

ANÁLISIS Y PREDICCIÓN DE LOS PRECIOS DEL COBRE Y DEL DÓLAR EN CHILE UTILIZANDO UN MODELO VARMA

Manuel Alejandro González Vargas

Supervisor:
Dra. Orietta Nicolis

Trabajo de titulación para optar para el título de:

Ingeniero Estadístico

Valparaíso, Chile, 2018

Agradecimientos

En primer lugar agradezco a mi madre por su esfuerzo y colaboración para ser un gran futuro profesional y sobre todo una gran persona, su labor como madre finaliza en esta etapa y solo decir gracias mamá por todo, tu constante esfuerzo y sacrificio da sus frutos. También agradecer a mi hermana que también senti su apoyo en este largo, sufrido y esforzado camino. Agradecer a una persona muy importante en mi vida que llego a mi lado para ser un pilar fundamental, pasamos por muchos procesos y permanece a mi lado como el primer día de universidad, mi novia, gracias por permanecer a mi lado y caminar juntos en todo lo que nos propongamos. Solo agradecer por su constante apoyo, sé que este título profesional es tan importante para mi como para ustedes. Mención aparte para un gran amigo que me ayudo bastante en este proceso y sobre todo este año, gracias Jose.

No quiero dejar de agradecer a mi profesora guía, Dra. Orietta Nicolis, por brindarme su constante apoyo, motivación y sabiduría en esta área de la estadística.

Resumen

La economía mundial se encuentra en constantes cambios, todo ello impulsado a través de diversos factores que la afectan, dos grandes factores que influyen en la economía a nivel mundial son las transacciones de la divisa norteamericana llamada dólar y el metal rojo llamado cobre. Chile no queda exento ante estos agentes, ya que el cobre y dólar influyen de gran manera en toda nuestra economía, el cobre por que es la principal materia prima que produce el país para exportar y el dolar que es la moneda de cambio más transada a nivel mundial, es por esto que en la economía chilena es de gran interés el comportamiento de los precios que estos tienen, es por esto que en este trabajo se presentan 3 modelos (dos univariados y uno multivariado) de series de tiempo para predecir el precio del cobre y del dólar en Chile. La construcción de los modelos se realiza primero de manera univariada, es decir que se ocupan modelos ARIMA para predecir el precio de cada una de las variables de manera separada y por otra parte se utiliza un modelo vectorial o multivariante de series de tiempo para predecir de forma conjunta cada uno de los precios de las variables de interés (cobre y dólar).

Para la construcción del modelo se utilizó la información dispuesta de manera diaria de los diferentes parámetros (cobre y dólar) desde el 01 de enero del 2006 hasta el 31 de diciembre del 2014, los datos son proporcionados por la bolsa de metales de Londres para el caso del cobre y por le banco central de Chile para el caso del dólar. Se realizaron diferentes análisis con el fin de describir el comportamiento de ambas variables, en donde se probaron distintos modelos con la intención de lograr una predicción confiable para los precios.

Los capítulos de este proyecto describen la teoría que permita comprender la base de los estudios de series temporales para luego aplicar lo teórico con una base sólida y con los conceptos necesarios para realizar un estudio de este tipo, aplicación que se realiza a través sel software R-project.

Introducción

A lo largo de los últimos años Chile se ha convertido en el mayor exportador de Cobre a nivel mundial, es por este motivo que gran parte de los ingresos económicos del país dependen del precio este metal. El precio del cobre se encuentra estrechamente relacionado con el del dólar debido a que esta es la divisa de mayor circulación en el mundo. La formulación de este proyecto de titulación nace debido a las interrogantes que presenta tanto el precio del cobre como el del dólar para el desarrollo de la economía en Chile, por lo tanto se busca diseñar un modelo basado en series de tiempo para analizar y pronosticar ambos precios (cobre y dólar) en Chile. En este trabajo se construirán 3 modelos, en el cual los dos modelos iniciales apuntan a la predicción del cobre y dólar separadamente, y un tercer modelo que nace de los precios de ambos y construir un modelo conjunto.

Hipótesis:

- Se podrá realizar un pronóstico de predicción confiable del precio del cobre y del precio del dólar a través de un modelo de series de tiempo VARMA (vectorial Autorregresivo-Medias Móviles).

Objetivo general:

- Diseñar un modelo de pronóstico para el precio del cobre y dólar en Chile.

Objetivos específicos:

- Realizar un análisis descriptivo del precio del cobre y del precio del dólar.
- Estimar un modelo VARMA a las variables en estudio.
- Aplicar técnicas de diagnósticos.
- Utilizar el modelo estimado para predecir los valores futuros del cobre y dólar.
- Comparar la predicción del modelo VARMA con las predicciones de los modelos univariados para las dos series.

Este proyecto de investigación se fundamenta de la metodología de la series de tiempo y por causa de esto dicho proyecto se compone de los siguientes capítulos:

En el primer capítulo se realiza un introducción a la importancia del cobre y dólar, sea su importancia y efecto que tienen a nivel mundial y la trascendencia que tiene a nivel de economía en Chile.

En el segundo capítulo se estudian los procesos estocástico de series de tiempo, con la importancia de estimar un modelo univariado de predicción. A través de este capítulo se estudian procesos estocásticos como ruido blanco, modelos autorregresivos, modelos media móvil, modelo autorregresivo de media móvil y modelo autorregresivo integrado de media móvil para luego dar alcance a la estimación, validación y predicción de los modelos.

En el tercer capítulo se describen los procesos vectoriales autorregresivo media móvil, este se puede presentar en forma de proceso vectorial autorregresivo, media móvil o simplemente un conjunto de ambos.

El cuarto capítulo es la aplicación de los conceptos estudiados en capítulos anteriores mediante una base de datos construida con precios del cobre y dólar que son obtenidos a través de la bolsa de metales de Londres y banco central de Chile respectivamente.

El quinto capítulo expone las conclusiones finales del trabajo de titulación.

Índice general

Resumen	III
Introducción	IV
1. Cobre y dólar	1
1.1. Cobre	1
1.1.1. Cobre y su origen en Chile	1
1.1.2. Minería del cobre chileno en la actualidad	2
1.1.3. Cobre en el mundo	3
1.2. Dólar	5
1.2.1. Dólar y su importancia	6
1.2.2. Dólar y su relacion con el cobre	6
2. Proceso estocástico de series de tiempo	8
2.1. Definición de proceso estocástico	9
2.2. Proceso estocástico estacionario	10
2.2.1. Proceso ruido blanco	11
2.2.2. Modelo Autorregresivo (AR)	11
2.2.3. Modelo media Móvil (MA)	16
2.2.4. Modelo Autorregresivo de Media Móvil (ARMA)	20
2.3. Proceso estocástico no estacionario	23
2.3.1. Modelo Autorregresivo Integrado de Media Móvil (ARIMA)	28
2.4. Estimación, validación y predicción de un modelo	29
2.4.1. Estimación de parámetros	30
2.4.2. Comparación de modelos.	31
2.4.3. Diagnóstico del modelo.	32
2.4.4. Pronóstico	34
3. Proceso Vector autorregresivo media móvil	36
3.1. Procesos vectoriales estacionarios	36
3.2. Proceso VARMA	37

3.2.1.	Autocorrelaciones y autocovarianzas de un proceso VARMA (p, q)	38
3.2.2.	Implicaciones del modelo VARMA	41
3.2.3.	Estimación de parámetros	42
3.2.4.	Comparación de modelos	43
3.2.5.	Diagnóstico	44
3.2.6.	Pronóstico	44
3.3.	Proceso vectorial autorregresivo	46
3.4.	Proceso vectorial media móvil	47
4.	Análisis de serie de tiempo del cobre y dólar	49
4.1.	Introducción	49
4.2.	Análisis de serie de tiempo univariada del cobre	50
4.2.1.	Datos del cobre	50
4.2.2.	Análisis descriptivo de los datos	51
4.2.3.	Análisis de serie temporal del cobre	53
4.2.4.	Construcción del modelo	56
4.2.5.	Identificación del modelo y estimación de parámetros	58
4.2.6.	Diagnóstico del modelo	62
4.2.7.	Pronóstico	64
4.3.	Análisis de serie de tiempo univariada del dólar	67
4.3.1.	Datos del dólar	67
4.3.2.	Análisis descriptivo de los datos	67
4.3.3.	Análisis de serie temporal del dólar	69
4.3.4.	Construcción del modelo	72
4.3.5.	Identificación del modelo y estimación de parámetros	73
4.3.6.	Diagnóstico del modelo	77
4.3.7.	Pronóstico	79
4.4.	Análisis de serie de tiempo multivariada del cobre y dólar	82
4.4.1.	Datos del cobre y dólar	82
4.4.2.	Análisis de serie temporal de cobre y dólar	82
4.4.3.	Construcción del modelo	84
4.4.4.	Identificación del modelo y estimación de parámetros	85
4.4.5.	Diagnóstico del modelo	88
4.4.6.	Pronóstico	90
5.	Conclusiones	93
	Bibliografía	95

Índice de figuras

1.1. Chuquicamata: Mina abierta mas grande del mundo y es la que produce mayor cobre en Chile [CODELCO]	3
1.2. Productores de cobre a nivel mundial[rankia]	4
1.3. 1 dólar estadounidense [Banco central Chile]	6
2.1. [Ejemplo series de tiempo.] Serie de tiempo de índice de producción minera en Chile (2003-2011). [INE]	9
2.2. Simulación de un proceso ruido blanco gaussiano con media 0 y varianza constante	10
2.3. Simulación de un proceso AR(1) y su función de autocorrelación parcial.	16
2.4. Simulación de un proceso MA(1) y su función de autocorrelación parcial.	20
2.5. Ejemplos de series de tiempo no estacionarias	24
4.1. Representaciones de curvas de asimetría	52
4.2. Representación de curvas de curtosis	52
4.3. Histograma de frecuencia de los valores del cobre	53
4.4. Precios del cobre durante los años 2006-2014 en Chile	53
4.5. Gráfico Q-Q normal de los datos del cobre	54
4.6. Gráfica de dispersión de la primera diferenciación regular del logaritmo del cobre	56
4.7. Función de autocorrelación simple del logaritmo con diferenciación regular del cobre	57
4.8. Función de autocorrelación parcial del logaritmo con diferenciación regular del cobre	57
4.9. Ajuste del modelo ARIMA(3,1,2) a los datos logaritmo del precio del cobre.	61
4.10. Ajuste del modelo ARIMA(1,1,1) a los datos logaritmo del precio del cobre.	61
4.11. Gráfico de dispersión de los residuos del modelo ARIMA(3,1,2)	62
4.12. Grafico de dispersión de los residuos del modelo ARIMA(1,1,1)	63
4.13. ACF de los residuos de modelos ARIMA propuestos	63

4.14. Pronóstico del modelo ARIMA(3,1,2)	64
4.15. Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(3,1,2)	65
4.16. Pronóstico del modelo ARIMA(1,1,1)	65
4.17. Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(1,1,1)	66
4.18. Histograma de frecuencia de los valores del dólar	69
4.19. Precios del dólar durante los años 2006-2014 en Chile	69
4.20. Gráfico Q-Q normal de los datos del dólar	70
4.21. Gráfico de dispersión de la primera diferenciación regular del logaritmo del dólar	72
4.22. Función de autocorrelación simple del logaritmo con diferenciación regular del dólar	73
4.23. Función de autocorrelación parcial del logaritmo con diferenciación regular del dólar	73
4.24. Ajuste del modelo ARIMA(3,1,4) a los datos del logaritmo del dólar .	76
4.25. Ajuste del modelo ARIMA(1,1,1) a los datos del logaritmo del dólar.	76
4.26. Gráfico de dispersión de los residuos del modelo ARIMA(3,1,4) . . .	77
4.27. Gráfico de dispersión de los residuos del modelo ARIMA(0,1,1) . . .	78
4.28. ACF de los residuos de modelos ARIMA propuestos	78
4.29. Pronóstico del modelo ARIMA(0,1,1)	79
4.30. Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(0,1,1)	80
4.31. Pronóstico del modelo ARIMA(3,1,4)	80
4.32. Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(3,1,4)	81
4.33. Precios del dólar y cobre durante los años 2005-2014 en Chile	82
4.34. Gráfico de correlaciones, correlación =-0,53	83
4.35. Gráfico de correlaciones cruzadas	85
4.36. Ajuste del modelo VARMA(1,2) a los datos del logaritmo del precio del cobre	87
4.37. Ajuste del modelo VARMA(1,2) a los datos del logaritmo del precio del dólar	87
4.38. Gráfico de dispersión de residuos modelo VARMA(1,2)	88
4.39. ACF de residuos modelo VARMA(1,2)	89
4.40. Pronóstico del modelo VARMA(1,2) para el cobre	90
4.41. Serie de precios del cobre en contraste con el pronóstico estimado por el modelo VARMA (1,2)	90
4.42. Pronóstico del modelo VARMA(1,2) para el dólar	91

4.43. Serie de precios del dólar en contraste con el pronóstico estimado por el modelo VARMA (1,2)	91
--	----

Índice de tablas

2.1. Transformaciones Box-Cox	25
4.1. Cantidad de datos por año.	50
4.2. Vista de los datos.	51
4.3. Estadísticos descriptivos precio del cobre	51
4.4. Resultados de las pruebas de raíz unitarias del precio cobre	54
4.5. Resultados de las pruebas de raíz unitarias de la diferenciación lo- garítmica del precio cobre	56
4.6. Resultados del ajuste de los modelos ARIMA a los datos del logaritmo del precio del cobre.	59
4.7. Contraste de significación de los parámetros del modelo ARIMA(3,1,2)	60
4.8. Contraste de significación de los parámetros del modelo ARIMA(1,1,1)	60
4.9. Resultado de test Ljung-Box aplicado a los modelos seleccionados . .	64
4.10. Valores de los errores de pronóstico de los modelo ARIMA para el cobre.	66
4.11. Cantidad de datos por año.	67
4.12. Vista de los datos.	68
4.13. Estadísticos descriptivos precio del dólar	68
4.14. Resultados de las pruebas de raíz unitarias del precio dólar	71
4.15. Resultados de las pruebas de raíz unitarias a la diferenciación lo- garítmica del precio dólar	71
4.16. Resultados del ajuste de los modelos ARIMA a los datos del logaritmo del precio del dólar.	74
4.17. Contraste de significación de los parámetros del modelo ARIMA(3,1,4)	75
4.18. Contraste de significación de los parámetros del modelo ARIMA(0,1,1)	75
4.19. Resultado de test Ljung-Box aplicado a los modelos seleccionados . .	79
4.20. Valores de los errores de pronóstico de los modelo ARIMA para el dólar.	81
4.21. Cantidad de datos por año.	82
4.22. Resultados de las pruebas de raíz unitarias a la diferenciación lo- garítmica del precio dólar	83

4.23. Resultados de las pruebas de raíz unitarias de la diferenciación logarítmica del precio cobre	84
4.24. Resultados del ajuste de los modelos VARMA a los datos del logaritmo del precio del cobre y dólar	86
4.25. Contraste de significación de los parámetros del modelo VARMA(1,1)	86
4.26. Resultado de test Ljung-Box aplicado a los residuos del modelo VARMA(1,1)	89
4.27. Valores de los errores de pronóstico de los modelo VARMA para el cobre.	92
4.28. Valores de los errores de pronóstico de los modelo VARMA para el dólar.	92

Capítulo 1

Cobre y dólar

1.1. Cobre

1.1.1. Cobre y su origen en Chile

El cobre es un mineral metálico básico, el cual no es posible encontrarlo en estado de pureza. Dicho metal presenta un color rojizo y una de sus principales características es ser un gran conductor de electricidad. Es por ello que es común la fabricación de cables eléctricos, entre otros componentes electrónicos y eléctricos.

En Chile la principal actividad económica y en la cual el país se sustenta es la minería del cobre. Chile a principios del siglo XX contaba con las mayores reservas mundiales de cobre, esto se debe a que cuenta con la cordillera de los Andes. El primer yacimiento de extracción del llamado metal rojo se estableció en la región del Libertador General Bernardo O'Higgins, en el complejo minero el teniente en el año 1904, la explotación fue iniciada por la compañía estadounidense Braden Copper Co. Luego en el año 1910 se inició la explotación de Chuquibambilla, ubicada en el norte del país, también por una compañía norteamericana llamada Chile Exploration Company. Durante años la explotación del cobre fue realizada por empresas extranjeras, a través de ello las compañías obtenían cuantiosas rentabilidades, debido a que tenían bajos impuestos.

En el año 1971 se produce la nacionalización del cobre en donde el metal pasa a manos del estado, de manera que el gobierno administra e implementa nuevas políticas de regularización. Con esta nacionalización del cobre a lo largo de los años se han producido modificaciones a distintas leyes que permiten un mejor funcionamiento y estabilidad del cobre a través de la administración del estado.

1.1.2. Minería del cobre chileno en la actualidad

En la actualidad la principal empresa que explota el cobre se llama CODELCO, la cual es una empresa estatal que tiene 6 faenas mineras activas. Por otra parte también se tiene a Anglo American Chile que es otro de los grandes productores presentes en Chile, que tiene 5 yacimientos activos a lo largo del territorio nacional. La producción chilena de explotación de cobre representa alrededor del 30% de la producción mundial. Chile también destaca por la producción y explotación de otros metales como hierro, molibdeno, plomo, oro, plata entre otros. Hoy en día el cobre representa alrededor del 3% del PIB de la economía nacional, y sin duda este metal es un pilar fundamental dentro de la economía chilena.

En términos monetarios el cobre para el 2017 bordea los 3US\$/lb y a septiembre del año 2017 las exportaciones de cobre generaron ganancias de alrededor de los 23 mil millones de dólares, cifras que están muy por debajo de lo que fue el mayor pico de ganancias de la década hacia el año 2011 cuando se registraron ganancias de alrededor de 33 mil millones de dólares, eso sí con una diferencia ya que el cobre marcaba un valor cercano a los 4 US\$/lb. Es así como se estima por los expertos que a finales del 2017 y principios del 2018 se registren los mayores precios para el cobre alcanzando la barrera de los 3.5US\$/lb.

