamos fuera procesos del tipo MA.

En general distinguimos que existen dos tipos de modelos:

1. Modelos caracterizados por una variable observable, por lo que tenemos certeza de los regímenes en cada uno de los momentos.
2. Modelos en los que el régimen no puede ser observado por una variable, pero si conocemos el proceso estocástico subyacente.

7.4.2 Regímenes determinados por información observable

En estos casos asumimos que el régimen ocurre en un momento t y puede ser determinado por una variable observable. Este modelo es conocido como el modelo autoregresivo con umbral (TAR, Threshold Autoregressive model). En este caso también diremos que cuando el régimen está determinado por la información de la misma serie será llamado Self-Exciting TAR (SETAR).

Veámos un caso particular. Supongamos que existe un umbral, c , para el régimen que está determinado por $q_t = y_{t-1}$ y que el estado de la naturaleza nos permite establecer dos estados o regímenes:

$$y_t = \begin{cases} \phi_{01} + \phi_{11}y_{t-1} + \varepsilon_t & \text{si } y_{t-1} \leq c \\ \phi_{02} + \phi_{12}y_{t-1} + \varepsilon_t & \text{si } y_{t-1} > c \end{cases} \quad (7.12)$$

Donde asumiremos que ε_t es i.i.d y que cumple con:

$$\mathbb{E}[\varepsilon_t | \Omega_{t-1}] = 0$$

Donde $\Omega_{t-1} = \{y_{t-1}, y_{t-2}, \dots\}$. Existe una variante de este modelo que suaviza la transición entre regímenes conocido como Smooth Transition AR (STAR) y puede ser especificada en su modalidad de dos regímenes como:

$$y_t = (\phi_{01} + \phi_{11}y_{t-1})(1 - G(y_{t-1}; \gamma, c)) + (\phi_{02} + \phi_{12}y_{t-1})G(y_{t-1}; \gamma, c) + \varepsilon_t \quad (7.13)$$

Donde $G(y_{t-1}; \gamma, c)$ es una función de distribución de probabilidad que suaviza la transición entre regímenes. La práctica común es suponer que esta tiene una forma logística:

$$G(y_{t-1}; \gamma, c) = \frac{1}{1 + e^{-\gamma(y_{t-1}-c)}} \quad (7.14)$$

Es posible hacer extensiones de lo anterior a modelos de orden superior; dando como resultado:

$$y_t = \begin{cases} \phi_{01} + \phi_{11}y_{t-1} + \phi_{21}y_{t-2} + \dots + \phi_{p_1 1}y_{t-p_1} + \varepsilon_t & \text{si } y_{t-1} \leq c \\ \phi_{02} + \phi_{12}y_{t-1} + \phi_{22}y_{t-2} + \dots + \phi_{p_2 2}y_{t-p_2} + \varepsilon_t & \text{si } y_{t-1} > c \end{cases} \quad (7.15)$$

En el segundo caso:

$$\begin{aligned} y_t = & (\phi_{01} + \phi_{11}y_{t-1} + \phi_{21}y_{t-2} + \dots + \phi_{p_1 1}y_{t-p_1})(1 - G(y_{t-1}; \gamma, c)) \\ & + (\phi_{02} + \phi_{12}y_{t-1} + \phi_{22}y_{t-2} + \dots + \phi_{p_2 2}y_{t-p_2})G(y_{t-1}; \gamma, c) \\ & + \varepsilon_t \end{aligned}$$

De igual forma que en el caso de los modelos ARIMA y VAR, el número de rezagos utilizados es determinado mediante el uso de criterios de información como el de Akaike:

$$AIC(p_1, p_2) = n_1 \ln(\hat{\sigma}_1^2) + n_2 \ln(\hat{\sigma}_2^2) + 2(p_1 + 1) + 2(p_2 + 1) \quad (7.16)$$

Los modelos SETAR y STAR generan procesos estacionarios siempre que cumplan ciertas condiciones. En estas notas nos enfocaremos únicamente en el modelo SETAR; el cual genera en un proceso estacionario cuando:

1. $\phi_{11} < 1, \phi_{12} < 1, \phi_{11} \cdot \phi_{12} < 1$
2. $\phi_{11} = 1, \phi_{12} < 1, \phi_{01} > 0$
3. $\phi_{11} < 1, \phi_{12} = 1, \phi_{02} < 0$
4. $\phi_{11} = 1, \phi_{12} = 1, \phi_{02} < 0 < \phi_{01}$
5. $\phi_{11} \cdot \phi_{12} = 1, \phi_{11} < 0, \phi_{02} + \phi_{12} \cdot \phi_{01} > 0$

Finalmente, en ocasiones podemos estar interesados en modelos donde los regímenes sean más de 2, es decir, digamos m umbrales bajo un modelo SETAR o STAR. Por ejemplo, en el caso de un modelo SETAR podemos verificar que m regímenes implican $m + 1$ umbrales: c_0, c_1, \dots, c_m . En cuyo caso:

$$-\infty = c_0 < c_1 < \dots < c_{m-1} < c_m = \infty$$

Así, tendríamos ecuaciones:

$$y_t = \phi_{0j} + \phi_{ij}y_{t-1} + \varepsilon_t \text{ si } c_{j-1} < y_{t-1} < c_j \quad (7.17)$$

Para $j = 1, 2, \dots, m$. De forma similar, podemos recomponer el modelo STAR.

7.4.3 Regímenes determinados por variables no observables

Este tipo de modelos asume que el régimen ocurre en el momento t y que no puede ser observado, ya que este es determinado por un proceso no observable, el cual denotamos como s_t . En el caso de dos regímenes, s_t puede ser asumido como que toma 2 valores: 1 y 2, por ejemplo. Supongamos que el proceso subyacente tiene una forma del tipo AR(1) dado por:

$$y_t = \begin{cases} \phi_{01} + \phi_{11}y_{t-1} + \varepsilon_t & \text{si } s_t = 1 \\ \phi_{02} + \phi_{12}y_{t-1} + \varepsilon_t & \text{si } s_t = 2 \end{cases} \quad (7.18)$$

O en un formato más corto de notación:

$$y_t = \phi_{0s_t} + \phi_{1s_t}y_{t-1} + \varepsilon_t \quad (7.19)$$

Para complementar el modelo, las propiedades del proceso s_t necesitan ser especificadas. El modelo más popular dentro de esta familia es el propuesto por James Hamilton en 1989, el cual es conocido como Markov Switching Model (MSM), en el cual el proceso s_t se asume como un proceso de Markov de primer orden. Esto implica que el régimen actual s_t sólo depende del período s_{t-1} .

Así, el modelo es completado mediante la definición de las probabilidades de transición para moverse del un estado a otro:

$$\begin{aligned} \mathbb{P}(s_t = 1 | s_{t-1} = 1) &= p_{11} \\ \mathbb{P}(s_t = 2 | s_{t-1} = 1) &= p_{12} \\ \mathbb{P}(s_t = 1 | s_{t-1} = 2) &= p_{21} \\ \mathbb{P}(s_t = 2 | s_{t-1} = 2) &= p_{22} \end{aligned}$$

Así, p_{ij} es igual a la probabilidad de que la cadena de Markov pase del estado i en el momento $t-1$ al estado j en el tiempo t . En todo caso asumiremos que $p_{ij} > 0$ y que:

$$\begin{aligned} p_{11} + p_{12} &= 1 \\ p_{21} + p_{22} &= 1 \end{aligned}$$

Otro tipo de probabilidades a analizar son las probabilidades incondicionales de $\mathbb{P}(s_t = i)$, $i = 1, 2$. Usando la teoría ergódica de las cadenas de Markov, estas probabilidades están dadas por:

$$\begin{aligned} \mathbb{P}(s_t = 1) &= \frac{1 - p_{22}}{2 - p_{11} - p_{22}} \\ \mathbb{P}(s_t = 2) &= \frac{1 - p_{11}}{2 - p_{11} - p_{22}} \end{aligned}$$

Un caso más general es el de múltiples regímenes en el cual s_t puede tomar cualquier valor $m > 2$, $m \in \mathbb{N}$. Este modelo se puede escribir como:

$$y_t = \phi_{0j} + \phi_{1j}y_{t-1} + \varepsilon_t \text{ si } s_t = j, \text{ para } j = 1, 2, \dots, m \quad (7.20)$$

Donde las probabilidades de transición estarán dadas por:

$$p_{ij} = \mathbb{P}(s_t = j | s_{t-1} = i) \text{ para } i, j = 1, 2, \dots, m \quad (7.21)$$

Donde la ecuación anterior satisface que $p_{ij} > 0$, $\forall i, j = 1, 2, \dots, m$ y que:

$$\sum_{j=1}^m p_{ij} = 1 \text{ para } i = 1, 2, \dots, m$$

Finalmente, plantearemos el procedimiento empírico seguido para la estimación de este tipo de modelos:

1. Estimar un proceso AR(p)
2. Probar una hipótesis nula de no linealidad de acuerdo con alguno de los modelos SERAR, STAR o MSM
3. Estimar los parámetros
4. Evaluar los resultados del modelo
5. Ajustar, en su caso, la estimación o modelo
6. Pronosticar o realizar análisis impluso-respuesta

7.5 Ejemplo: Modelos de Cambio de Régimen (TAR)

Dependencies and Setup:

Data: Monthly rates of deaths due to flu in the United States for 11 years

```
## X.8113721
## 1 0.4458291
## 2 0.3415985
## 3 0.2774243
## 4 0.2484958
## 5 0.2525427
## 6 0.2466902
```

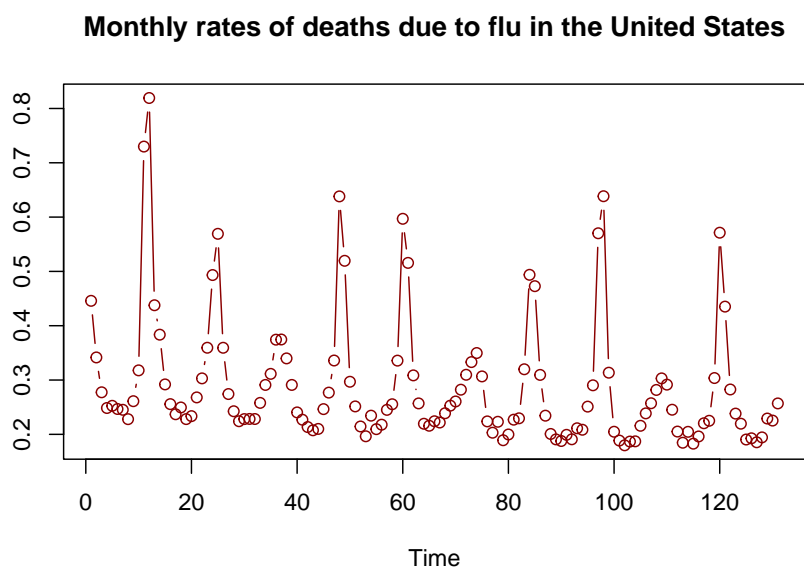
Convert time series

```
flu <- ts(flu)

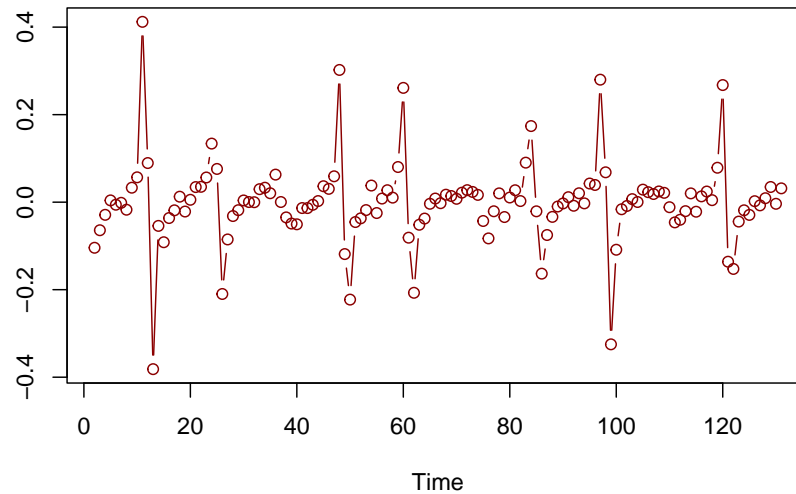
D_flu = diff(flu, lag = 1)
```