En cuanto a la participación del cobre en cuanto a exportaciones al mundo, este representa alrededor del 40% del total. Si bien ha disminuido en los últimos años, en comparación al año anterior estas han aumentado. El mayor punto de exportaciones del metal rojo al mundo se vio reflejado en la primera década del presente siglo, las cuales representaban cerca de un 60%. En la actualidad en Chile se exportan 5.5 millones de toneladas aproximadamente, de ello en Antofagasta, que es la capital del círculo minero presente en Chile, se establece cerca del 60% de la producción cuprífera del país.



Figura 1.1: Chuquibambilla: Mina abierta mas grande del mundo y es la que produce mayor cobre en Chile [CODELCO]

1.1.3. Cobre en el mundo

Chile se sitúa entre las 40 economías con mayores exportaciones, la principal exportación que realiza son minerales, la más destacada es el cobre y a los principales países que se exporta son los ubicados en Asia, países como China, Japón y Corea del Sur, también se tienen otras grandes economías como la italiana, la estadounidense y la alemana.

Para el 2017 el mayor exportador de cobre en el mundo es Chile, seguido de Perú que desplazó al tercer lugar a China quien por años se mantuvo en el segundo lugar.

El precio del cobre a nivel mundial se establece en la bolsa de metales de Londres, entre los principales factores de influencia se puede mencionar que se tiene una menor producción del mineral en la principales faenas del mundo, la faena más grande del mundo se ubica en Chile que tiene por nombre la escondida, la segunda es llamada The Grasberg mine que está en Indonesia y por último se tiene entre las tres más grandes una producción minera que se ubica en Perú y se llama cerro verde.

Otro punto que influye es la variación del valor del dólar, las constantes alzas y bajas condicionan de manera importante el precio del cobre.

La economía de las grandes potencias mundiales son otro factor muy importante en la evolución del precio del cobre a lo largo del tiempo, es por ello que el valor de este metal se ve condicionado por el desempeño económico de China y Estados Unidos a través de su crecimiento, mercado inmobiliario, políticas fiscales y monetarias, tasas de interés, confianza empresarial a la inversión.

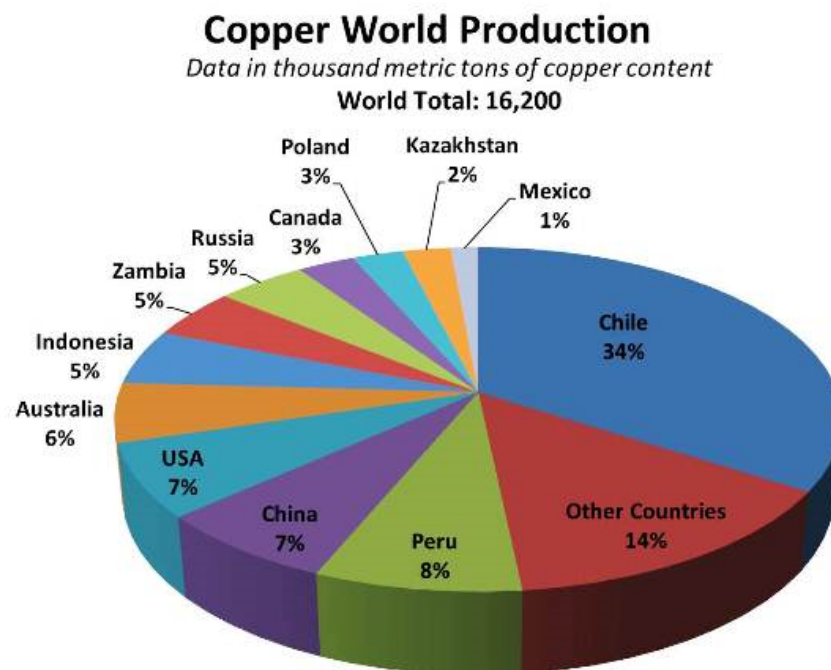


Figura 1.2: Productores de cobre a nivel mundial[rankia]

1.2. Dólar

El dólar estadounidense es la moneda de mayor circulación que existe en el mundo, esto va ligado estrechamente a que se utiliza en la gran mayoría de transacciones a nivel mundial, por ejemplo para la compra y venta de petróleo, oro, inmobiliarios, acciones y como se estudia en el caso de este trabajo, el precio del cobre. El valor de la divisa norteamericana se determina a través del mercado, específicamente mediante la oferta y demanda. En el caso de Chile el precio del dólar está influenciado directamente por las exportaciones e inversión extranjera.

Los principales actores en la oferta de dólares en la economía chilena están determinadas esencialmente por las exportaciones e inversión extranjera, así mismo por el ingreso en divisas producto de los créditos y otro gran factor es la inversión de capitales.

La demanda de dólares en el territorio nacional chileno se determina principalmente por la inversión en suelo extranjero, importaciones, pago de interés y capital de deudas adquiridas.

Para la economía en Chile un factor de peso en el precio del dólar es el valor que adquiere el cobre a través del tiempo. En el tiempo se ha demostrado que existe una relación inversa entre el precio de la divisa norteamericana y el valor del metal rojo, al menos en Chile, cuando sube el precio del cobre ocurre lo contrario con el costo del dólar.

Otros factores que inciden en el valor que adquiere el dólar es la diferencia en la tasa de interés en Chile con respecto a la que está presente en Estados Unidos, es pocas palabras este suceso se puede explicar a que los inversionistas buscan depósitos a plazos, por lo tanto indagan en países donde la tasa de interés sea mayor a la que tienen los estadounidenses, con ello los inversionistas acuden a realizar negocios a estos países y se encuentra con un aumento en la oferta de dólares.

1.2.1. Dólar y su importancia

El dólar es una divisa sumamente importante en el planeta ya que es la moneda de cambio y reserva internacional que se encuentra presente en casi todos los países, se utiliza para transacciones comerciales y la gran mayoría de los países tiene sus reservas en dólares, es posible decir que alrededor del 75% En la actualidad cerca del 85% La importancia del dólar a nivel mundial es muy grande, porque gran parte de negocios, transacciones, entre otros se realizan en dólar. La economía que mueve el dólar es la estadounidense, es por ello que se comenta que un desplome de este país pone en jaque la economía a nivel mundial, se dice que si se cae la economía de Estados Unidos en conjunto con el dólar es como si estallase una bomba en la economía, específicamente en el sistema capitalista.



Figura 1.3: 1 dólar estadounidense [Banco central Chile]

1.2.2. Dólar y su relación con el cobre

Entre el cobre y el dólar existe una relación bastante estrecha. En la economía chilena es común observar que mientras el precio del cobre presenta un alza, el valor de la divisa de mayor circulación en el mundo (dólar) exhibe una baja, es decir, se presenta una correlación inversa entre ambas variables. Cuando se aprecia una variación negativa del dólar permite que se tenga una mayor demanda del cobre, esto provoca que muchas veces haya una sobredemanda en la adquisición del metal y ante lo cual se elevan los precios de la libra de cobre en su venta. Y en el caso de que el dólar sufra variaciones positivas y aumente su valor se tiene que se compre y exporte en menor cantidad el cobre, con lo cual se ve en necesidad de disminuir su precio por libra.

Si bien un alza en el precio de la divisa del dólar favorece las exportaciones nacionales agrícolas, frutícolas y mineras, por otra parte una decadencia en el precio del cobre perjudica de manera directa las arcas fiscales, esto se debe a que gran parte de las utilidades del cobre van en directo aporte al gobierno. Mencionar que por cada centavo de dólar que baje el cobre el fisco percibe alrededor de 50 millones de dólares menos, todo aplicado a que Codelco, la mayor empresa faenera de Chile, aporta menor recaudación de dinero por ende se cobran menos impuestos asociados al metal.

Existe una real preocupación de la economía cuando los valores del cobre y dólar no están equilibrados, ya que la economía chilena se sustenta en gran parte por el valor que tienen mencionado metal y divisa correspondiente.

Capítulo 2

Proceso estocástico de series de tiempo

Una serie de tiempo está dada por un conjunto de datos u observaciones que son medidos en instantes de tiempo, y equidistantes entre sí. El principal objetivo que resalta del análisis de series de tiempo son: realizar predicciones sobre las variables de interés a partir de información pasada y con ello poder planificar, prevenir y controlar diversas situaciones de interés.

Es común encontrar series de tiempo en distintas áreas de estudio. Por ejemplo en el área de la geofísica, demografía, marketing, telecomunicaciones, transporte, meteorología y economía. En áreas como estas podemos encontrar ejemplos claros del uso de series de tiempo, algunos de ellos son: sismos, tasas de natalidad y mortalidad, desempleo, ventas de tiendas, precios de productos, lluvias, contaminación, inflación, crecimiento de un país, valores de divisas mundiales, acciones de las empresas, valores de metales, y así una innumerable cantidad de ejemplos claros en los cuales se aprecia el uso de series de tiempo para la vida cotidiana. Al observar las distintas áreas de estudio se pueden encontrar series de tiempo univariadas en las cuales el análisis se basa en solo una variable de interés, por otra parte se encuentran estudios los cuales se basan en varias variables de interés, es decir, que se complementan observaciones de distintos tipos para analizar una variable final. De manera más formal una serie de tiempo es una secuencia de observaciones registradas en el tiempo de manera ordenada, representada mediante Z_t .

Donde las observaciones son consideradas variables aleatorias denotadas por $Z_1, Z_2, Z_3, \dots, Z_{t-1}, Z_t, Z_{t+1}, \dots$; donde t indica el tiempo en el cual fue tomada la observación ($t \in T \supseteq \mathbb{Z}$). En general, se supone que las series de tiempo son del tipo discreta aun cuando estas sean continuas. Esto quiere decir, que las observaciones son tomadas en periodos de tiempo específicos, por lo tanto, t varía con respecto a los

valores enteros. Por otra parte, los datos de las series de tiempo son equispaciados y equidistantes, donde existe un $h \in \mathbb{R}^+$ tal que t_{i+1} para todo $1 < i < n - 1$ [Chatfield, 2000].

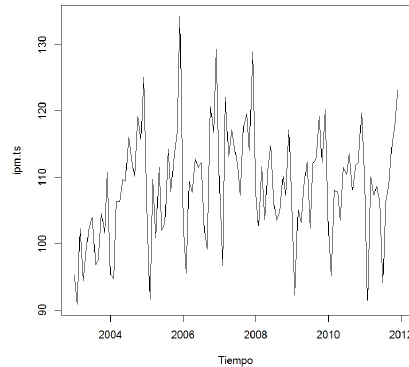


Figura 2.1: [Ejemplo series de tiempo.] Serie de tiempo de índice de producción minera en Chile (2003-2011). [INE]

2.1. Definición de proceso estocástico

Un proceso estocástico es un conjunto de variables aleatorias Z_t donde el índice t toma valores en un cierto conjunto C . Para este caso, este conjunto es ordenado y corresponde a los instantes temporales (días, meses, años, etc.). Para cada valor t del conjunto C (para cada instante temporal) está definida una variable aleatoria, $Z(t)$, y los valores observados de las variables aleatorias en distintos instantes forman una serie temporal. Es decir, una serie de T datos, $(Z_1, \dots, Z_t, \dots, Z_T)$, es una muestra de tamaño uno del vector de T variables aleatorias ordenadas en el tiempo correspondientes a los momentos $t = 1, \dots, T$, y la serie observada se considera una realización o trayectoria del proceso estocástico [Pena, 2010].

Un proceso estocástico también se define a través de sus momentos, como se puede observar a continuación:

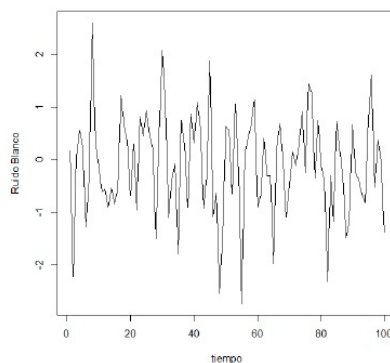


Figura 2.2: Simulación de un proceso ruido blanco gaussiano con media 0 y varianza constante

- Media: $E(Z_t) = \mu_t, \forall t$
- Varianza: $Var(Z_t) = \sigma_t^2, \forall t$
- Autocovarianza: $Cov(Z_t, Z_{t+h}) = E(Z_t - \mu)(Z_{t+h} - \mu) = \gamma_k, \forall t$

Por lo tanto se puede decir que un proceso estocástico se representa como una sucesión de variables aleatorias en el tiempo, independientes entre sí.

2.2. Proceso estocástico estacionario

Al estudiar las series de tiempo es posible apreciar que están se pueden clasificar en dos grandes tipos, una es del tipo no estacionaria, que se define más adelante, y la otra es la estacionaria, que según sus características puede ser estacionaria en sentido estricto o estacionaria en sentido débil.

a) **Estacionariedad en sentido estricto:** Hace mención a aquellas series de tiempo que poseen un comportamiento constante o no se modifican si se trasladan a través del tiempo, es decir:

$$\blacksquare F(Z_{t_1}, Z_{t_2}, \dots, Z_{t_k}) = F(Z_{t_1+h}, Z_{t_2+h}, \dots, Z_{t_k+h})$$

b) **Estacionariedad en sentido débil:** Dicha condición hace evocación e involucra la estabilidad de la media, la varianza y la estructura de covarianzas a través del tiempo, entonces se dice que un proceso es estacionario en sentido si, para todo t :

$$\blacksquare \mu_t = \mu = cte$$

- $\sigma_t^2 - \sigma^2 = cte$
- $\gamma(t, t - k) = E[(Z_t - \mu)(Z_{t-k} - \mu)] = \gamma_k, k = 0, \pm 1, \pm 2$

2.2.1. Proceso ruido blanco

Se tiene que una serie de tiempo $Z_t : t = 1, 2, \dots, n$ es un ruido blanco si sus variables $Z_1, Z_2, Z_3, \dots, Z_n$ son independientes idénticamente distribuidas (iid) con media cero, además de una varianza σ_ε^2 y covarianza para distintos tiempos igual a cero ($cov(Z_t, Z_{t-k}), \forall k \neq t - k$), de manera más formal un proceso ruido blanco es definido por las siguientes condiciones:

- a) $E(Z_t) = 0, t = 1, 2, \dots$
- b) $Var(Z_t) = \sigma_\varepsilon^2, t = 1, 2, \dots$
- c) $Cov(Z_t, Z_{t-kh}) = \sigma^2, k = 0, \pm 1, \pm 2, \dots$

Con esto se puede decir que algunos procesos de series de tiempo temporales siguen un modelo ruido blanco con los parámetros vistos y se denota como:

$$rb(0, \sigma_\varepsilon^2) \tag{2.1}$$

2.2.2. Modelo Autorregresivo (AR)

Un modelo Autorregresivo AR describe una clase particular de proceso que las observaciones en un momento dado son predecibles a partir de las observaciones previas del proceso más un término de error, este proceso es siempre invertible. En otros términos se puede describir un proceso autorregresivo la dependencia de sus valores con respecto a su pasado.

Un proceso Autorregresivo de orden p, es representado por ARMA $(p, 0)$ o simplemente AR (p) .

$$AR(p) = X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \tag{2.2}$$

Donde $\phi_1, \phi_2, \dots, \phi_p$ son parámetros del modelo y $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$.

Otra manera de escribir de forma generalizada el proceso autorregresivo es considerar un operador retardo B , llamado *Backward shift*, el cual tiene como fin retrasar

el índice de tiempo del proceso estocástico, en términos matemáticos se expresa como $(B)X_t = X_{t-1}$ de forma generalizada se define el polinomio autorregresivo como $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$, por lo tanto se puede reescribir un proceso AR (p) de la siguiente manera:

$$\phi_p(B)X_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)X_t = \varepsilon_t \quad (2.3)$$

Quedando de manera generalizada como:

$$\phi_p(B)X_t = \varepsilon_t \quad (2.4)$$

Un proceso autorregresivo de orden p es siempre invertible, y esto se debe a que $\sum_{i=1}^p |\phi_i| < \infty$. Además se debe cumplir que un proceso autorregresivo sea estacionario, y esto se cumple si y solo si todas las raíces del polinomio autorregresivo están dentro del círculo unitario, $(\phi_p B = 0) > 1$. En relación a lo anterior se desprenden y definen las siguientes propiedades:

- *Esperanza* : $E(X_t) = 0$
- *Varianza* : $Var(X_t) = \gamma_0 = \phi_1 \gamma_1 + \dots + \phi_p \gamma_p$
- *Autocovarianza* : $Cov(X_t, X_{t-h}) = \gamma_h = \phi_h \gamma_{h-1} + \dots + \phi_p \gamma_{h-p}, h > 0$

Función de autocorrelación simple (ACF)

En variadas ocasiones las series de tiempo presentan que las variables aleatorias en el tiempo dependen entre sí, es decir que una observación depende de los valores que lo anteceden en el tiempo, la función de autocorrelación entre pares de observaciones de una serie de tiempo que están separadas por h período X_t, X_{t+h} . De una manera más clara, si se tiene una serie de tiempo donde las observaciones están dadas por

$$X_1, X_2, X_3, \dots, X_{t-11}, X_t, X_{t+1}, \dots, \quad (2.5)$$

Siendo así, la función de autocorrelación simple determina la correlación entre dos variables sucesivas en el tiempo, entonces de una manera más clara esta función puntualiza la correlación entre X_1 y X_2 , entre X_2 y X_3 y así paulatinamente hasta X_{t-1} y X_t . La función de autocorrelación queda dada como:

$$\rho_h = cor(X_t, X_{t+h}) = \frac{Cov(X_t, X_{t+h})}{\sigma_{X_t} \sigma_{X_{t+h}}} = \frac{\gamma_x(k)}{\gamma_x(0)} \quad (2.6)$$

Se hace necesario mencionar lo siguiente:

- $\rho_0 = 1$
- Si $\rho_h = 0$, no existe correlación lineal
- Si $\rho_h = 1$, existe correlación lineal
- Si $\rho_h = -1$, existe correlación lineal negativa
- $-1 \leq \rho_h \leq 1$

Otra manera de escribir el ACF:

$$r_h = \frac{\sum_{t=1}^{n-h} (x_t - \bar{x})(x_{t+h} - \bar{x})}{\sum_{t=1}^n (x_t - \bar{x})^2} \quad (2.7)$$

Donde $\bar{x} = \frac{\sum x_i}{n}$ es la media de los datos de la serie de tiempo, x_t es la observación en el periodo t y x_{t+h} es la observación en el período $t + h$; la fórmula Ecuación 2.7 es llamada Correlograma.