Plotting

```
plot(flu, type = "b", col = "darkred", ylab = "",
     main = "Monthly rates of deaths due to flu in the United States")
```



```
plot(D_flu, type="b", col = "darkred", ylab = "",
     main = "Diff Monthly rates of deaths due to flu in the United States")
```

Diff Monthly rates of deaths due to flu in the United States

The tsDyn package in R has simplified this code into a handful of steps:

```
##?setar
```

```
D_flu_tar4_05 <- setar(D_flu, m = 4, thDelay = 0, th = 0.05)
```

```
## Warning:
```

```
## With the threshold you gave (0.05) there is a regime with less than trim=15% observations
```

```
## Warning: Possible unit root in the high regime. Roots are: 0.6182 0.6244 0.6244
```

```
## 0.6182
```

```
summary(D_flu_tar4_05)
```

```
##
```

```
## Non linear autoregressive model
```

```
##
```

```
## SETAR model ( 2 regimes)
```

```
## Coefficients:
```

```
## Low regime:
```

```
##      const.L      phiL.1      phiL.2      phiL.3      phiL.4
```

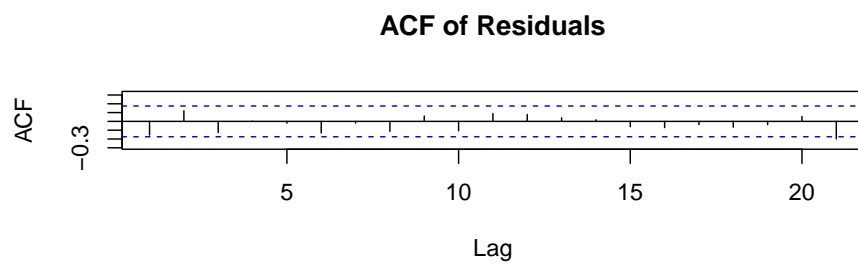
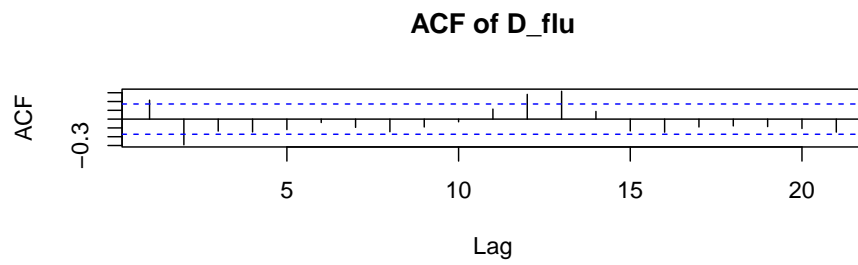
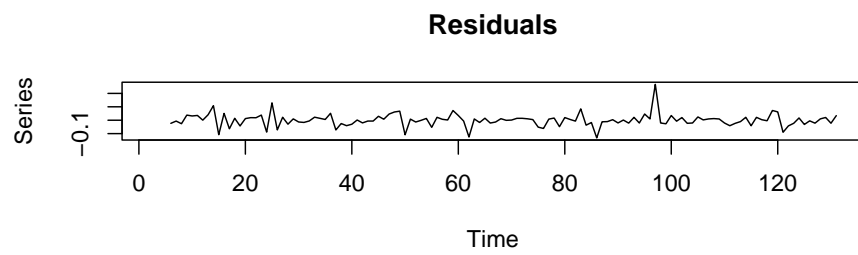
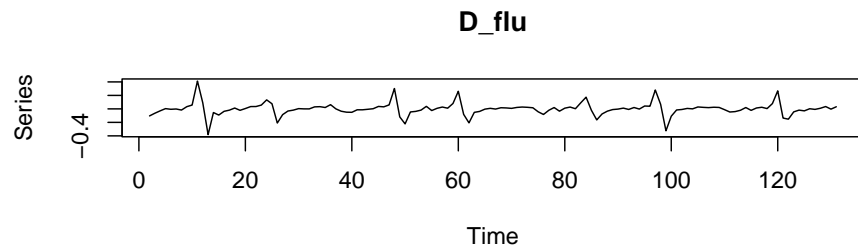
```
## 0.004432028 0.501179574 -0.206693004 0.120140800 -0.122254140
```

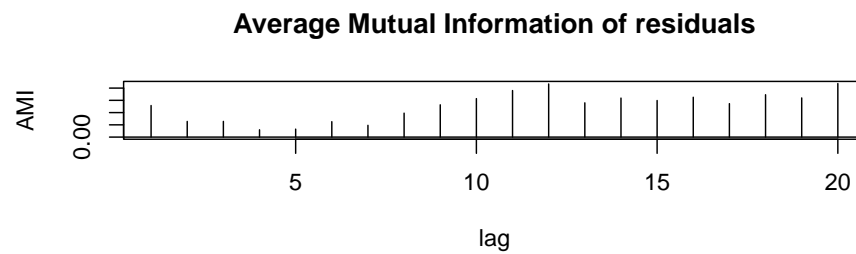
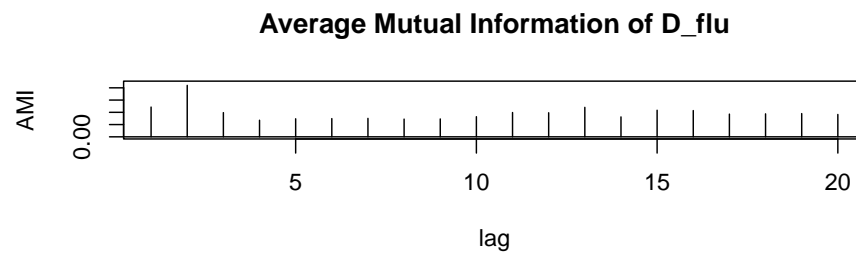
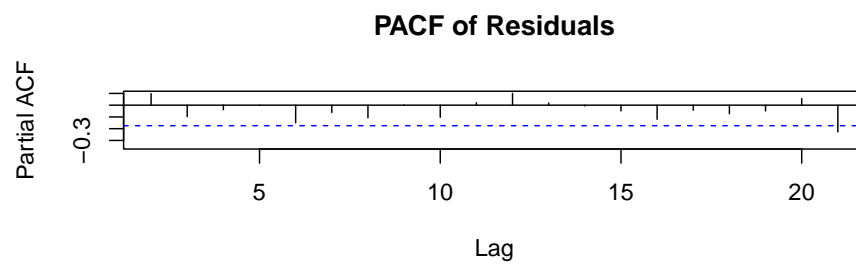
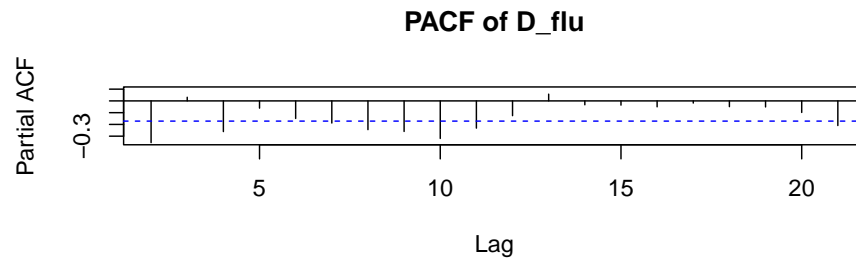
```
##
```

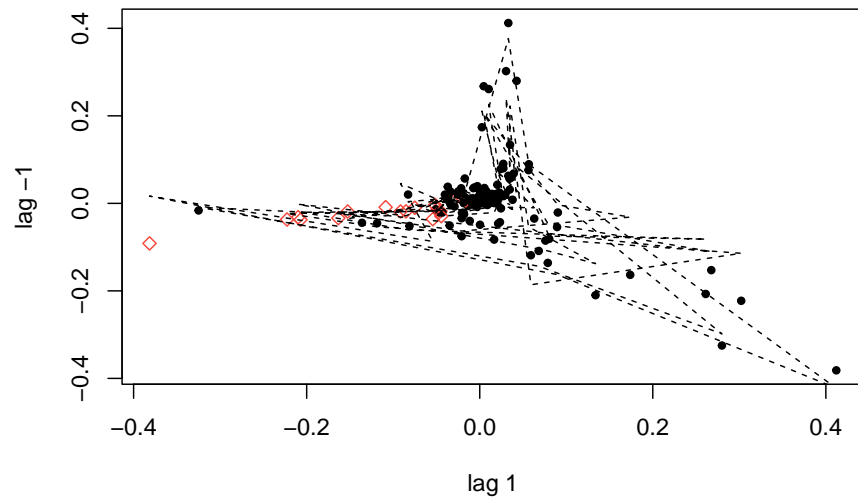
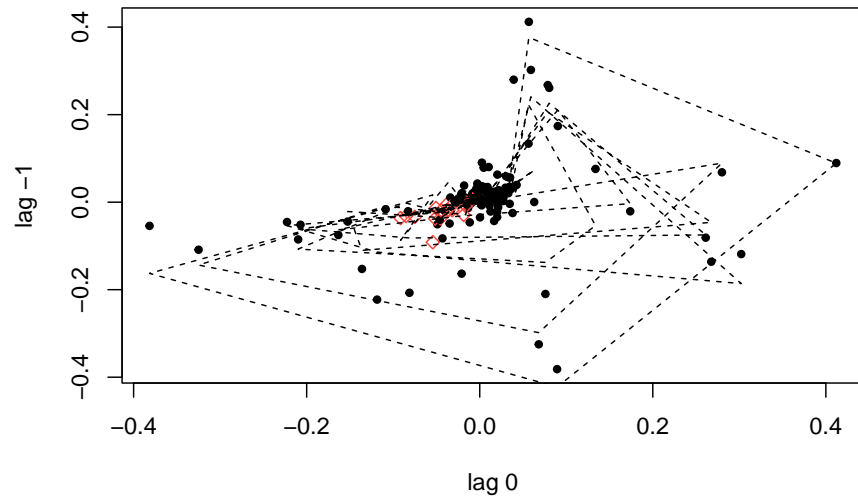
```
## High regime:
```