Función de autocorrelación parcial (PACF)

Cuando se estudian los procesos autorregresivos y se intenta identificar el orden que sigue el proceso a partir de la función de autocorrelación simple resulta bastante complejo. Para disminuir dicha dificultad se emplea la función de autocorrelación parcial. La función de autocorrelación parcial mide la correlación entre variables pero a diferencia de la ACF no se considera la dependencia que es provocada por la distancia o retardos intermedios que existe entre las variables. En palabras mas claras que X_t influye en X_{t-1} y X_{t-1} incide en X_{t-2} , en su equivalente se dice que X_t afecta en X_{t-2} y así sucesivamente.

El PACF se calcula de la siguiente manera:

- $\phi_{11} = \rho_1$
- $\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$

$$\begin{aligned} \blacksquare \phi_{33} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} \\ \blacksquare \phi_{hh} &= \frac{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_2 \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & \rho_h \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{h-1} \\ \rho_1 & 1 & \cdots & \rho_{h-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & 1 \end{vmatrix}} \end{aligned}$$

Esta ultima es conocida como la función de autocorrelación general.

Proceso autorregresivo de primer orden AR (1)

Se considera un proceso AR (1) como el siguiente:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t$$

Donde $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$. Un proceso autorregresivo de primer orden también se puede escribir como:

$$\phi_p(B)X_t \varepsilon_t$$

Donde $\phi(B) = (1 - \phi B)$, es decir que el proceso autorregresivo queda expresado en funcion del operador retardo, llamado en algunas literaturas como operador *Backward shift*.

Por lo tanto con el proceso AR (1) se puede demostrar que:

- *Esperanza* :
 $E(X_t) = \mu = 0$
- *Varianza* :
 $Var(X_t) = \gamma_0 = E[X - E(X_t)]^2$

$$\begin{aligned}
 &= E[\phi X_{t-1} + \varepsilon_t - 0]^2 \\
 &= \phi^2 E[X_{t-1}]^2 + \sigma_\varepsilon^2
 \end{aligned}$$

Bajo el supuesto que el proceso sigue la condición de Estacionariedad se tiene que

$$E[X_{t-1}]^2 = Var[X_{t-1}] = Var[X_t] = \gamma_0$$

Por lo tanto se tiene:

$$\gamma_0 = \phi^2 \gamma_0 + \sigma_\varepsilon^2$$

Y despejando γ_0 se obtiene:

$$\gamma_0 = \frac{\sigma_\varepsilon^2}{1 - \phi^2}$$

■ Autocovarianza

$$\gamma_h = E[X_t - E(X_t)][X_{t-h} - E(X_{t-h})]$$

Con lo que se sabe que $E[X_t]$ y $E[X_{t-h}]$ son cero,

$$\gamma_h = E[X_t X_{t-h}]$$

$$\gamma_h = E[(\phi X_{t-1} + \varepsilon_t) X_{t-h}]$$

$$\gamma_h = \phi E[X_{t-1} X_{t-h}] + E[\varepsilon_t X_{t-h}]$$

$$\gamma_h = \phi \gamma_{h-1}$$

Por lo tanto se llega a la conclusión que:

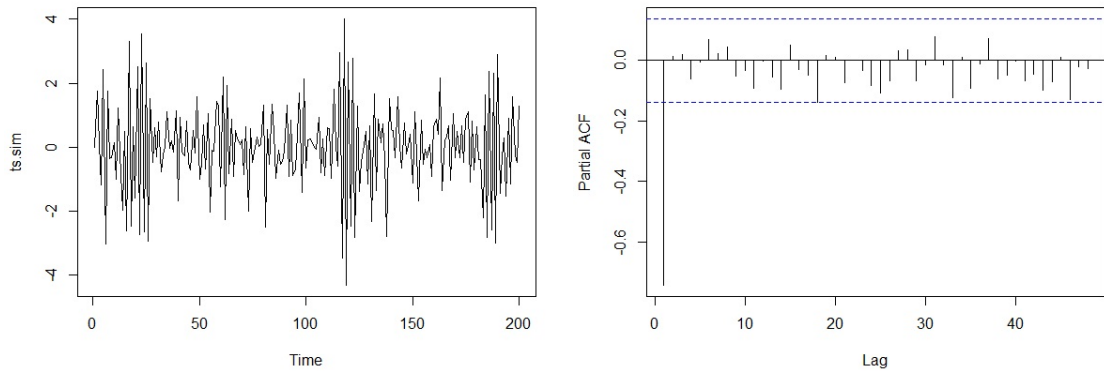
$$\gamma_1 = E[X_t X_{t-1}] = \phi \gamma_0$$

$$\gamma_2 = E[X_t X_{t-2}] = \phi \gamma_1$$

$$\gamma_3 = E[X_t X_{t-3}] = \phi \gamma_2$$

$$\gamma_h = \begin{cases} \phi & h = 1 \\ 0 & h > 1 \end{cases}$$

En la Figura 2.3 se observa la simulación de un proceso AR(1), al observar la PACF resulta fácil identificar el modelo, pero resulta que en algunos casos no es tan inmediata esta identificación, por lo que en la literatura se sugiere usar los criterios de información para hallar el proceso que siguen los datos.



(a) AR(1) con parámetro $\phi_1 = -0,9$

(b) Función de autocorrelación parcial AR(1)

Figura 2.3: Simulación de un proceso AR(1) y su función de autocorrelación parcial.

2.2.3. Modelo media Móvil (MA)

Un modelo de medias móviles MA describe una serie temporal estacionaria, en este modelo el valor actual puede predecirse a partir de la componente aleatoria en este momento y en menor medida de los impulsos aleatorios anteriores.

$$MA(q) = X_t = \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2.8)$$

Donde $\theta_1, \theta_2, \dots, \theta_q$ son parámetros del modelo y $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$.

Otra manera de expresar un proceso de media móvil es a través del operador

retardo, quedando como $\varepsilon_t = \varepsilon_{t-1}$. Se define el polinomio de media móvil de forma generalizada $\theta(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$, por lo tanto se puede reescribir el proceso media móvil como:

$$MA(q) = X_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \varepsilon_t = \theta(B) \varepsilon_t \quad (2.9)$$

Mencionar que el orden q del proceso media móvil depende de la cantidad de los errores anteriores.

Por otra parte se tiene que el proceso de media móvil es siempre estacionario debido a que $1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2 < \infty$. Y también se tiene que el proceso es invertible si y solo si las raíces del polinomio están fuera del círculo unitario, $\theta_q B = \theta > 1$.

Considerando lo anterior se desprenden las siguientes propiedades:

- *Esperanza* : $E(X_t) = 0$
- *Varianza* : $Var(X_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2) \sigma_\varepsilon^2$
- *Autocovarianza* : $Cov(X_t, X_{t-h})$

$$\gamma_h = \begin{cases} (-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{q-h}) \sigma_\varepsilon^2 & \text{si } h = 1, 2, \dots, q \\ 0 & \text{si } h > q \end{cases}$$

Función de autocorrelación (ACF)

$$\rho_h = \begin{cases} \frac{(-\theta_h + \theta_1 \theta_{h+1} + \dots + \theta_q \theta_{q-h})}{(1 + \theta_1^2 + \dots + \theta_q^2)} & \text{si } h = 1, 2, \dots, q \\ 0 & \text{si } h > q \end{cases}$$

Función de autocorrelación parcial (PACF)

$$\phi_{hh} = \frac{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_1 \\ \rho_1 & 1 & \cdots & \rho_2 \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & \rho_h \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \cdots & \rho_{h-1} \\ \rho_1 & 1 & \cdots & \rho_{h-2} \\ \vdots & \vdots & \cdots & \vdots \\ \rho_{h-1} & \rho_{h-2} & \cdots & 1 \end{vmatrix}} \quad (2.10)$$

Proceso media móvil de primer orden MA (1)

Se considera un proceso MA (1) como el siguiente:

$$X_t = \varepsilon_t - \theta_1 \varepsilon_{t-1}$$

Donde $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$. Un proceso media móvil de primer orden también se puede escribir como a través del operador retardo $(B)\varepsilon_t = \varepsilon_{t-1}$

$$X_t = (1 - \theta_1 B)\varepsilon_t$$

Como se mencionó anteriormente en la definición de un proceso media móvil que dicho proceso es siempre estacionario, no así en el caso de la invertibilidad.

Para comprobar la condición de invertibilidad se tiene lo siguiente:

$$1 - \theta_1 B = 0$$

$$\theta_1 B = 1$$

$$B = \frac{1}{\theta_1}$$

Por lo tanto para que se cumpla la condición de invertibilidad $B = \frac{1}{\theta_1} > 1$. Y con respecto a lo anterior se desprenden las siguientes propiedades:

- *Esperanza* :

$$E(X_t) = 0$$

- *Varianza* :

$$\text{Var}(X_t) = \gamma_0 = (1 + \theta_1^2)\sigma_\varepsilon^2$$

- *Autocovarianza* :

$$\text{Cov}(X_t, X_{t-h}) = \gamma_h = \begin{cases} \gamma_0 & \text{si } h = 0 \\ -\theta_1\sigma_\varepsilon^2 & \text{si } h = 1 \\ 0 & \text{si } h > 1 \end{cases}$$

Función de autocorrelación simple (ACF):

$$\rho_h = \begin{cases} \frac{-\theta_1}{1+\theta_1^2} & \text{si } h = 1 \\ -0 & \text{si } h > 1 \end{cases}$$

En este caso se tiene que solo la primera autocorrelación es distinta de cero.

Función de autocorrelación parcial (PACF):

$$\phi_{hh} = \frac{-\theta_1^h(1 - \theta_1^2)}{1 - \theta_1^{2(h+1)}} \quad h > 0$$

En este caso se tiene que las autocorrelación decaen de manera exponencial lentamente hacia cero.

En la Figura 2.4 se observa la simulación de un proceso $MA(1)$, al observar la ACF resulta fácil identificar el modelo, pero en algunos casos no es tan inmediata esta identificación, por lo que en la literatura se sugiere usar los criterios de información para hallar el proceso que siguen los datos.

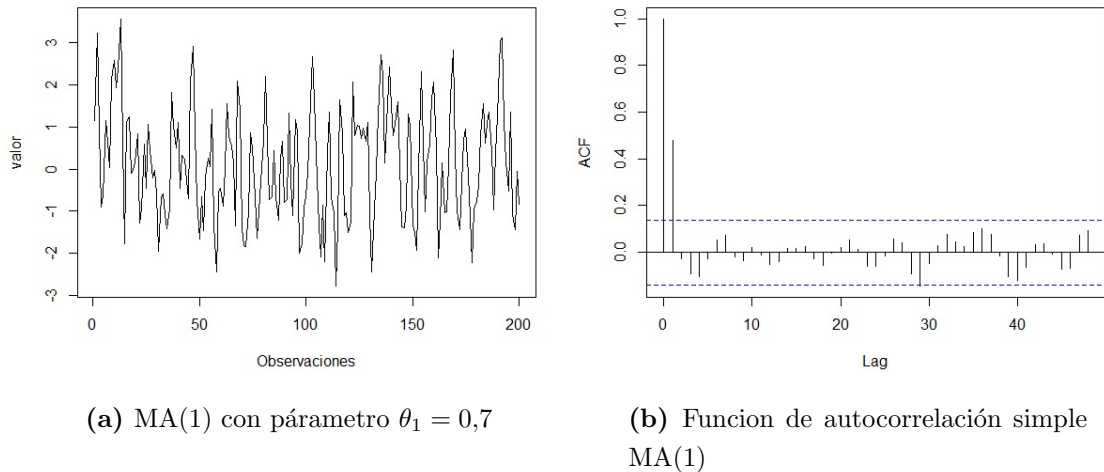


Figura 2.4: Simulación de un proceso $MA(1)$ y su función de autocorrelación parcial.

2.2.4. Modelo Autorregresivo de Media Móvil (ARMA)

La combinación de los procesos $AR(p)$ y $MA(q)$ da lugar a los procesos mixtos $ARMA(p, q)$, este modelo incluye tantos términos regresivos como medias móviles. Este proceso es estacionario si lo es su componente autorregresiva, y es invertible si lo es su componente de medias móviles.

$$ARMA(p, q) = X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2.11)$$

Donde $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ son parámetros del modelo y $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$.

Como en este modelo se mezclan dos procesos como lo son procesos autorregresivos y de medias móviles por lo tanto los pronósticos van a depender de los valores del pasado de X_t , así como de los valores pasados y actuales de los errores.

Si se tiene en consideración que un modelo $ARMA(p, q)$ puede ser expresado en función del operador *Backward shift* denotado con la letra B y considerado como $(B)X_t = X_{t-1}$, la parte autorregresiva del modelo queda expresada de la siguiente

forma:

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \quad (2.12)$$

Para el caso de la parte media móvil se considera el operador retardo como $(B)\varepsilon_t = \varepsilon_{t-1}$, por lo tanto queda expresado de la siguiente manera:

$$\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) \quad (2.13)$$

En tanto el proceso ARMA se puede expresar como:

$$\phi_p(B)X_t = \theta_q(B)\varepsilon_t \quad (2.14)$$

Destacar que la condición de Estacionariedad de un proceso ARMA (p, q) está dada por la parte autorregresiva, puesto que la parte media móvil es siempre estacionaria. Por ende el proceso ARMA es estacionario si y solo si las raíces del polinomio autorregresivo se encuentran fuera del círculo unitario, $\phi_p(B) = 0$.

Caso contrario ocurre con la condición de invertibilidad que está dada por la parte media móvil, ya que según definición un proceso autorregresivo es siempre invertible, por lo tanto se cumple esta condición si y solo si las raíces del polinomio unitario media móvil se encuentran fuera del círculo unitario $\theta_q(B) = 0$.

Para tener una visión más generalizada de un proceso ARMA, a continuación se realiza un modelo ARMA(1, 1), resaltar que un modelo ARMA(1, 0) en un proceso autorregresivo de primer orden y que un modelo ARMA(0, 1) es un proceso media móvil de primer orden.

Proceso ARMA(1,1)

Un proceso ARMA(1, 1) se expresa de la siguiente manera:

$$ARMA(1, 1) = X_t = \phi_1 X_{t-1} + \theta_1 \varepsilon_{t-1} + \varepsilon_t$$

Donde $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$.

Destacar que el proceso ARMA(1, 1) cumple con la condición de estacionariedad si $|\phi| < 1$ y la condición de invertibilidad viene dada si se cumple que $|\theta| < 1$. Por lo tanto las propiedades de un proceso ARMA(1, 1) son:

- *Esperanza* :

$$E(X_t) = 0$$

- *Varianza* :

$$\text{Var}(X_t) = \gamma_0 = \frac{(1 + \theta^2 - 2\theta\phi)}{(1 + \phi^2)} \sigma_\varepsilon^2$$

- *Autocovarianza* :

$$\text{Cov}(X_t, X_{t-h}) = \gamma_h = \begin{cases} \gamma_0 & \text{si } h = 0 \\ \frac{(\phi - \theta)(1 - \phi\theta)}{(1 + \theta^2)} \sigma_\varepsilon^2 & \text{si } h = 1 \\ \theta\gamma_{(h-1)} & \text{si } h > 1 \end{cases}$$

Función de autocorrelación:

$$\rho_h = \begin{cases} 1 & \text{si } h = 0 \\ \frac{(\phi - \theta)(1 - \phi\theta)}{(1 + \theta^2 - 2\phi\theta)} \sigma_\varepsilon^2 & \text{si } h = 1 \\ \phi\rho_{(h-1)} & \text{si } h > 1 \end{cases}$$

Mencionar que las características de ACF y PACF poseen atributos tanto del proceso MA como AR, por lo tanto en la literatura se observa que el comportamiento del ACF decrece de manera exponencial y dependiendo directamente del proceso autorregresivo. Para el caso del PACF también decae de forma exponencial pero dependiendo directamente del proceso media móvil.

2.3. Proceso estocástico no estacionario

Hasta el momento se han visto modelos que siguen un proceso estacionario, pero es común que un proceso sea no estacionario, es decir que la media y la varianza no sean constantes, o las autocorrelaciones dependan del tiempo.

Cuando el nivel de la serie no es estable en el tiempo, pudiendo en particular tener tendencia creciente o decreciente, se dice que la serie es no estacionaria en media. Cuando la variabilidad o las autocorrelaciones se modifican con el tiempo, se dice que la serie es no estacionaria en la varianza o autocorrelaciones. Finalmente, si la distribución de la variable en cada instante varía con el tiempo, se dice que la serie es no estacionaria en distribución. Los procesos no estacionarios más importantes son los procesos integrados, que tienen la propiedad fundamental de que al diferenciarlos se obtienen procesos estacionarios. Una propiedad importante que diferencia a los procesos integrados de los estacionarios es la forma en que desaparece la dependencia con el tiempo. En los procesos estacionarios ARMA las autocorrelaciones disminuyen geoméricamente y se hacen cero a los pocos retardos. En los procesos integrados las autocorrelaciones disminuyen linealmente con el tiempo y es posible encontrar coeficientes de autocorrelación distintos de cero hasta para retardos muy altos [Pena, 2010].

Para determinar si la serie es estacionaria o no estacionaria existen pruebas como la de Dickey-Fuller aumentada (ADF), Kwiatkowski-Phillips-Schmidt-Shin (KPSS), Phillips-Perron (PP), todas ellas basándose en la teoría de la presencia de raíz unitaria, teoría que se aborda en profundidad más adelante. Otro método para determinar la estacionariedad de los datos es mediante el gráfico de datos. En variadas ocasiones se aprecia que las series de tiempo son no estacionarias, los datos pueden ser no estacionaria en media, en varianza o autocorrelación. Para eliminar la tendencia de los datos se estabiliza la media mediante la diferenciación, y para el caso de estabilizar la dispersión de los datos, ayudar a que la varianza sea constante, se utiliza el método de la transformación de Box-Cox.