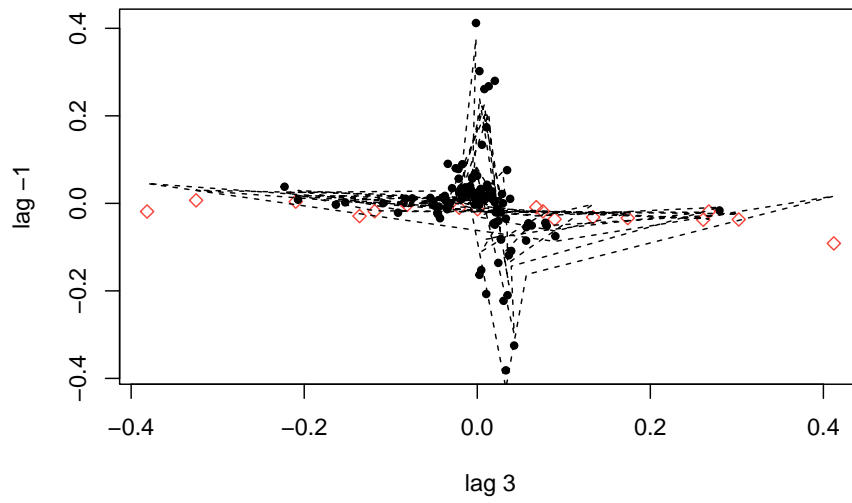
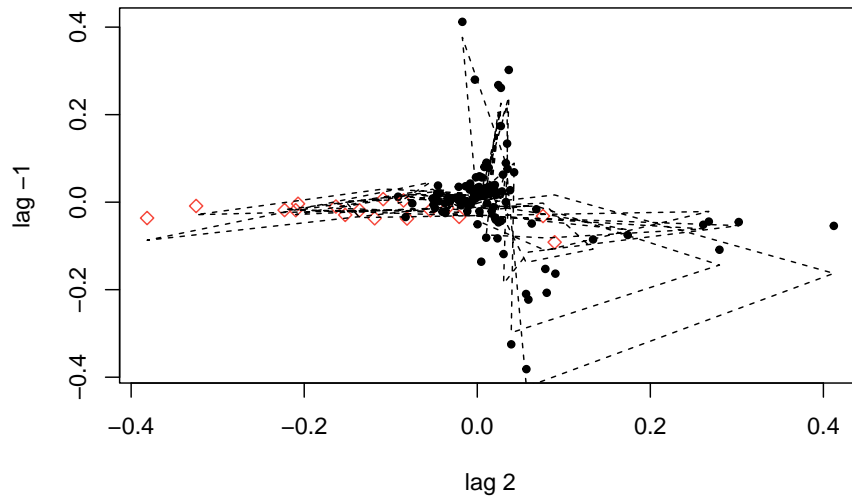
```
##      const.H      phiH.1      phiH.2      phiH.3      phiH.4
## 0.4079353 -0.7483325 -1.0323129 -2.0450407 -6.7117769
##
## Threshold:
## -Variable: Z(t) = + (1) X(t)+ (0)X(t-1)+ (0)X(t-2)+ (0)X(t-3)
## -Value: 0.05 (fixed)
## Proportion of points in low regime: 86.51%      High regime: 13.49%
##
## Residuals:
##      Min      1Q      Median      3Q      Max
## -0.1319314 -0.0217073  0.0015801  0.0196093  0.2674245
##
## Fit:
## residuals variance = 0.002166,  AIC = -778, MAPE = 412.4%
##
## Coefficient(s):
##
##      Estimate Std. Error t value Pr(>|t|)
## const.L 0.0044320  0.0051791   0.8558 0.3938357
## phiL.1  0.5011796  0.0833821   6.0106 2.043e-08 ***
## phiL.2 -0.2066930  0.0608839  -3.3949 0.0009313 ***
## phiL.3  0.1201408  0.0576505   2.0839 0.0392868 *
## phiL.4 -0.1222541  0.0522363  -2.3404 0.0209137 *
## const.H 0.4079353  0.0314121  12.9866 < 2.2e-16 ***
## phiH.1 -0.7483325  0.1118329  -6.6915 7.455e-10 ***
## phiH.2 -1.0323129  0.1420203  -7.2688 4.018e-11 ***
## phiH.3 -2.0450407  0.7055163  -2.8986 0.0044573 **
## phiH.4 -6.7117769  0.8367945  -8.0208 7.911e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold
## Variable: Z(t) = + (1) X(t) + (0) X(t-1)+ (0) X(t-2)+ (0) X(t-3)
##
## Value: 0.05 (fixed)
```

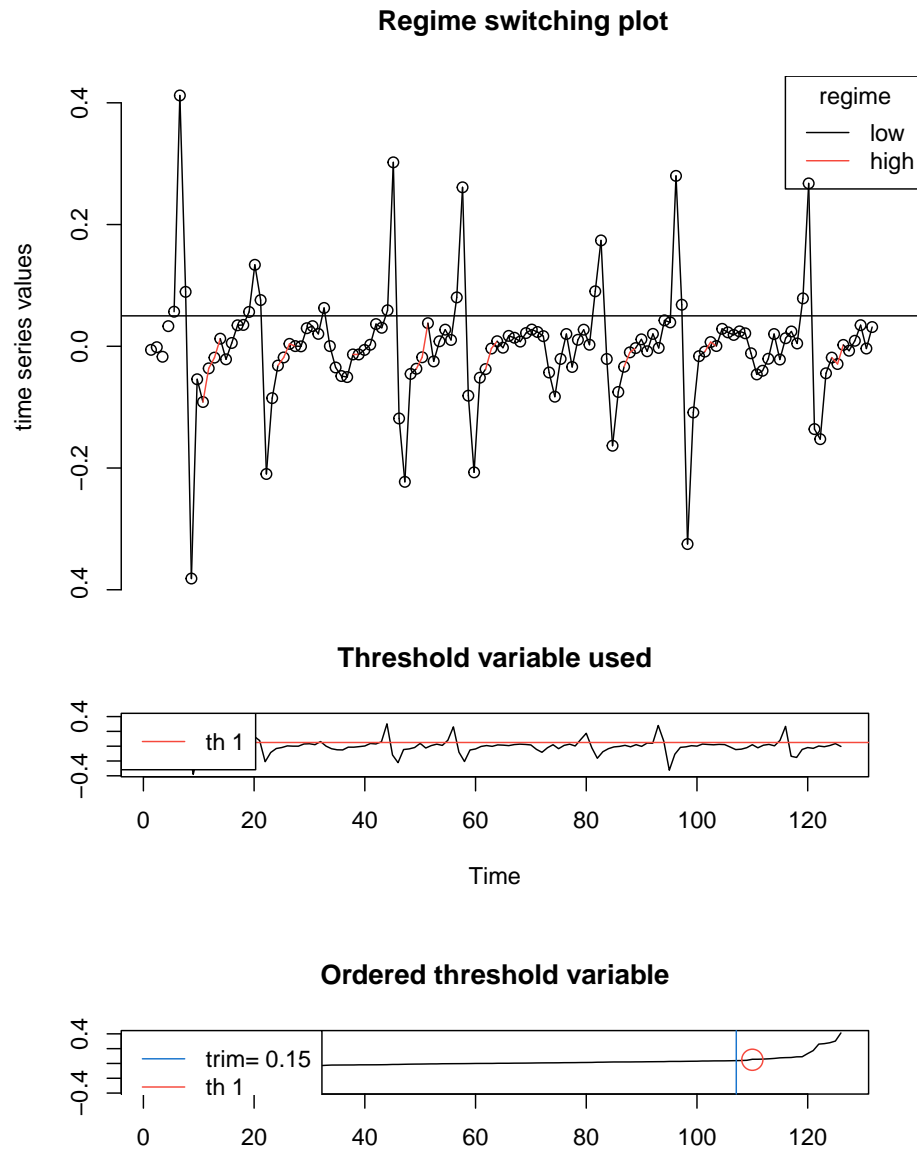
```
plot(D_flu_tar4_05)
```











If we do not provide a threshold to the `th` option, `setar` searches over a grid to choose a threshold ~ 0.038 :

```
D_flu_tar4 <- setar(D_flu, m = 4, thDelay = 0)
```

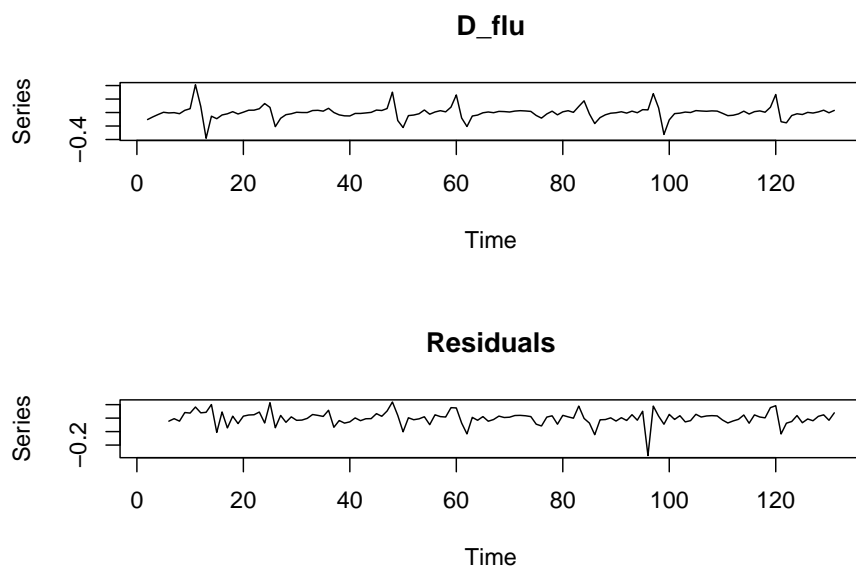
```
## Warning: Possible unit root in the high regime. Roots are: 0.6316 0.7222 0.7222
## 0.6316
```