No estacionaria en media

Una serie de tiempo que es no estacionaria en media es más fácil de identificar mediante el gráfico de los datos, se dice que una serie de tiempo tiene un comportamiento de ser no estacionaria en media cuando se tiene la presencia de una tendencia determinista. Cuando la serie de tiempo presenta no estacionariedad en media es posible realizar la transformación de la serie de tiempo a través del método de la

diferenciación que se define como:

$$\nabla^d = X_t = X_t - X_{t-1} = (1 - B)X_t \quad (2.15)$$

Donde el parámetro d indica la cantidad de diferenciaciones que se realizan. Cuando $d = 1$, hace referencia a la primera diferenciación y con ello se elimina la tendencia lineal. Cuando $d = 2$, se refiere a la segunda diferenciación y se elimina una tendencia cuadrática y así sucesivamente. El operador retardo se asocia a B .

No estacionaria en varianza

Una serie de tiempo que es no estacionaria en varianza presenta oscilaciones o dispersión de los datos, a diferencia de la no estacionariedad en media esta se hace un poco más complicado determinarla mediante gráficos, por lo tanto se realizan pruebas para determinar la presencia o no de estacionariedad en la varianza, dichas pruebas son Dickey-Fuller aumentada (ADF), Kwiatkowski–Phillips–Schmidt–Shin (KPSS), Phillips-Perron (PP), entre las más conocidas.

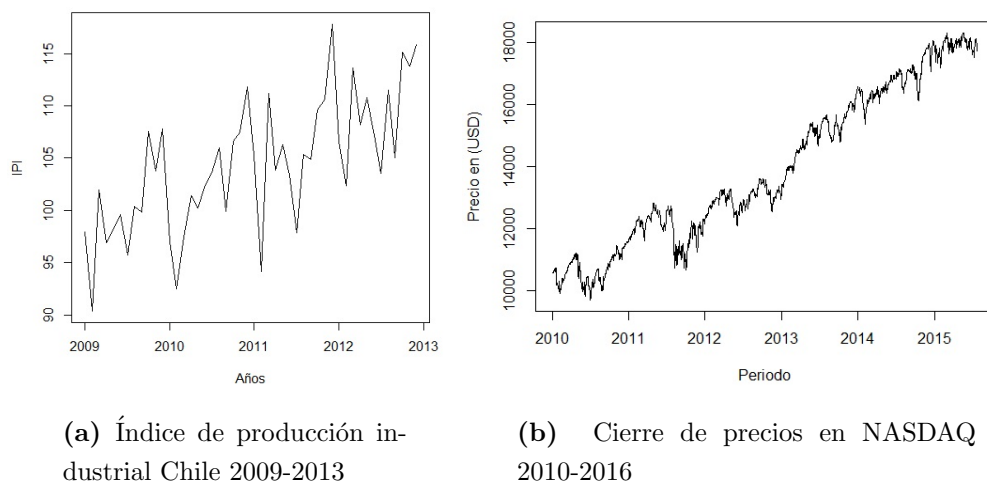


Figura 2.5: Ejemplos de series de tiempo no estacionarias

Cuando se encuentra una serie de tiempo no estacionaria en varianza es posible estabilizar la alta variabilidad mediante la transformación de Box-Cox que está dada por:

$$T_\lambda(X_t) \begin{cases} \frac{X_t^\lambda - 1}{\lambda} & \text{si } \lambda \neq 0 \\ \ln(X_t) & \text{si } \lambda = 0 \end{cases} \quad (2.16)$$

Se tiene que λ es el parámetro de transformación, donde los valores más utilizados se presentan a continuación:

λ	Transformación
-1	$\frac{1}{X_t}$
-0,5	$\frac{1}{\sqrt{X_t}}$
0	$\ln(X_t)$
0,5	$\sqrt{X_t}$
1	X_t

Tabla 2.1: Transformaciones Box-Cox

Para utilizar la transformación de Box-Cox los valores de la serie de tiempo deben ser positivos. Mencionar que si se desea utilizar una transformación de Box-Cox y una diferenciación de los datos en simultáneo, se debe aplicar la transformación Box-Cox primero y después la diferenciación.

Test de estacionariedad

Una manera de observar si la serie de tiempo es estacionaria es a través de los gráficos, pero cuando se hace difícil identificar que la serie de tiempo sea de tipo estacionaria, en la literatura resalta un procedimiento de manera formal para identificar si una serie de tiempo es estacionaria, es por esto que destacan las pruebas de raíz unitaria, a continuación se destacan los test más utilizados.

Test Dickey-Fuller (DF)

Mediante este test se propone un modelo autorregresivo de primer orden:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t \quad (2.17)$$

Mediante este modelo se definen las hipótesis de interés para determinar la estacionariedad de la serie de tiempo, entonces las hipótesis se definen de la siguiente manera:

$$H_0 : \phi = 1 \text{ (No estacionario)}$$

$$H_1 : \phi < 1 \text{ (Estacionario)}$$

Test Dickey-Fuller aumentado (ADF)

Esta prueba de raíz unitaria es una versión aumentada del *test* Dickey-Fuller, el cual propone el siguiente modelo:

$$\nabla X_t = \gamma + \alpha X_{t-1} + \beta_1 \nabla X_{t-1} + \beta_2 \nabla X_{t-2} + \dots + \beta_{h-1} \nabla X_{t-h+1} + \varepsilon_t \quad (2.18)$$

Donde se define γ como constante y α es el coeficiente de raíz unitaria, a causa

de esto se definen las siguientes hipótesis de interés:

$$H_0 : \alpha = 0 \wedge \gamma = 0 \text{ (No estacionario)}$$

$$H_1 : \alpha < 0 \wedge \gamma \neq 0 \text{ (Estacionario)}$$

Test Kwiatkowski–Phillips–Schmidt–Shin (KPSS)

A través de este *test* se contrasta la hipótesis nula donde la serie de tiempo es estacionaria en torno a una tendencia determinista, por lo tanto el modelo se puede definir como:

$$X_t = \gamma D_t + \mu_t \quad (2.19)$$

Donde:

- $\mu_t = \mu_{t-1} + \varepsilon_t$
- $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$

Además D_t contiene la componente determinística (tendencia). A raíz de este modelo se plantean las siguientes hipótesis de interés:

$$H_0 : \sigma_\varepsilon^2 = 0 \text{ (Estacionaria)}$$

$$H_1 : \sigma_\varepsilon^2 > 0 \text{ (No estacionario)}$$

Test Phillips-Perron (PP)

Este test es una extensión de *test* Dickey-Fuller aumentado. La diferencia se basa en que el *test* de Phillips-Perron supone que los errores se distribuyen de manera dependiente y heterogénea, en cambio el *test* ADF supone que los errores se distribuyen de forma independiente y homogénea. Ante ello se genera que la hipótesis nula (H_0) presenta una raíz unitaria con tendencia (No estacionaria) y la hipótesis alternativa (H_1) no presenta una raíz unitaria (Estacionaria).

En la práctica si al aplicar estos *test* resulta que alguno de ellos dice que la serie de tiempo no es estacionaria se deben aplicar los distintos métodos para lograr la transformación de los datos y lograr que la serie de tiempo sea estacionaria.

2.3.1. Modelo Autorregresivo Integrado de Media Móvil (ARIMA)

Este tipo de modelos se dan cuando no se cumplen las condiciones de estacionariedad, ya sea en sentido estricto o débil, estas series de tiempo se pueden observar de gran forma en procesos económicos, en temas de producción, entre otros.

Cuando se cumple que la serie no es estacionaria se pueden aplicar diversos métodos para transformar la serie de tiempo, es en este caso cuando se da el pie para presentar un modelo autorregresivo integrado media móvil (*ARIMA*) con sus respectivos componentes (p, d, q) . Donde se sabe que p es el componente autorregresivo, q es el componente media móvil, y d es el número de diferenciación que se aplica para que la serie de tiempo sea estacionaria.

Se define un *ARIMA*(p, d, q):

$$ARIMA(p, d, q) = X_t = \phi_1 \nabla^d X_{t-1} + \dots + \phi_p \nabla^d X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (2.20)$$

Donde:

- $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$
- $\nabla^d X_t = X_t - X_{t-d}$
- $d =$ Número de diferenciaciones

El proceso *ARIMA* se puede expresar en función del operador retardo, y queda de la siguiente manera:

$$\phi_p(B)(1 - B)^d X_t = \theta_q(B)\varepsilon_t \quad (2.21)$$

Donde:

- $\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ es el polinomio autorregresivo
- $\theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ es el polinomio media móvil

Además las raíces de las ecuaciones $\phi_p(B) = 0$ y $\theta_q(B) = 0$, están fuera del círculo unitario. Mencionar que cuando $d = 0$ el proceso es estacionario y que cuando $d > 0$ es un proceso no estacionario diferenciado d veces. Por otra parte se tiene que un proceso *ARIMA* puede tener ausencia de componente media móvil o componente autorregresivo.

2.4. Estimación, validación y predicción de un modelo

Metodología de Box-Jenkins

En el mundo del análisis de series de tiempo una de las metodologías más conocidas para realizar un ajuste a los datos de una serie de tiempo es la conocida metodología de Box-Jenkins, llamada así en honor a los autores George Box y Gwilym Jenkins, quienes desarrollaron este método con el fin de realizar predicciones más confiables.

La metodología Box-Jenkins consta de 5 etapas fundamentales para un desarrollo adecuado:

1. Verificar estacionariedad de la serie.
2. Identificar el mejor modelo posible.
3. Estimar el modelo.
4. Diagnosticar si el modelo es el adecuado.
5. Utilizar el modelo para predicciones.

Entonces como primer paso se encuentra verificar estacionariedad de la serie de tiempo, esta se puede realizar mediante la inspección de los gráficos, otro análisis necesario para verificar la estacionariedad de los datos es a través de las pruebas de raíz unitaria. Si la serie de tiempo arroja que no estacionaria se deben convertir los datos a través de métodos propuestos, diferenciación y/o Box-Cox. Una vez que se tiene la serie de tiempo estacionaria, se procede a identificar el mejor modelo posible que se adapte a los datos, esto se puede realizar mediante las funciones de autocorrelación simple (ACF) y autocorrelación parcial (PACF). Algunos de los procesos más conocidos poseen características propias y que son identificables a simple vista mediante el correlograma.

Por ejemplo para un proceso autorregresivo en el ACF se observa que las correlaciones disminuyen exponencialmente hacia cero sin llegar a anularse, en el PACF las correlaciones se anulan para los retardos superiores a p . Para el caso de un proceso media móvil en el ACF se observa que las correlaciones se anulan para los retardos superiores al parámetro q , y en el PACF las correlaciones disminuyen exponencialmente hacia cero sin llegar a anularse.

2.4.1. Estimación de parámetros

El siguiente paso una vez que se tiene identificado el orden del modelo es el que se deben estimar los parámetros del modelo. En este caso se presenta el método más común asociado a la estimación de dichos parámetros, que es el llamado método de máxima verosimilitud. En general se considera un modelo ARMA (p,q), ya que el modelo es estacionario o sino ha sido transformado adecuadamente para que sea estacionario.

Estimador de máxima verosimilitud condicional

Se tiene el modelo general estacionario ARMA (p,q):

$$X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

Donde $\varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$ y la densidad de probabilidad condicional de $\varepsilon = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$ se expresa como:

$$P(\varepsilon | \phi, \mu, \theta, \sigma_\varepsilon^2) = (2\pi\sigma_\varepsilon^2)^{-\frac{n}{2}} \exp \left[-\frac{1}{2\sigma_\varepsilon^2} \sum_{i=1}^n \varepsilon_i^2 \right] \quad (2.22)$$

Lo anterior se reescribe así:

$$\varepsilon_t = \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + X_t - \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} \quad (2.23)$$

Se puede escribir la verosimilitud en función de los parámetros $(\phi, \mu, \theta, \sigma_\varepsilon^2)$.

Se tiene $X = (X_1, X_2, \dots, X_n)'$ y se asume la condición inicial $X_* = (X_{1-p}, \dots, X_{-1}, X_0)'$ y también $\varepsilon_* = (\varepsilon_{1-q}, \dots, \varepsilon_{-1}, \varepsilon_0)$ son conocidas. La función condicional de log verosimilitud es:

$$\ln L_*(\phi, \mu, \theta, \sigma_\varepsilon^2) = -\frac{n}{2} \ln 2\pi\sigma_\varepsilon^2 - \frac{S_*(\phi, \mu, \theta)}{2\sigma_\varepsilon^2} \quad (2.24)$$

Donde

$$S_*(\phi, \mu, \theta) = \sum_{i=1}^n \sigma_\varepsilon^2(\phi, \mu, \theta | X_*, \varepsilon_*, X) \quad (2.25)$$

Es la función condicional de suma de cuadrados. Las cantidades de $\hat{\phi}, \hat{\mu}$ y $\hat{\theta}$ maximizan la función de log-verosimilitud y son los llamados estimadores condicionales de máxima verosimilitud. Porque $\ln L_*(\phi, \mu, \theta, \sigma_\varepsilon^2)$ involucra los datos solo a través de $S_*(\phi, \mu, \theta)$, estos estimadores son los mismos que los estimadores condicionales mínimos cuadrados obtenidos al minimizar la función de suma de cuadrados de

$S_*(\phi, \mu, \theta)$, que se observa, que no contiene el parámetro σ_ε^2 . Luego de obtener los estimadores de máxima verosimilitud, se puede conseguir el estimador de σ_ε^2 usando:

$$\sigma_\varepsilon^2 = \frac{S_*(\phi, \mu, \theta)}{d.f} \quad (2.26)$$

Donde d.f son los grados de libertad, siendo el número de términos usados en la suma de $S_*(\phi, \mu, \theta)$ menos el número de parámetros estimados, entonces $d.f = (n - p) - (p + q + 1) = n - (2p + q + 1)$ [Wei, 2006].

2.4.2. Comparación de modelos.

Una vez que se han identificado los parámetros del modelo de la serie de tiempo temporal, ya sea en un proceso $AR(p)$, $MA(q)$, $ARMA(p, q)$, $ARIMA(p, d, q)$, entre otros se procede a seleccionar el modelo que se ajuste de mejor manera a los datos. Para realizar la comparación de modelos existen criterios de la información, que logran encontrar el modelo más parsimonioso, en los criterios de la información se busca una medida de equilibrio entre un ajuste de los datos y un número mínimo de parámetros, para ello se pueden evaluar los datos a través de dos criterios de información como los son criterio de información Akaike (AIC) o el criterio de información Bayesiano (BIC), destacar que se escoge el modelo que presenta el menor valor de AIC o BIC porque es el que presenta la mejor calidad en el conjunto de modelos evaluados.

Criterio de la información de Akaike (AIC)

El criterio de la información de Akaike, fue iniciado en series temporales por el estadístico japonés Akaike, quien fuera pionero en los métodos de selección de modelos. La idea principal de este criterio es penalizar un exceso de parámetros ajustados. El *AIC* es un estimador muestral de $E[\ln f(X|\theta)]$, esperanza de la log-verosimilitud, y la cual viene dada por la expresión general:

$$AIC(k) = -2\ln\mathcal{L}[\hat{\theta}(k)] + 2k \quad (2.27)$$

Donde $\mathcal{L}[\theta(k)]$ es la función de verosimilitud de las observaciones, $\hat{\theta}(k)$ se asocia a la estimación máximo verosímil de parámetros θ y k es el número de parámetros independientes estimados dentro del modelo.

Criterio de la información Bayesiano (BIC)

En el año 1978 el matemático Schwarz sugiere que el *AIC* puede ser no asintóticamente justificable, por lo tanto presenta un criterio de información alternativo a partir de un enfoque bayesiano, más conocido como BIC. Este criterio penaliza el número de parámetros con \ln (logaritmo neperiano), por lo tanto *BIC* se describe como:

$$BIC(k) = -2\ln\mathcal{L}[\hat{\theta}(k)] + \ln(n)k \quad (2.28)$$

Donde $\mathcal{L}[\theta(k)]$ es la función de verosimilitud de las observaciones, $\hat{\theta}(k)$ se asocia a la estimación máximo verosímil de parámetros θ y k es el número de parámetros independientes estimados dentro del modelo, mientras que n es el tamaño muestral.

2.4.3. Diagnóstico del modelo.

Una vez estimado el modelo de series de tiempo, se deben comprobar ciertos supuestos sobre los residuos del modelo, y afirmar que estos se cumplan con integridad.

Lo primero que se debe comprobar es que los residuos deben tener media cero, varianza constante e independencia de estos.

- Media cero:

Para comprobar que la media de los residuos es cero, se realiza la prueba t-student. A través de la cual se plantean las siguientes hipótesis de interés:

$$H_0 : E(\varepsilon_t) = 0$$

$$H_1 : E(\varepsilon_t) \neq 0$$

- Varianza constante:

Para determinar la homocedasticidad de la varianza, es posible a través del comportamiento del ACF de los residuos al cuadrado. Mediante ello se debe apreciar que no existan autocorrelaciones significativas, y si esto se aprecia entonces se presenta que la varianza es constante u homocedástica. Si se observan autocorrelaciones significativas entonces se tiene la presencia de varianza heterocedástica, con lo cual se debe ajustar un modelo heterocedástico, ya que se tiene presencia de volatilidad de los datos.

- Independencia:

Para corroborar la independencia de los residuos es posible aplicar el test de Ljung-Box, la cual es una prueba estadística que mide que las autocorrelaciones de una serie de tiempo sean distintas de cero. El estadístico de Ljung-Box, se define como:

$$Q(h) = n(n+2) \sum_{i=1}^h \frac{\hat{r}_i^2}{n-i}$$

Donde n es el tamaño de la muestra, \hat{r}_i es la autocorrelación de la muestra en el retraso i , h es el número de retardos, mencionar que este estadístico se distribuye asintóticamente con χ^2 con m grados de libertad. A esto se asocian las siguientes hipótesis de interés:

$$H_0 : \varepsilon_t \sim rb(0, \sigma_\varepsilon^2)$$

$$H_1 : \varepsilon_t \text{ no son ruido blanco}$$

Otra manera de observar la independencia de los residuos es calcular el ACF y PACF sobre estos, y observar que si estos no presentan autocorrelaciones significativas entonces se cumple la independencia entre las observaciones.