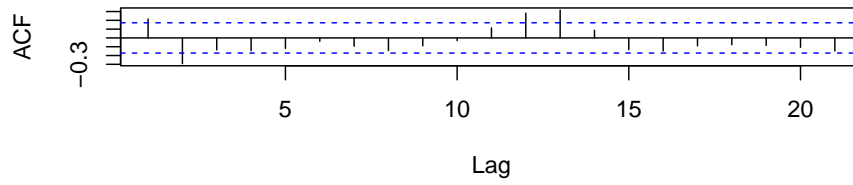
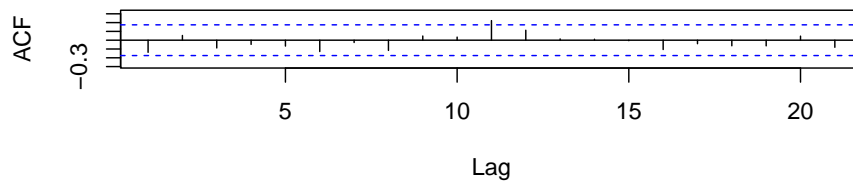
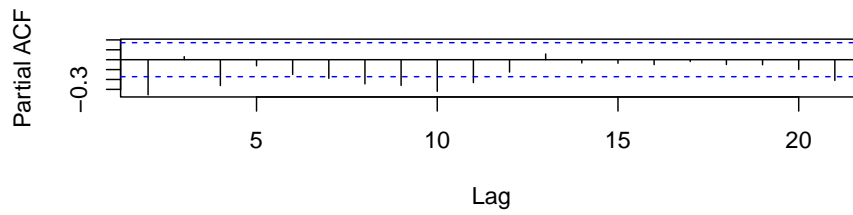
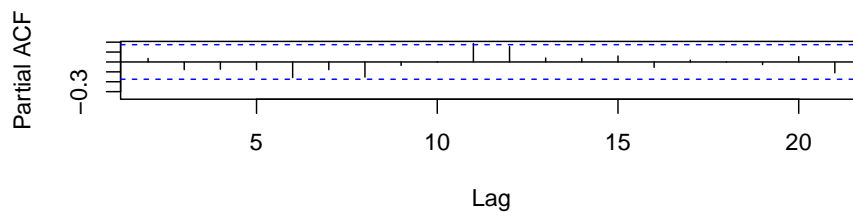
```
summary(D_flu_tar4)
```