Lo segundo que se debe comprobar es que los residuos sigan una distribución normal. Para comprobar la normalidad existen diversos test, en esta ocasión se presentan dos test, los cuales tienen por nombre Shapiro-Wilk y Jarque-Bera.

- Shapiro-Wilk:

Este es un test que contrasta la normalidad de un conjunto de datos, tiene como condición que el tamaño de muestra (n) sea menor o igual a 50. Se plantean las siguientes hipótesis de interés:

$$H_0 : X \sim normal$$

$$H_1 : e.o.c$$

Si el valor p es mayor al nivel de significación se concluye que los residuos presentan una distribución normal.

- Jarque-Bera:

Este test de normalidad contrasta que los residuos poseen la misma asimetría y curtosis de una distribución normal, es decir que tengan valores 0 y 3 respectivamente. Se plantean las siguientes hipótesis de interés:

$$H_0 : X \sim normal$$

$$H_1 : e.o.c.$$

Si el valor p es mayor al nivel de significación se concluye que los residuos presentan una distribución normal.

2.4.4. Pronóstico

El principal y fundamental objetivo de comenzar un análisis de series de tiempo es lograr desembrollar el comportamiento de la información con la finalidad de obtener predicciones hacia un futuro cercano. En todo pronóstico de datos existe un error de pronóstico, el cual se define como:

$$\varepsilon_t = X_{n+T} - \hat{X}_{n+T} \quad (2.29)$$

Esto representa las diferencias entre los datos originales y los valores pronosticados. Para la representación de errores de pronósticos existen diversas formas las cuales se presentan a continuación:

- Cuadrado del error medio (MSE)

$$MSE = \frac{1}{M} \sum_T^M e_T^2 \quad (2.30)$$

Mide el promedio de los errores al cuadrado.

- Raíz del error cuadrático medio (RMSE)

$$RMSE = \sqrt{\frac{1}{M} \sum_T^M e_T^2} \quad (2.31)$$

Mide la raíz del promedio de los errores al cuadrado.

- Error medio porcentual absoluto (MAPE)

$$MAPE = \left(\frac{1}{m} \sum_{T=1}^m \left| \frac{e_T}{X_{n+T}} \right| \right) 100 \quad (2.32)$$

Mide la magnitud de error en términos porcentuales.

- Error medio (ME)

$$ME = \frac{1}{M} \sum_{T=1}^M (\hat{X}_{n+T} - X_{n+T}) \quad (2.33)$$

Mide la desviación promedio entre las observaciones y la predicción.

- Error medio absoluto (MAE)

$$MAE = \frac{1}{M} \sum_{T=1}^M |e_T| \quad (2.34)$$

Mide el promedio absoluto de los errores.

M representa la totalidad de los valores que se utilizan para calcular el error. Es decir que de la totalidad de los datos solo se toman en cuenta una proporción de ellos, esto se produce ya que se realiza una comparación de los datos originales con los datos pronosticados.

Capítulo 3

Proceso Vector autorregresivo media móvil

Hasta ahora se han visto procesos de series de tiempo univariados, pero también es posible observar los datos de una manera multivariante, que consiste en dar una mirada a las observaciones de varias variables y no tan solo de una por si sola y se concentra en la relación de una variable dependiente con un conjunto de variables independientes. Por ejemplo en el área de economía es común el uso de modelos vectoriales de series de tiempo, el simple hecho del precio de un producto se puede ver influenciado por la demanda, precios, inflación, entre otros factores determinantes a la hora de fijar el activo.

En este capítulo se aprecia que la metodología a utilizar es una generalización de los procesos univariantes en series de tiempo y que se pueden utilizar herramientas similares para identificar el modelo, estimarlo y realizar diagnósticos.

3.1. Procesos vectoriales estacionarios

Se tiene un vector de k series temporales $X_t = (x_{1t}, \dots, x_{kt})'$, en la cual se define su esperanza como el vector de esperanzas de cada una de las series del proceso. Se define que un proceso vectorial es estacionario si se cumplen las siguientes condiciones:

- $E(X_t) = \mu$, donde el vector de medias es constante.
- $\Gamma(m) = E[(X_{t-m} - \mu)(X_t - \mu)'] = E[(X_t - \mu)(X_{t-m} - \mu)'] = \gamma_{ij}(m)$, donde sus matrices de autocovarianzas dependen del retardo.

3.2. Proceso VARMA

Una manera general de describir el modelo vectorial autorregresivo media móvil es:

$$\phi(B)X_t = \theta(B)Z_t \quad (3.1)$$

Donde,

- $Z_t \sim rb(0, \Sigma)$.
- $\phi(x) := I - \phi_1 X - \dots - \phi_p X^p$, siendo la parte autorregresiva.
- $\theta(x) := I - \theta_1 X - \dots - \theta_q X^q$, siendo la parte media móvil.
- I es la matriz identidad $m \times m$.
- B es el operador retardo.

Otra manera de encontrar descrito un proceso VARMA en la literatura es como se plantea a continuación:

$$VARMA(p, q) = X_t = \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3.2)$$

Donde $\phi_1, \phi_2, \dots, \phi_p, \theta_1, \theta_2, \dots, \theta_q$ son parámetros del modelo y $\varepsilon_t \sim rb(0, \Sigma)$.

Un proceso VARMA es estacionario e invertible si y solo si se cumple las siguientes condiciones:

$$DET\phi(x) \neq 0; \forall x \text{ tal que } |x| \leq 1 \quad (3.3)$$

$$DET\theta(x) \neq 0; \forall x \text{ tal que } |x| \leq 1 \quad (3.4)$$

Existen casos particulares en los cuales un proceso VARMA se anulan los componentes autorregresivos o componentes medias móviles, es decir, cuando $q = 0$ se dice que la serie temporal X_t sigue un proceso autorregresivo de orden p , $VAR(p)$, y si se da el caso de que $p = 0$, se dice que la serie temporal X_t sigue un proceso medias móviles de orden q , $VMA(q)$.

3.2.1. Autocorrelaciones y autocovarianzas de un proceso VARMA (p, q)

Autocovarianzas

Las autocovarianzas en un proceso vectorial autorregresivo media móvil se calculan de la siguiente manera:

$$\Gamma_x(h) = A\Gamma_x(h-1) \quad h = 1, 2, 3, \dots \quad (3.5)$$

$$\Gamma_x(0) = \text{vec}(\Gamma_x(0)) = (I - A \otimes A)^{-1} \text{vec}(\Sigma_\mu) \quad (3.6)$$

Donde Σ_μ es la matriz de covarianzas del residuo que sigue un proceso ruido blanco.

Lo que es necesario tener en consideración al momento de calcular las autocovarianzas:

- Composición de la matriz A

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

Las dimensiones de la matriz son: $[K(p+q) \times K(p+q)]$

Donde,

$$* A_{11} := \begin{bmatrix} A_1 & \cdots & A_{p-1} & A_p \\ I_k & & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_k & 0 \end{bmatrix}$$

Las dimensiones de la matriz son: $(Kp \times Kp)$

$$* A_{12} := \begin{bmatrix} M_1 & \cdots & M_{q-1} & M_q \\ 0 & \cdots & 0 & 0 \\ \vdots & & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \end{bmatrix}$$

Las dimensiones de la matriz son: $(Kp \times Kq)$

* $A_{21} := 0$

Las dimensiones de la matriz son: $(Kq \times Kp)$

* $A_{22} := \begin{bmatrix} 0 & \cdots & 0 & 0 \\ I_k & \cdots & 0 & 0 \\ & \ddots & & \vdots \\ 0 & \cdots & I_k & 0 \end{bmatrix}$

Las dimensiones de la matriz son: $(Kq \times Kq)$

- Composición del producto Kronecker:

$$A = \begin{bmatrix} 3 & 4 & -1 \\ 2 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 5 & -1 \\ 3 & 3 \end{bmatrix}$$

$$A \otimes B = \begin{bmatrix} 15 & -3 & 20 & -4 & -5 & 1 \\ 9 & 9 & 12 & 12 & -3 & -3 \\ 10 & -2 & 0 & 0 & 0 & 0 \\ 6 & 6 & 0 & 0 & 0 & 0 \end{bmatrix}$$

- Composición del operador *vec*:

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{bmatrix}$$

$$\text{vec}(A') = \begin{bmatrix} \alpha_{11} \\ \alpha_{12} \\ \alpha_{21} \\ \alpha_{22} \\ \alpha_{31} \\ \alpha_{32} \end{bmatrix}$$

Autocorrelaciones

Para calcular las autocorrelaciones en un proceso vectorial autorregresivo media móvil se tiene la siguiente formula:

$$R_x(h) = D^{-1}\Gamma_x(h)D^{-1} \quad (3.7)$$

Donde D es una matriz diagonal con las raíces cuadradas de los elementos de la diagonal de $\Gamma_x(0)$, denotando la covarianza entre $X_{i,t}$ y $X_{j,t-h}$ por $\gamma_{ij}(m)$.

Tal que:

$$D^{-1} = \begin{pmatrix} \frac{1}{\sqrt{\gamma_{11}(0)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sqrt{\gamma_{kk}(0)}} \end{pmatrix} \quad (3.8)$$

En un proceso vectorial de series de tiempo las matrices de autocorrelación se definen de la siguiente manera:

$$R(m) = \rho_{ij}(m) \quad (3.9)$$

Donde los términos de esta matriz están dados por:

$$\rho_{ij}(m) = \frac{\gamma_{ij}(m)}{\sqrt{\gamma_{ii}(0)\gamma_{jj}(0)}} \quad (3.10)$$

3.2.2. Implicaciones del modelo VARMA

En la literatura revisada se hace mención a que los procesos VARMA presentan flexibilidad entre variables. En dichos modelos generalmente se abarcan modelos econométricos y estadísticos de uso común. En esta sección se discutirán las relaciones implícitas entre los componentes del modelo VARMA(p,q).

Causalidad de Granger

”La representación MA de un modelo VARMA sigue siendo una manera fácil de inferir las relaciones dinámicas entre variables. Sin pérdida de generalidad, se asume que $E(z_t) = 0$, y se considera el caso simple de particionamiento $z_t = (x'_t, y'_t)$. Donde $k_1 + k_2 = k$. Para ello se particiona ψ – pesos, de la matriz de polinomios AR y MA, y la innovación $\varepsilon_t = (u'_t, v'_t)$. La representación de MA se escribe como:

$$\begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \psi_{xx}(B) & \psi_{xy}(B) \\ \psi_{yx}(B) & \psi_{yy}(B) \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} \quad (3.11)$$

Similar al caso VAR, si $\psi_{xy}(B) = 0$, pero $\psi_{yx}(B) \neq 0$, entonces existe una relación unidireccional de x_t a y_t . En este caso particular, x_t no depende de ninguna información pasada de y_t , pero y_t depende de algunos valores rezagados de x_t . En consecuencia, x_t causa y_t , en el sentido de causalidad de Granger, y el modelo VARMA implica la existencia de un modelo de función de transferencia lineal.

La condición que $\psi_{xy}(B) = 0$ pero $\psi_{yx}(B) \neq 0$ puede ser expresada en términos de la matriz polinomial de AR y MA de un modelo VARMA(p,q). La representación VARMA es:

$$\begin{bmatrix} \phi_{xx}(B) & \phi_{xy}(B) \\ \phi_{yx}(B) & \phi_{yy}(B) \end{bmatrix} \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} \theta_{xx}(B) & \theta_{xy}(B) \\ \theta_{yx}(B) & \theta_{yy}(B) \end{bmatrix} \begin{bmatrix} u_t \\ v_t \end{bmatrix} \quad (3.12)$$

Usando la definición de $\psi(B)$, se tiene:

$$\begin{bmatrix} \phi_{xx}(B) & \phi_{xy}(B) \\ \phi_{yx}(B) & \phi_{yy}(B) \end{bmatrix}^{-1} \begin{bmatrix} \theta_{xx}(B) & \theta_{xy}(B) \\ \theta_{yx}(B) & \theta_{yy}(B) \end{bmatrix} = \begin{bmatrix} \psi_{xx}(B) & \psi_{xy}(B) \\ \psi_{yx}(B) & \psi_{yy}(B) \end{bmatrix} \quad (3.13)$$

Usando propiedades de la matriz particionada, se tiene:

$$[\phi(B)]^{-1} = \begin{bmatrix} D(B) & -D(B)\phi_{xy}(B)\phi_{yy}^{-1}(B) \\ -\phi_{yy}^{-1}(B)\phi_{yx}(B)D(B) & \phi_{yy}^{-1}(B)[I_{k_2} + \phi_{yx}(B)D(B)\phi_{xy}(B)\phi_{yy}^{-1}(B)] \end{bmatrix}$$

Donde $D(B) = [\phi_{xx}(B) - \phi_{xy}(B)\phi_{yy}^{-1}(B)\phi_{yx}(B)]^{-1}$, en consecuencia, la condición

que $\psi_{xy}(B) = 0$ es equivalente a

$$D(B)[\theta_{xy}(B) - \phi_{xy}(B)\phi_{yy}^{-1}(B)\theta_{yy}(B)] = 0$$

o equivalentemente,

$$\theta_{xy}(B) - \phi_{xy}(B)\phi_{yy}^{-1}(B)\theta_{yy}(B) = 0 \quad (3.14)$$

Usando las mismas tecnicas pero expresando de manera diferente $[\phi(B)]^{-1}$, se obtiene que la condición $\psi_{yx}(B) \neq 0$ es equivalente a

$$\theta_{yx}(B) - \phi_{yx}(B)\phi_{xx}^{-1}(B)\theta_{xx}(B) \neq 0 \quad (3.15)$$

Para la ecuación (3.14), la condición $\phi_{xy}(B) = \theta_{xy}(B) = 0$ que resulta en tener matrices triangulares inferiores simultaneas en los polinomios de la matriz AR y MA de z_t , es solo una condición suficiente para la existencia de causalidad de Granger para x_t a y_t .

Si la ecuación (3.14) contiene a la ecuación (3.15), cambia la desigualdad a la igualdad, entonces x_t y y_t están desacoplados en el sentido de que no están correlacionados dinamicamente.

3.2.3. Estimación de parámetros

En esta sección se estimarán los parámetros del modelo. Para este caso se presenta uno de los métodos mas comunes asociado a la estimación de parámetros, la llamada función condicional de máxima verosimilitud.

Estimador de máxima verosimilitud condicional

”Para un modelo VARMA(p,q) estacionario e invertible, la funcion condicional de verosimilitud de los datos puede ser evaluada recursivamente asumiendo que $\varepsilon_t = 0$ para $t \leq 0$ y $X_t = \bar{x}$ para $t \leq 0$, donde \bar{x} es la media de la muestra de X_t . Mas simple, se asume que $E(X_t) = 0$. Entonces todos los valores de la pre-muestra son 0, que es, $\varepsilon_t = 0 = X_t$ para $t \leq 0$.

Se tiene $[X_t|t = 1, \dots, T]$, se definen dos variables aleatorias de dimensión kT: $X_t = (x'_1, \dots, x'_t)'$ y $A = (\varepsilon'_1, \dots, \varepsilon'_t)'$. Bajo las condiciones que $\varepsilon_t = 0 = X_t$ para $t \leq 0$, se tiene:

$$\phi X = \theta A \quad (3.16)$$

Cuando $kT \times kT$ las matrices ϕ y θ son:

$$\phi = \begin{bmatrix} I_k & 0_k & 0_k & \cdots & 0_k \\ -\phi_1 & I_k & 0_k & \cdots & 0_k \\ -\phi_2 & -\phi_1 & I_k & \cdots & 0_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_k & 0_k & 0_k & \cdots & I_k \end{bmatrix}, \theta = \begin{bmatrix} I_k & 0_k & 0_k & \cdots & 0_k \\ -\theta_1 & I_k & 0_k & \cdots & 0_k \\ -\theta_2 & -\theta_1 & I_k & \cdots & 0_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0_k & 0_k & 0_k & \cdots & I_k \end{bmatrix}$$

Cuando 0_k denota la matriz 0 de $k \times k$. De manera mas precisa, ambas ϕ y θ son matrices de bloques triangulares bajo los bloques diagonales de I_k . Para ϕ los primeros bloques fuera de la diagonal son $-\phi_1$ y los segundos bloques fuera de la diagonal son $-\phi_2$ y asi continua. Todos los j -enésimos bloques fuera de la diagonal son 0 para $j < q$. La matriz θ es definida de manera similar con ϕ_i reemplazando por θ_i y p sustituido por q . La ecuación $\phi X = \theta A$ no es nada mas que el resultado de una expresión recursiva del modelo $\phi(B)X_t = \theta(B)Z_t$ bajo las condiciones previas de la muestra. Si se define $T \times T$ matriz de retardo $\mathbf{L} = [L_{ij}]$, donde $L_{i,i-1} = 1$ para $i = 2, \dots, T$ y 0 en otro caso, entonces se tiene que $\phi = (\mathbf{I}_T \otimes \mathbf{I}_k) - \sum_{i=1}^p (\mathbf{L}^i \otimes \phi_i)$ y $\theta = (\mathbf{I}_T \otimes \mathbf{I}_k) - \sum_{i=1}^q (\mathbf{L}^i \otimes \theta_i)$.

Se mantiene $W = \phi X$. Se tiene $W = \theta A$ o su equivalente $A = \theta^{-1}W$, se tiene en cuenta que θ^{-1} existe por $|\theta| = 1$. En consecuencia, la transformación de A a W tiene una unidad jacobiana. Bajo la suposición de normalidad, se tiene que $A \sim N(0, \mathbf{I}_T \otimes \Sigma_\varepsilon)$. Por lo tanto la función condicional de log-verosimilitud de los datos es:

$$\begin{aligned} l(\beta, \Sigma_\varepsilon : X) &= -\frac{T}{2} \log(|\Sigma_\varepsilon|) - \frac{1}{2} \sum_{t=1}^T (\varepsilon_t' \Sigma_\varepsilon^{-1} \varepsilon_t) \\ &= -\frac{T}{2} \log(|\Sigma_\varepsilon|) - \frac{1}{2} A' (\mathbf{I}_T \otimes \Sigma_\varepsilon^{-1}) A \end{aligned} \quad (3.17)$$

Donde $\beta = \text{vec}[\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q]$ denotan los parámetros del vector AR y MA, y el termino involucrado 2π es omitido.” [Tsay,2004]

3.2.4. Comparación de modelos

Una vez que se han identificado los parámetros del modelo de series de tiempo multivariante, se procede a seleccionar el modelo que mejor se ajuste a la información. Para realizar la comparación de modelos existen criterios de información, al igual que en proceso univariado, estos criterios son de Akaike (AIC) y Bayesiano (BIC), y se seleccionará el modelo que presente un menor valor de AIC y BIC.

Criterio de información de Akaike (AIC)

Como se mencionó en el capítulo dos la idea de este criterio es penalizar un exceso de parámetros ajustados, el estimador muestral AIC para el caso multivariante viene dado por:

$$AIC = -2\ln\mathcal{L}(\phi_1, \dots, \phi_p, \Sigma) + \frac{2(pm^2 + 1)nm}{nm - pm^2 - 2} \quad (3.18)$$

Donde $\mathcal{L}(\phi_1, \dots, \phi_p, \Sigma)$ es la función de verosimilitud del modelo, n es el tamaño de la muestra, m es el número de parámetros y p es el número de variables.

Criterio de información Bayesiano (BIC)

Este criterio penaliza el número de parámetros con \ln (logaritmo neperiano), y el criterio de información bayesiano se describe como:

$$BIC = -2\ln\mathcal{L}(\phi_1, \dots, \phi_p, \Sigma) + \ln(n) \frac{(pm^2 + 1)nm}{nm - pm^2 - 2} \quad (3.19)$$

Donde $\mathcal{L}(\phi_1, \dots, \phi_p, \Sigma)$ es la función de verosimilitud del modelo, n es el tamaño de la muestra, m es el número de parámetros y p es el número de variables.

3.2.5. Diagnóstico

Después de estimar el modelo de series de tiempo multivariado, se comprueban los supuestos sobre los residuos y afirmar que se cumplan con integridad.

Para el diagnóstico del modelo multivariado se debe cumplir que los residuos de cada variables en estudio acaten los siguientes puntos (que son detallados en la sección 2.4.3):

- Media cero
- Varianza constante
- Independencia

Otro supuesto que se debe cumplir es que los residuos presenten normalidad, es necesario mencionar que este supuesto no es tan fuerte y que si no se cumple, se puede continuar con el proceso de encontrar el modelo que mejor ajuste la información.

3.2.6. Pronóstico

Como se mencionó en el capítulo dos el objetivo de analizar series de tiempo es lograr predecir el comportamiento futuro de la información. En la predicción de

datos existe un error de pronóstico, el cual se define como:

$$\varepsilon_t = X_{n+T} - \hat{X}_{n+T} \quad (3.20)$$

Esto representa las diferencias entre los datos originales y lo valores pronosticados. Para la representación de errores de pronósticos existen diversas formas las cuales se presentan a continuación:

- Cuadrado del error medio (MSE)

$$MSE = \frac{1}{M} \sum_T^M e_T^2 \quad (3.21)$$

Mide el promedio de los errores al cuadrado.

- Raíz del error cuadrático medio (RMSE)

$$RMSE = \sqrt{\frac{1}{M} \sum_T^M e_T^2} \quad (3.22)$$

Mide la raíz del promedio de los errores al cuadrado.

- Error medio porcentual absoluto (MAPE)

$$MAPE = \left(\frac{1}{m} \sum_{T=1}^m \left| \frac{e_T}{X_{n+T}} \right| \right) 100 \quad (3.23)$$

Mide la magnitud de error en términos porcentuales.

- Error medio (ME)

$$ME = \frac{1}{M} \sum_{T=1}^M (\hat{X}_{n+T} - X_{n+T}) \quad (3.24)$$

Mide la desviación promedio entre las observaciones y la predicción.

- Error medio absoluto (MAE)

$$MAE = \frac{1}{M} \sum_{T=1}^M |e_T| \quad (3.25)$$

Mide el promedio absoluto de los errores.

M representa la totalidad de los valores que se utilizan para calcular el error.

3.3. Proceso vectorial autorregresivo

En el capítulo dos se ha abordado series de tiempo univariantes, una extensión de las series univariantes son las series de tiempo vectorial o multivariante, es por esto que a continuación se da a conocer el modelo vectorial autorregresivo VAR , la cual aborda de una manera multivariante los procesos autorregresivos. Un proceso vectorial autorregresivo contiene observaciones en las cuales su principal objetivo es predecir a través de observaciones previas.

Un proceso Vectorial Autorregresivo de orden p , es representado por $VARMA(p, 0)$ o simplemente $VAR(p)$. Con lo cual el modelo planteado es:

$$X_t - \mu = \phi_1(X_{t-1} - \mu) + \dots + \phi_p(X_{t-p} - \mu) + \varepsilon_t \quad (3.26)$$

Donde,

- X_t es el vector de k series temporales.
- μ es el vector de medias.
- ϕ_p es el componente autorregresivo del proceso.
- $\varepsilon_t \sim rb(0, \Sigma)$

Por otra parte se puede escribir el proceso VAR (p) en función de los operadores de retardo, por lo tanto se tiene que:

$$\begin{aligned} X_t - \mu &= (\phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) + \varepsilon_t \\ \phi(B) &= X_t - \mu = \varepsilon_t \end{aligned} \quad (3.27)$$

- Donde $\phi(B) = I_k - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$

Un proceso $VAR(p)$ es estable si se cumple que las raíces de la ecuación $|\phi(B)| = 0$ están fuera del círculo unitario, es decir:

$$DET(I_k - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \neq 0; \forall |x| \leq 1 \quad (3.28)$$

Autocorrelaciones de un proceso $VAR(p)$

Para calcular las autocorrelaciones en un proceso vectorial autorregresivo media móvil se tiene la siguiente formula:

$$R_x(h) = D^{-1}\Gamma_x(h)D^{-1} \quad (3.29)$$

Donde D es una matriz diagonal con las raíces cuadradas de los elementos de la diagonal de $\Gamma_x(0)$

3.4. Proceso vectorial media móvil

Al igual que un proceso $VAR(p)$ es una extensión multivariada de un proceso autorregresivo, existe una extensión multivariante de un proceso media móvil, con lo cual un proceso vectorial medias móviles de orden q , es representado por un proceso $VARMA(0, q)$ o simplemente $VMA(q)$. El modelo planteado es:

$$VMA(q) = X_t = \theta_1\varepsilon_{t-1} + \theta_2\varepsilon_{t-2} + \dots + \theta_q\varepsilon_{t-q} + \varepsilon_t \quad (3.30)$$

Donde,

- X_t es el vector de k series temporales.
- θ_q es el componente autorregresivo del proceso.
- $\varepsilon_t \sim rb(0, \Sigma)$

Otra manera de expresar un proceso $VMA(q)$ es a través del operador retardo, quedando como $B\varepsilon_t = \varepsilon_{t-1}$. Se define el polinomio de media móvil de forma generalizada $\theta(B) = (1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q)$, por lo tanto se puede reescribir el proceso vector media móvil como:

$$VMA(q) = X_t = (1 - \theta_1B - \theta_2B^2 - \dots - \theta_qB^q)\varepsilon_t = \theta(B)\varepsilon_t \quad (3.31)$$

Las matrices de autocovarianzas en un proceso media móvil vectorial se calculan como:

$$\Gamma(m) = E[(X_t - \mu)(X_{t+m} - \mu)'] \quad (3.32)$$

Con lo que se obtiene lo siguiente:

$$\Gamma(m) = -\Sigma\theta'_m + \theta_1\Sigma\theta'_{m+1} + \dots + \theta_{q-m}\Sigma\theta'_q, m \leq q \quad (3.33)$$

$$\Gamma(m) = 0, m > q \quad (3.34)$$

A través de esto se llega a la conclusión que un proceso vectorial de medias móviles tendrá q matrices de autocovarianzas distintas de cero.

Capítulo 4

Análisis de serie de tiempo del cobre y dólar

4.1. Introducción

El cobre y dólar son dos pilares fundamentales de la economía, tanto a nivel mundial como local, es por ello que en este capítulo se llevara a cabo un análisis sobre las series temporales del cobre y dólar, ya sea de manera univariada como multivariada, fundamentando el análisis y complementando con lo visto en los capítulos anteriores.

Desde un punto de vista general se busca estudiar los rendimientos del análisis y predicción sobre los precios en ambas variables de interés, todo ello en la economía chilena. Entonces la idea principal esta centrada en hallar un modelo en series de tiempo para el cobre y dólar de manera univariada y otro modelo utilizando las dos variables, esto se realiza a través de un modelo en series de tiempo multivariado o vectorial, siendo la variable dependiente el valor que adquiere el cobre, y por otra parte la variable independiente que es el precio del dólar. En ambos análisis se abordan bajo la metodología de Box-Jenkins, uno de los principales objetivos de este enfoque es modelar una serie temporal que depende en una parte de sus observaciones pasadas y otra parte que es impredecible. Todo el análisis estadístico se realizó mediante el Software R-project, el cual es un programa gratuito y de libre distribución que es posible descargar desde su página <http://r-project.org>.

4.2. Análisis de serie de tiempo univariada del cobre

4.2.1. Datos del cobre

La recopilación de información para la serie de tiempo del cobre se realizó mediante los datos que proporciona la bolsa de metales de Londres (LME) a través de su página web (<http://lme.com>). La importancia en el mundo de esta bolsa de mercados en que se transan metales, que es la más grande e importante a nivel global y en donde se realizan importantes transacciones de metales, como el cobre, oro y plata.

Para el análisis del metal que tiene mayor transacción en el mercado Chile se utilizan 3.648 datos, que corresponden al período comprendido entre el 1 de Enero del 2006 y el 31 de Diciembre del 2014. En la siguiente tabla es posible apreciar como se distribuyen los datos a través de los años, es decir, la cantidad de observaciones y datos faltantes:

Año	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Cantidad de datos	365	365	366	365	365	365	366	365	365	3.287
Datos faltantes	0	0	0	0	0	0	0	0	0	0
Porcentaje (%)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Tabla 4.1: Cantidad de datos por año.

Es posible apreciar y resalta que no se observan datos faltantes, es cierto que la bolsa de metales no realiza transacciones en días Sábado, Domingo y festivos, pero es necesario mencionar que para los casos en que la bolsa de metales Londres tenga cerrado dichos días el valor que adquiere al mundo el cobre se mantiene del día hábil anterior. Por ejemplo el precio que perdura el metal para días Sábado y Domingo es el que se fija el día Viernes en caso de no ser festivo.

4.2.2. Análisis descriptivo de los datos

Al tener la totalidad de los datos y sin observaciones que falten es posible comenzar un completo análisis en el precio del cobre, los datos se almacenan en un archivo Excel en formato CSV y los formatos de fecha están expresadas como dd-mm-aaaa, además de apreciar la fecha se muestra el valor del cobre en ese periodo; el calculo del valor que tiene el metal para cada fecha es obtenido a través del precio que presenta el dolar en dicho instante, de manera mas clara por ejemplo el primer valor del precio cobre (CLP) se obtiene a través de la multiplicación del precio cobre (USD/lb) y dolar (CLP) - recordar que CLP es la abreviación de la moneda local en Chile y USD/lb es el valor por libra del metal expresado en dolares-. El cuadro que se presenta a continuación es posible apreciar una vista de la visualización que tienen los datos y como están ordenados en el archivo Excel:

N°	Fecha	Precio cobre (USD/lb)	Precio dólar (CLP)	Precio cobre (CLP/lb)
1	01-01-2006	2,07	514,21	1064,41
2	02-01-2006	2,07	512,50	1060,88
3	03-01-2006	2,05	513,18	1052,02
4	04-01-2006	2,12	515,45	1098,24
5	05-01-2006	2,11	516,40	1087,24
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
3283	27-12-2014	2.88	609.17	1754.41
3284	28-12-2014	2.88	609.17	1754.41
3285	29-12-2014	2.87	608.15	1745.39
3286	30-12-2014	2.87	607.38	1743.18
3287	31-12-2014	2.88	607.38	1749.25

Tabla 4.2: Vista de los datos.

Una vez definidos los datos cuales se realiza el análisis de series de tiempo para el cobre se procede a realizar un cuadro descriptivo de la variable de interés, tabla que se presenta a continuación:

Variable	Min.	Q_1	Mediana	Media	Q_3	Máx.	Curtosis	Asimetría	Varianza
Cobre	782,00	1587,74	1717,21	1666,04	1817,24	2200,46	4,66	-1,21	106140,14

Tabla 4.3: Estadísticos descriptivos precio del cobre

En el cuadro anterior es posible apreciar que es precio mas bajo en las observaciones utilizadas tiene un valor de 782 pesos por libra de cobre, y alcanzando un máximo de 2200,46 pesos por libra de cobre, así mismo se puede apreciar que en las 3648 observaciones se tiene un precio medio de 1666,04 pesos por libra de cobre. Por otra parte se tiene la asimetría que es -1,21, este valor quiere decir que los valores se tienden a reunir en la parte derecha del valor de la media, para una mayor referencia acerca de la asimetría se presenta la Figura 4.1 en donde su curva es negativa si el valor de asimetría es menor a 0; simétrica si es igual a 0 y positiva si mayor

a 0. La curtosis es de 4,66 con lo cual se dice que la distribución de los datos es Leptocúrtica, a través de ello se puede identificar que gran parte de los valores se encuentran alrededor de la media. En la Figura 4.2 se observan los tipos de curtosis existentes, en donde una alta concentración de valor en torno a un valor central se denomina Leptocúrtica (coeficiente de curtosis mayor a 0); una concentración normal de los datos se llama Mesocúrtica (coeficiente de curtosis igual a 0) y una baja concentración de los datos en torno a un valor central (media) se llama Platicúrtica (coeficiente de curtosis menor a 0)

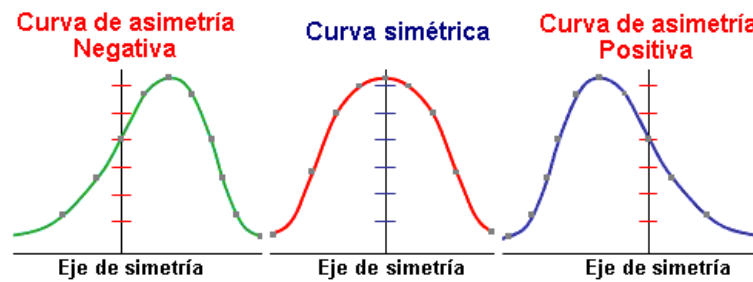


Figura 4.1: Representaciones de curvas de asimetría

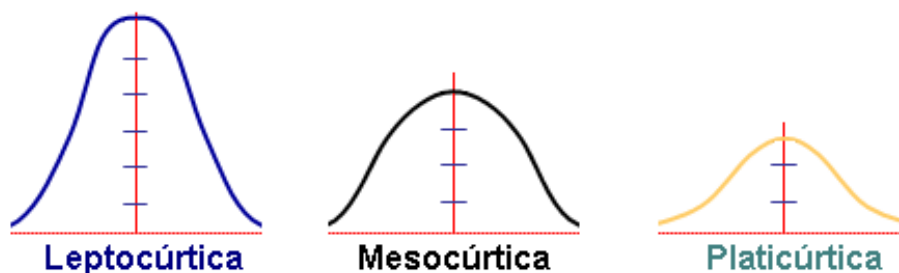


Figura 4.2: Representación de curvas de curtosis

En la Figura 4.3 se observa el histograma de frecuencias de los precios del cobre, a través de esta figura se confirma el valor obtenido en el coeficiente de asimetría, con lo cual es posible observar una curva de asimetría negativa, con ello se concluye que gran cantidad de datos tienen valores altos.

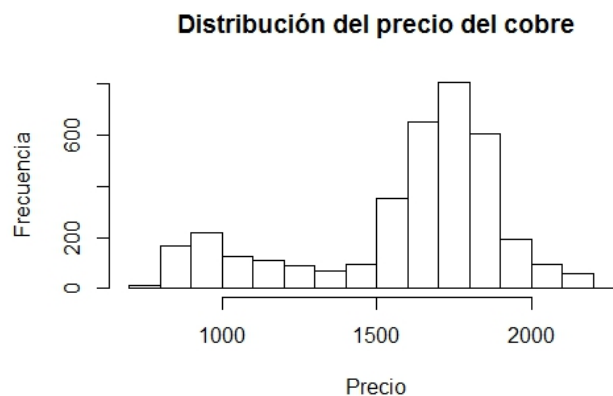


Figura 4.3: Histograma de frecuencia de los valores del cobre

4.2.3. Análisis de serie temporal del cobre

Esta parte del capítulo busca encontrar el mejor modelo que se ajuste a las observaciones del precio del cobre, para después pronosticar los precios del metal rojo, por consiguiente para realizar todo este análisis se aplicara la teoría del capítulo 2 y el paso a paso que conlleva encontrar el modelo mas parsimonioso.

Una etapa fundamental antes de comenzar el análisis de series temporales es graficar los datos, para obtener la serie y observar el comportamiento que presenta a través del tiempo, en la Figura 4.4 se aprecia el comportamiento de los datos.