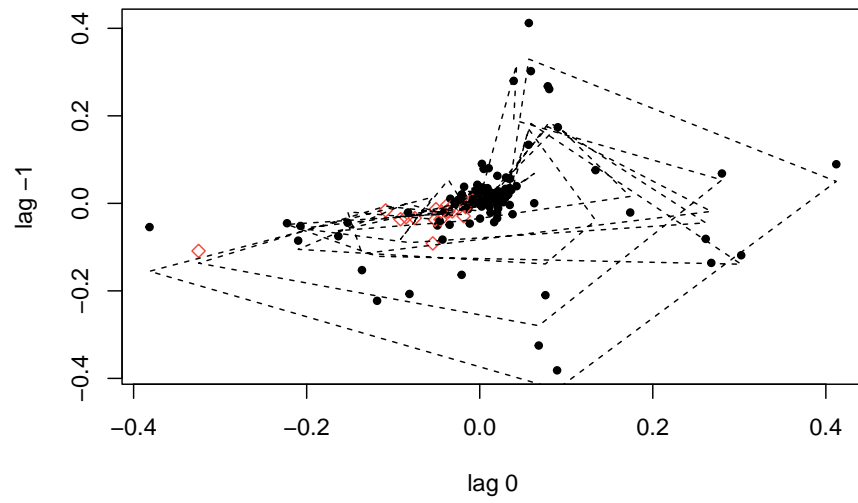
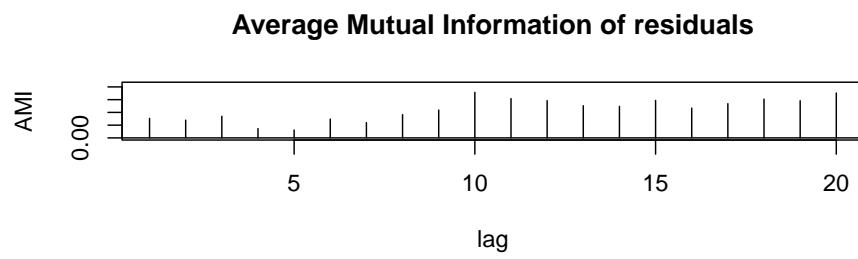
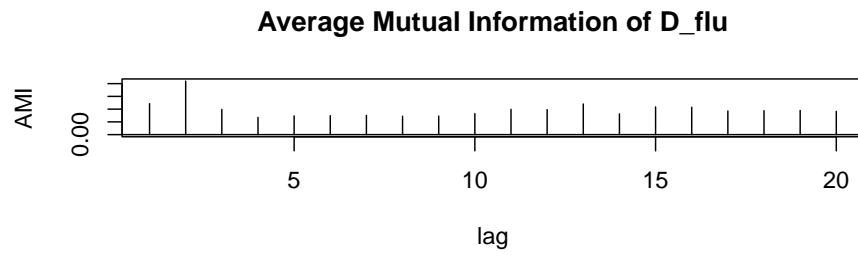
```
##
## Non linear autoregressive model
##
## SETAR model ( 2 regimes)
## Coefficients:
## Low regime:
##      const.L      phiL.1      phiL.2      phiL.3      phiL.4
## -0.00006287604  0.44640751264 -0.23158878472  0.10701408961 -0.14406085874
##
## High regime:
##      const.H      phiH.1      phiH.2      phiH.3      phiH.4
##  0.3486150 -0.5903335 -1.0318488 -2.4053812 -4.8052422
##
## Threshold:
## -Variable: Z(t) = + (1) X(t)+ (0)X(t-1)+ (0)X(t-2)+ (0)X(t-3)
## -Value: 0.03798
## Proportion of points in low regime: 84.92%    High regime: 15.08%
##
## Residuals:
##      Min      1Q    Median      3Q      Max
## -0.277686 -0.020554  0.005116  0.020762  0.119627
##
## Fit:
## residuals variance = 0.002405,  AIC = -762, MAPE = 372.5%
##
## Coefficient(s):
##
##      Estimate  Std. Error  t value  Pr(>|t|)
## const.L -0.000062876  0.005572290  -0.0113  0.9910158
## phiL.1  0.446407513  0.089159422   5.0068  1.921e-06 ***
## phiL.2 -0.231588785  0.064328877  -3.6001  0.0004636 ***
## phiL.3  0.107014090  0.061034049   1.7534  0.0820953 .
## phiL.4 -0.144060859  0.055198346  -2.6099  0.0102109 *
## const.H  0.348615027  0.027460007  12.6954 < 2.2e-16 ***
## phiH.1 -0.590333531  0.109881869  -5.3724  3.869e-07 ***
## phiH.2 -1.031848767  0.149570541  -6.8987  2.641e-10 ***
## phiH.3 -2.405381238  0.691344783  -3.4793  0.0007014 ***
## phiH.4 -4.805242210  0.828241480  -5.8017  5.451e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

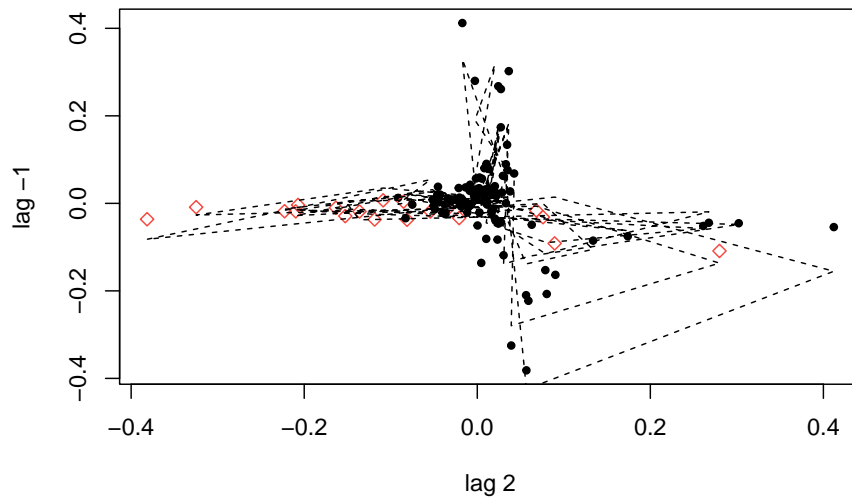
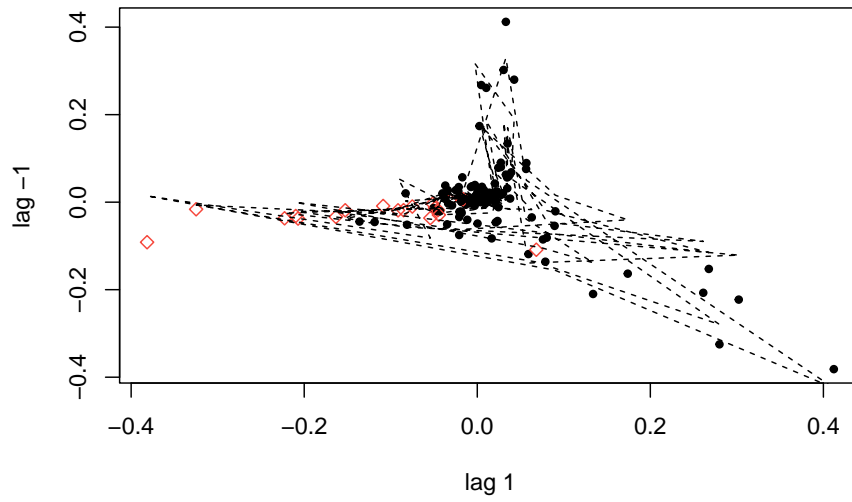
```
##  
## Threshold  
## Variable:  $Z(t) = + (1) X(t) + (0) X(t-1) + (0) X(t-2) + (0) X(t-3)$   
##  
## Value: 0.03798
```

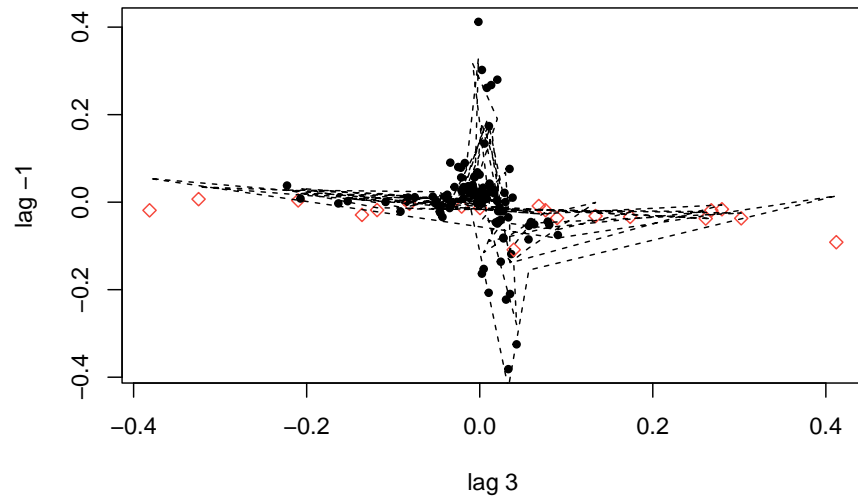
```
plot(D_flu_tar4)
```



ACF of D_flu**ACF of Residuals****PACF of D_flu****PACF of Residuals**







Regime switching plot