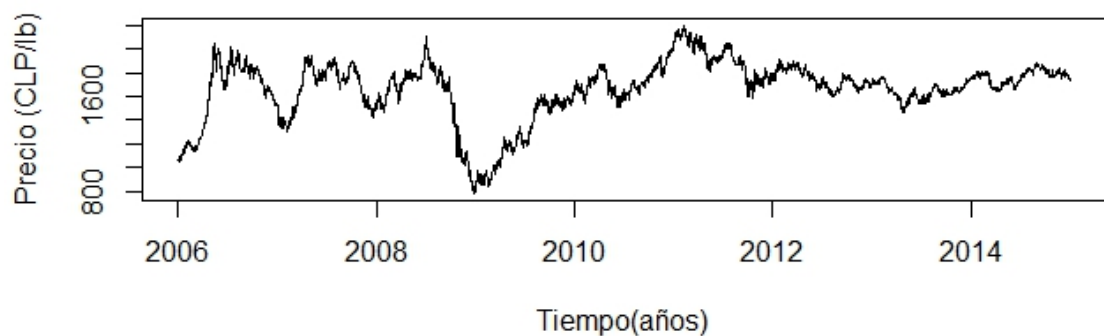


Figura 4.4: Precios del cobre durante los años 2006-2014 en Chile

En la figura anterior se observa que los menores precios alcanzados por el metal fueron en el año 2006 y 2009, para el año 2009 se observa una baja bastante importante, esto se debe a diversos factores presentes en la economía mundial (bajas inversiones de China, crisis inmobiliaria de Estados Unidos). Los valores mas altos

se observan para el año 2011. A través de la Figura 4.4 se aprecia la presencia de tendencia, además se observa que la serie no es estacionaria ni en media ni varianza

En la Figura 4.5 se presenta un gráfico de cuantiles de las observaciones en comparación con los cuantiles teóricos de la distribución normal, dejando en evidencia que los datos no presentan normalidad. Para confirmar esto se aplica el test de Jarque-Bera, el resultado es un p-valor muy pequeño, con lo cual se rechaza la hipótesis nula (normalidad)

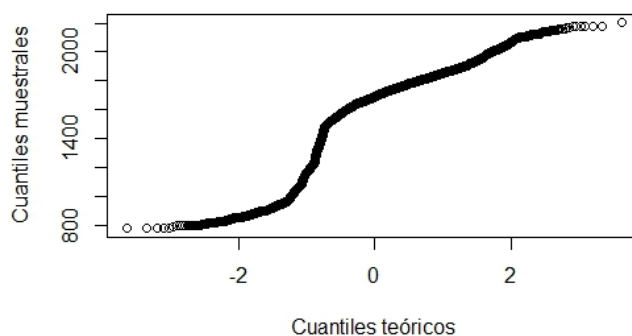


Figura 4.5: Gráfico Q-Q normal de los datos del cobre

Un paso importante en el análisis de series de tiempo es determinar la estacionariedad de los datos, para ello se aplican tres pruebas:

- Test Dickey-Fuller aumentado (ADF)
- Test Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
- Test Phillips-Perron (PP)

Es importante que al aplicar estas pruebas se compruebe que los datos son estacionarios,

Prueba	p-valor	H_0
ADF	0,13	Serie de tiempo no estacionaria
KPSS	0,01	Serie de tiempo estacionaria
PP	0,20	Serie de tiempo no estacionaria

Tabla 4.4: Resultados de las pruebas de raíz unitarias del precio cobre

En la Tabla 4.4 se aprecia que dos de las tres pruebas dicen que la serie no es estacionaria, por lo tanto se debe considerar aplicar alguna transformación, diferenciación o simplemente ambas, debido a que la estacionariedad de los datos se pone en duda. Entonces antes de comenzar con la construcción del modelo se realiza

una transformación logarítmica y posteriormente se realizó una diferenciación de los datos , con la diferenciación se logra eliminar la tendencia de los datos, mientras que con una transformación logarítmica la dispersión de los datos sea menor y los datos se aproximen a una distribución normal.

Con la aplicación de estas transformaciones los datos presentan una menor dispersión y se elimina la tendencia. Con esto se puede continuar para lograr construir un modelo de predicción parsimonioso para los del logaritmo del precio del cobre. Es importante mencionar que para la construcción del modelo, se trabajara con una menor cantidad de datos, todo esto con el fin de facilitar el trabajo del software R-project y disminuir tiempos. Se utiliza el 90% de los datos que corresponden a 2959 registros.

4.2.4. Construcción del modelo

Para efectos de tiempos y mejor uso del software R-project se decide trabajar con el 90 % de los datos, los cuales corresponden a 2959 observaciones (los cuales se llamaran datos de entrenamiento). Recordar a los datos originales se aplicaron una transformación logarítmica y una diferenciación, por lo tanto los datos cumplen la condición de estacionariedad como se muestra a continuación

Prueba	p-valor	H_0
ADF	0,01	Serie de tiempo no estacionaria
KPSS	0,1	Serie de tiempo estacionaria
PP	0,01	Serie de tiempo no estacionaria

Tabla 4.5: Resultados de las pruebas de raíz unitarias de la diferenciación logarítmica del precio cobre

Es importante observar como quedan los datos una vez que se logra obtener que la serie de tiempo sea estacionaria, por lo tanto en la Figura 4.6 se aprecia el comportamiento de los datos a través del tiempo en el cual se ha eliminado la tendencia y dispersión de los datos.

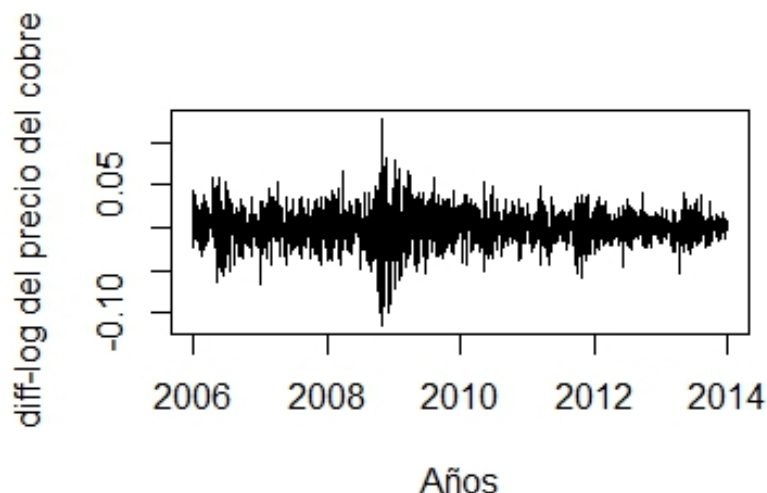


Figura 4.6: Gráfica de dispersión de la primera diferenciación regular del logaritmo del cobre

Cuando se busca identificar un modelo de serie temporal es fundamental determinar los componentes autorregresivo y media móvil que se encuentran presentes en la serie, para identificar los componentes autorregresivos es necesario analizar la

función de autocorrelación parcial de la serie de tiempo por lo tanto se debe identificar aquellos rezagos significativos, esos que están por fuera de la banda de confianza en el correlograma. Por otra parte el componente media móvil de la serie de tiempo se identifica mediante la función de autocorrelación simple también se debe observar los rezagos significativos.

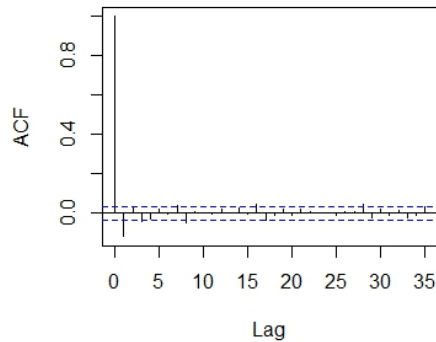


Figura 4.7: Función de autocorrelación simple del logaritmo con diferenciación regular del cobre

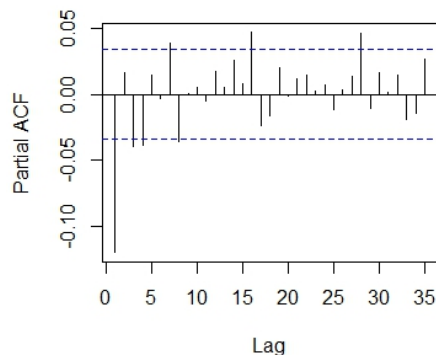


Figura 4.8: Función de autocorrelación parcial del logaritmo con diferenciación regular del cobre

En las figuras anteriores es posible observar las funciones de autocorrelación simple y parcial. "En la teoría y la metodología Box-Jenkins es que a través de estos gráficos se debiese identificar los componentes autorregresivos y de media móvil de la serie, pero en variadas ocasiones se dificulta. La idea principal en series de tiempo es identificar un modelo parsimonioso, es decir un modelo que explique la mayor cantidad de información con una cantidad reducida de parámetros, es por esto que el estadístico H. Akaike plantea una solución utilizando los criterios de la información (AIC). El criterio de la información (AIC) en conjunto al criterio de la información Bayesiana (BIC), permiten identificar el mejor modelo de serie. Por lo tanto para escoger el orden de los componentes del modelo se deben probar con

distintos procesos de orden MA y AR, para luego escoger el modelo que presente el menor valor en los criterios de la información. La diferencia entre el AIC y BIC es que el criterio Bayesiano es más restrictivo que el criterio Akaike, es decir el criterio BIC penaliza con respecto a la cantidad de parámetros en el modelo, cuando la muestra es pequeña este es riguroso, por otra parte el criterio AIC no es tan riguroso en el caso de muestras grandes y otorga una penalización menor por la cantidad de parámetros del modelo” [Shumway & Stoffe, 2010].

4.2.5. Identificación del modelo y estimación de parámetros

Una vez que se tiene una serie temporal la cual es estacionaria se procede a realizar un ajuste de modelos ARIMA. Para llegar a que la serie de tiempo cumpla la condición de estacionariedad se hizo necesario realizar una transformación logarítmica de los datos para lograr disminuir la dispersión de los datos, mediante solo la transformación logarítmica los datos aun siguen presentando no estacionariedad por lo tanto se aplica una diferenciación, a través de ello se busca eliminar la tendencia.

Identificación del modelo ARIMA

Anteriormente se realizaron diversos ajustes de la serie de tiempo, ante estos ajustes las observaciones presentan ideales que son aplicables a datos que siguen un proceso ARIMA.

En la Figura 4.7 y Figura 4.8 es posible apreciar los correlograma de la función de autocorrelación simple y parcial (ACF y PACF) respectivamente. Así mismo en la Figura 4.6 es posible apreciar el gráfico de dispersión de la diferenciación logarítmica del precio del cobre en los años interés para el estudio. Para lograr una correcta identificación del orden del modelo se realizan diversos ajustes a través del software R-project, a través de ello se logran distintos ajustes de modelos ARIMA, y para identificar el modelo más parsimonioso se presentan con sus respectivos valores obtenidos mediante los criterios de información, ya sea Akaike (AIC) o Bayesiano (BIC), y el modelo que presente los menores valores puede ser el modelo que mejor ajuste los datos y por lo tanto el más adecuado.

En la Tabla 4.6 se aprecian los distintos ajustes de los diversos modelos ARIMA con sus respectivos valores del criterio de información Akaike (AIC) y criterio de información Bayesiano (BIC). Como se mencionó anteriormente el modelo que

presente los menores valores de AIC y BIC puede ser un posible modelo que ajuste de mejor manera los datos. Para este caso, el modelo ARIMA (3,1,2) y el modelo ARIMA (1,1,1) son los que contienen el menor valor de los criterios de información Akaike y Bayesiano respectivamente.

Modelo	AIC	BIC
Arima(0,1,1)	-15347,29	-15335,33
Arima(1,1,1)	-15448,92	-15430,98
Arima(1,1,2)	-15446,98	-15423,06
Arima(1,1,3)	-15450,86	-15420,96
Arima(1,1,4)	-15452,79	-15416,91
Arima(2,1,0)	-15448,25	-15430,31
Arima(2,1,1)	-15446,95	-15423,04
Arima(2,1,2)	-15447,10	-15417,21
Arima(2,1,3)	-15453,95	-15418,07
Arima(2,1,4)	-15448,55	-15406,70
Arima(3,1,0)	-15450,17	-15426,25
Arima(3,1,1)	-15449,94	-15420,05
Arima(3,1,2)	-15454,38	-15418,50
Arima(3,1,3)	-15452,42	-15410,57
Arima(3,1,4)	-15450,17	-15402,33
Arima(4,1,0)	-15452,62	-15422,72
Arima(4,1,1)	-15453,96	-15418,09
Arima(4,1,2)	-15452,42	-15410,57
Arima(4,1,3)	-15450,54	-15402,71
Arima(4,1,4)	-15451,16	-15397,35

Tabla 4.6: Resultados del ajuste de los modelos ARIMA a los datos del logaritmo del precio del cobre.

En la Tabla 4.7 y la Tabla 4.8 se dan a conocer los contrastes de significación del modelo ARIMA (3,1,2) y el modelo ARIMA(1,1,1) respectivamente. Para el caso del modelo ARIMA(3,1,2) se aprecia que los seis parámetros del modelo son significativos, esto se debe a que tiene un p-valor mucho mayor que un nivel de significación de $\alpha = 0,5$. Para el caso del modelo ARIMA(1,1,1) se tiene que uno de los parámetros es no significativo.

Parámetro	Estimación	Error estándar	z value	Pr(> z)
ar1	0,335781	0,127052	2,6429	0,0082206
ar2	-0,643892	0,124188	-5,1848	2,162e-07
ar3	-0,132984	0,018722	-7,1029	1,221e-12
ma1	-0,455145	0,127474	-3,5705	0,0003563
ma2	0,718931	0,124391	5,7796	7,488e-09

Tabla 4.7: Contraste de significación de los parámetros del modelo ARIMA(3,1,2)

Parámetro	Estimación	Error estándar	z value	Pr(> z)
ar1	-0,33292	0,12864	-2,5880	0,009653
ma1	0,21520	0,13301	1,6179	0,105684

Tabla 4.8: Contraste de significación de los parámetros del modelo ARIMA(1,1,1)

Lo que se presenta a continuación es el ajuste de los modelos ARIMA(3,1,2) y ARIMA(1,1,1), para lograr tener una mejor visualización de los datos el ajuste se realizó solo a 100 datos para lograr tener una mejor visualización del ajuste del modelo.

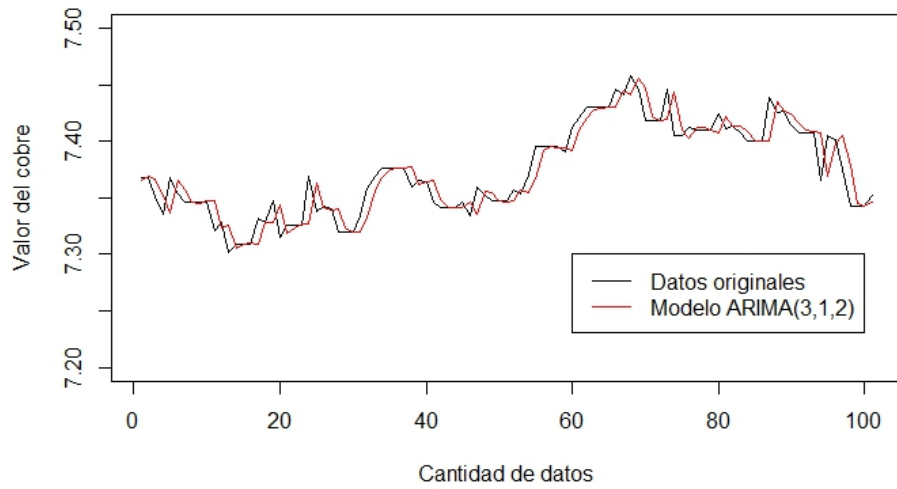


Figura 4.9: Ajuste del modelo ARIMA(3,1,2) a los datos logaritmo del precio del cobre.

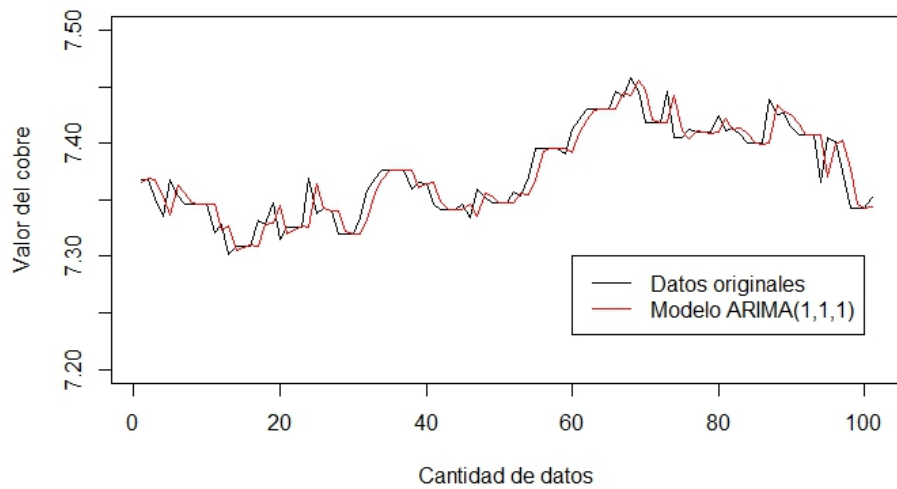


Figura 4.10: Ajuste del modelo ARIMA(1,1,1) a los datos logaritmo del precio del cobre.

Para determinar si estos modelos son los adecuados para representar a los datos se deben analizar los residuos, trabajo que se realiza en la siguiente sección.

4.2.6. Diagnóstico del modelo

Después de seleccionar los posibles mejores modelos se continua con el paso en el cual se debe realizar el método de diagnostico o validación del modelo, esto consiste principalmente en analizar los residuos del modelo. Un modelo es el adecuado cuando se cumple que los residuos se deben comportar como ruido blanco. (Recordar que para que se comporten como ruido blanco la media es igual a cero, una varianza constante y observaciones independientes). Para determinar si la media es cero y la varianza constante se analizara el gráfico de dispersión, por otro lado se observara el correlograma de la función de autocorrelación parcial y simple para la independencia de los residuos y por ultimo mediante la prueba de Ljung-Box se verificará que los residuos sean no autocorrelacionados. Entonces el modelo que cumpla con dichos requisitos se seleccionará como el modelo que represente a los datos de la diferenciación logarítmica del precio del cobre.

En la Figura 4.11 y Figura 4.11 se aprecian los gráficos de dispersión de los residuos de ambos modelos seleccionados y analizados. En ambos gráficos de dispersión es posible apreciar una media y varianza constante, por lo tanto se cumple la primera condición.

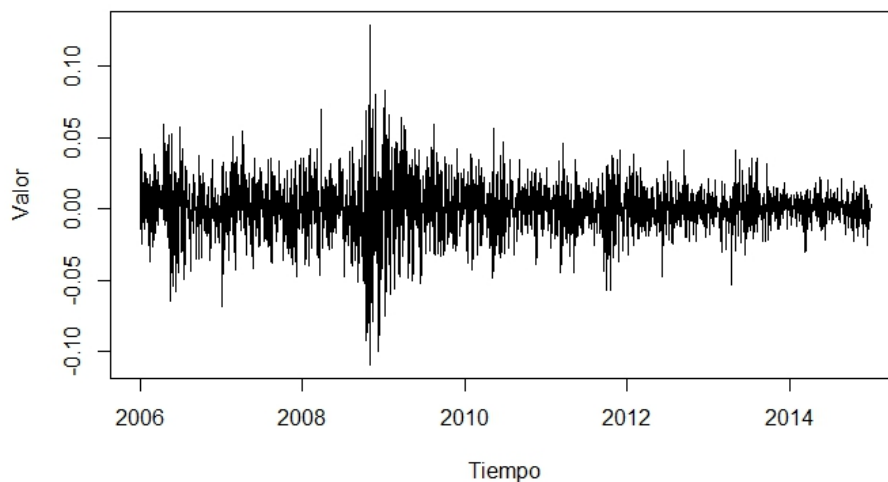


Figura 4.11: Gráfico de dispersión de los residuos del modelo ARIMA(3,1,2)

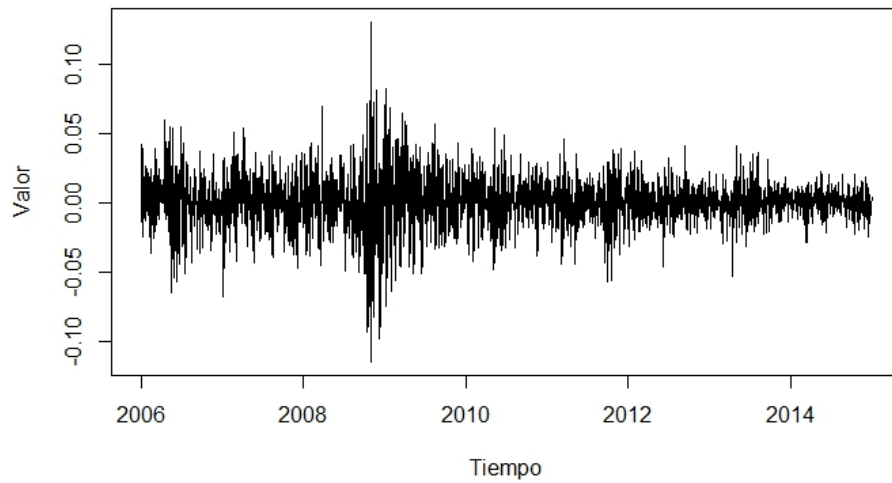


Figura 4.12: Grafico de dispersión de los residuos del modelo ARIMA(1,1,1)

En los gráficos que se presentan a continuación se aprecian las funciones de autocorrelaciones simples de ambos modelos propuestos, es posible apreciar que no se presentan gran cantidad de correlaciones significativas que estén por fuera de las bandas de confianza, por lo tanto los residuos a primera vista son independientes entre sí, es así como no existe autocorrelación en los residuos o dependencia en el tiempo en los residuos.

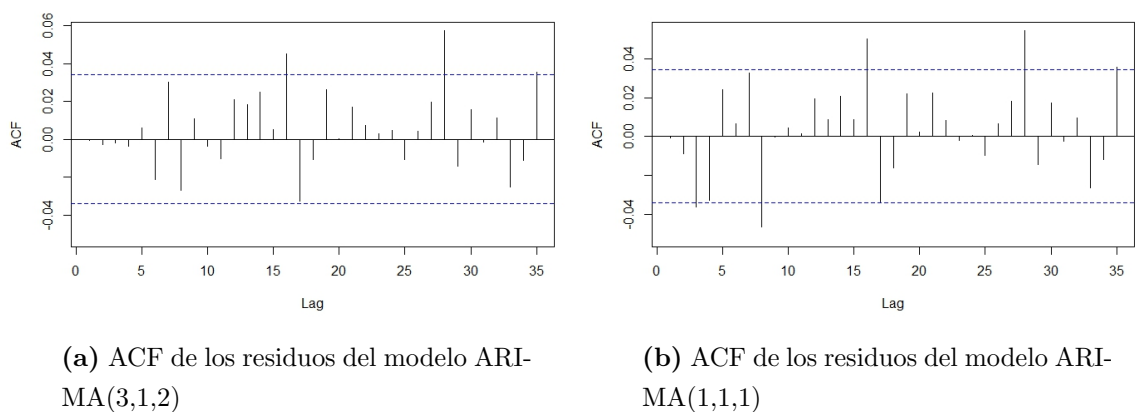


Figura 4.13: ACF de los residuos de modelos ARIMA propuestos

La prueba de Ljung-Box asegura de manera estadística la independencia de los residuos, es por esto que se aplica la prueba a los modelos seleccionados. En la Tabla 4.9 se aprecian los resultados, en donde no se rechaza la hipótesis nula de

Ljung-Box que dice que la distribución de los datos son independientes. Entonces se confirma que los residuos son independientes.

Modelo	p-valor
ARIMA(3,1,2)	0,9731
ARIMA(1,1,1)	0,9664

Tabla 4.9: Resultado de test Ljung-Box aplicado a los modelos seleccionados

Entonces al cumplir con las condiciones y tener los modelos con el menor AIC y BIC, desde ahora se trabajara con estos dos modelos, específicamente el modelo ARIMA(3,1,2) y ARIMA(1,1,1), ya que cumplen con todos los requisitos de los criterios de selección.

4.2.7. Pronóstico

En la sección 2.4 se habló de pronóstico y errores de predicción, en esta parte del capítulo se realiza el pronóstico de los datos, mencionar que se realizan predicciones sobre los modelos ARIMA seleccionados (ARIMA(1,1,1) y ARIMA(3,1,2)). Entonces se procederá a realizar el pronóstico del 10 % de los datos que y se determinará el mejor mediante el análisis de los errores de pronósticos.

En la Figura 4.14 se muestra la imagen donde se aprecia el pronóstico para 327 datos pronosticados del modelo ARIMA(3,1,2) y solo se muestra de la observación 2800 en adelante para apreciar de mejor manera los datos.

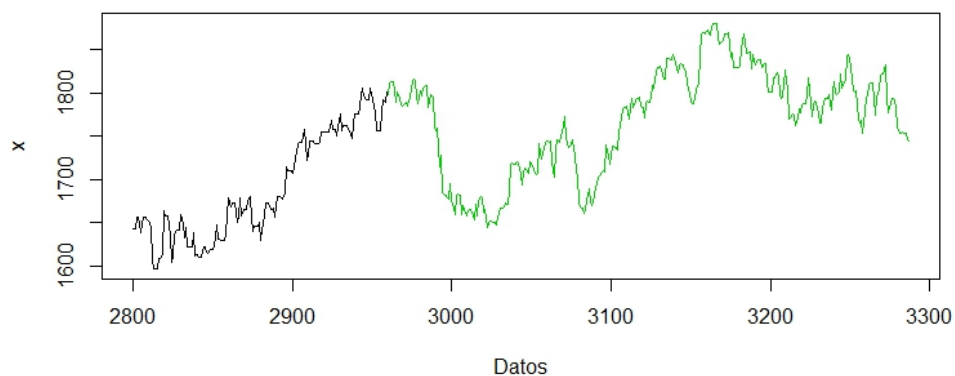


Figura 4.14: Pronóstico del modelo ARIMA(3,1,2)

En la figura que se muestra a continuación se aprecia un contraste entre los datos reales y los datos pronosticados para el precio del cobre.

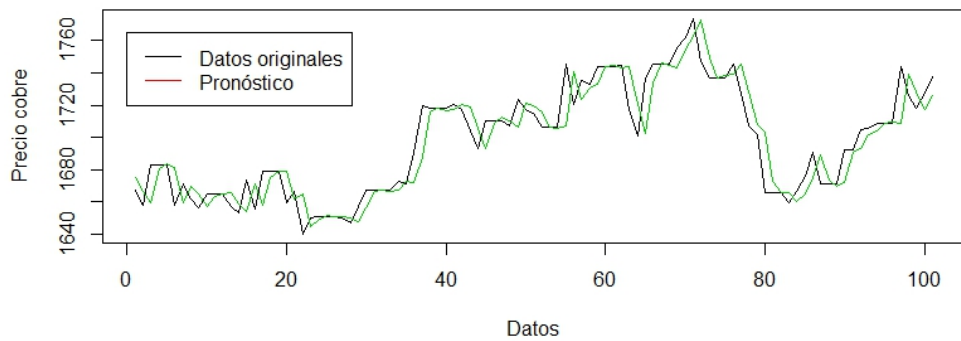


Figura 4.15: Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(3,1,2)

En la Figura 4.16 se muestra la imagen donde se aprecia el pronóstico para 327 datos pronosticados del modelo ARIMA(1,1,1) y solo se muestra de la observación 2800 en adelante para apreciar de mejor manera los datos.

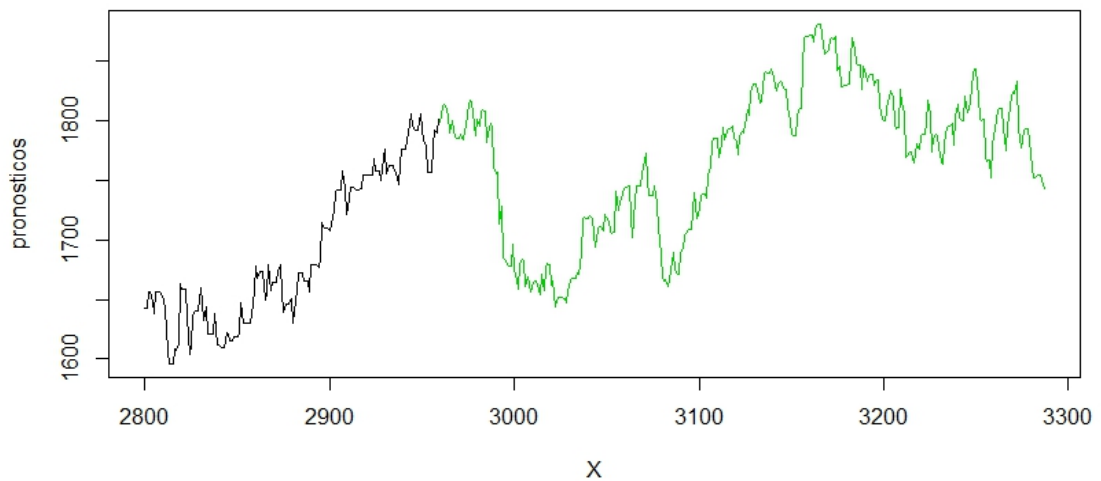


Figura 4.16: Pronóstico del modelo ARIMA(1,1,1)

En la figura que se muestra a continuación se aprecia un contraste entre los datos reales y los datos pronosticados para el precio del cobre.

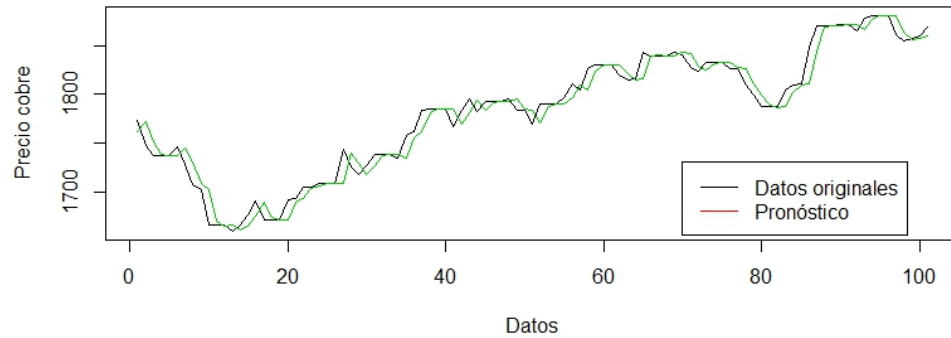


Figura 4.17: Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(1,1,1)

Para saber cual modelo predijo de manera mas precisa los datos se calculan cuatro medidas de errores de pronósticos. En la Tabla 4.10, se muestra el resultado del error medio (ME), error absoluto medio (MAE), raiz del error cuadrático medio (RMSE) y el error medio porcentual absoluto (MAPE). El modelo con mayor precisión al momento de pronosticar, es aquel que contiene los valores menores. Por ende y como se aprecia el modelo ARIMA(1,1,1) es aquel que contiene los menores valores de las medidas de error. Se concluye que el modelo ARIMA(1,1,1) es seleccionado como adecuado para realizar pronósticos.

Error de pronóstico	ARIMA(1,1,1)	ARIMA(3,1,2)
ME	0,1655	0,1693
RMSE	13,8651	13,8406
MAE	9,5255	9,6104
MAPE	0,5389	0,5437

Tabla 4.10: Valores de los errores de pronóstico de los modelo ARIMA para el cobre.

4.3. Análisis de serie de tiempo univariada del dólar

4.3.1. Datos del dólar

La recopilación de información para la serie de tiempo del dólar se realizó mediante los datos que proporciona el banco central de Chile a través de su página web (<http://bcentral.cl>). La importancia de esta entidad gubernamental es que vela por la estabilidad de la moneda y que se cumpla un funcionamiento de manera normal en cuanto a pagos internos y externos, también a través del banco se emiten billetes y monedas, entre otros objetivos no menos importantes. Además el banco central es el encargado de regular y estabilizar la tasa de cambio del dólar.

Para el análisis de la moneda de mayor circulación en el mundo se utilizan 3287 datos, que corresponden al período comprendido entre el 1 de Enero del 2006 y el 31 de Diciembre del 2014. En la siguiente tabla es posible apreciar como se distribuyen los datos a través de los años, es decir, la cantidad de observaciones y datos faltantes.

Año	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Cantidad de datos	365	365	366	365	365	365	366	365	365	3287
Datos faltantes	0	0	0	0	0	0	0	0	0	0
Porcentaje (%)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Tabla 4.11: Cantidad de datos por año.

Es posible apreciar y resalta que no se observan datos faltantes, es cierto que el banco central no realiza transacciones en días Sábado, Domingo y festivos, pero es necesario mencionar que para los casos en que el banco central de Chile tenga cerrado dichos días el valor que adquiere la divisa se mantiene del día hábil anterior.

4.3.2. Análisis descriptivo de los datos

Al tener la totalidad de los datos y sin observaciones que falten es posible comenzar un completo análisis en el precio del cobre, los datos se almacenan en un archivo Excel en formato CSV y los formatos de fecha están expresadas como dd-mm-aaaa, además de apreciar la fecha se muestra el valor del dólar en ese periodo.

N°	Fecha	Precio dólar (CLP)
1	01-01-2006	514,21
2	02-01-2006	512,50
3	03-01-2006	513,18
4	04-01-2006	518,04
5	05-01-2006	515,45
⋮	⋮	⋮
⋮	⋮	⋮
3283	27-12-2014	609,17
3284	28-12-2014	609,17
3285	29-12-2014	608,15
3286	30-12-2014	607,38
3287	31-12-2014	607,38

Tabla 4.12: Vista de los datos.

Una vez definidos los datos cuales se realiza el análisis de series de tiempo para el dólar se procede a realizar un cuadro descriptivo de la variable de interés, tabla que se presenta a continuación:

Variable	Min.	Q_1	Mediana	Media	Q_3	Máx.	Curtosis	Asimetría	Varianza
Cobre	431,20	485,20	517,42	520,01	540,62	676,75	3,99	0,89	1843,98

Tabla 4.13: Estadísticos descriptivos precio del dólar

En la Tabla 4.13 es posible apreciar que durante el período estudiado el valor mas bajo del precio de dólar es 431,20 pesos y alcanzando un máximo de 676,75 pesos, donde el precio medio del dolar es 520,01 pesos para las 3287 observaciones. Por otra parte se tiene que el valor del coeficiente de asimetría es 0,89, mediante este valor se puede inferir que gran cantidad de los datos están almacenados a la izquierda de la media. La curtosis es de 3,99 con este valor se dice que la distribución de los datos es Leptocúrtica, esto quiere decir que una gran concentración de los datos giran en torno a un valor central.

En la Figura 4.18 es posible observar el histograma de frecuencias del precio del dólar, a través de esta figura se confirma el valor obtenido mediante el coeficiente de asimetría, además del valor obtenido mediante el cálculo de la curtosis.

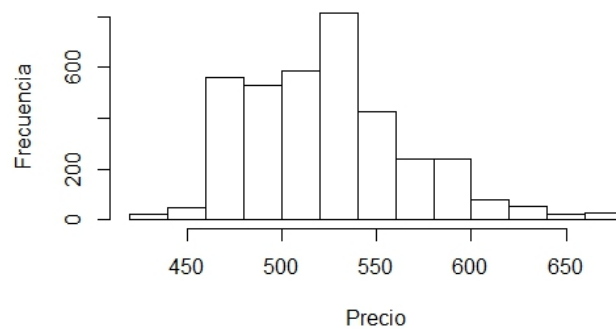


Figura 4.18: Histograma de frecuencia de los valores del dólar

4.3.3. Análisis de serie temporal del dólar

En esta sección del capítulo se aborda un análisis de series de tiempo para encontrar el modelo que ajuste de mejor manera los datos y así luego pronosticar valores a través de observaciones pasadas.

Una parte fundamental del análisis es observar como se reparten los datos a lo largo del tiempo, es por ello que a continuación se presenta la serie de tiempo del precio del dólar entre el período comprendido entre el 1 de Enero del 2006 y el 31 de Diciembre del 2014.

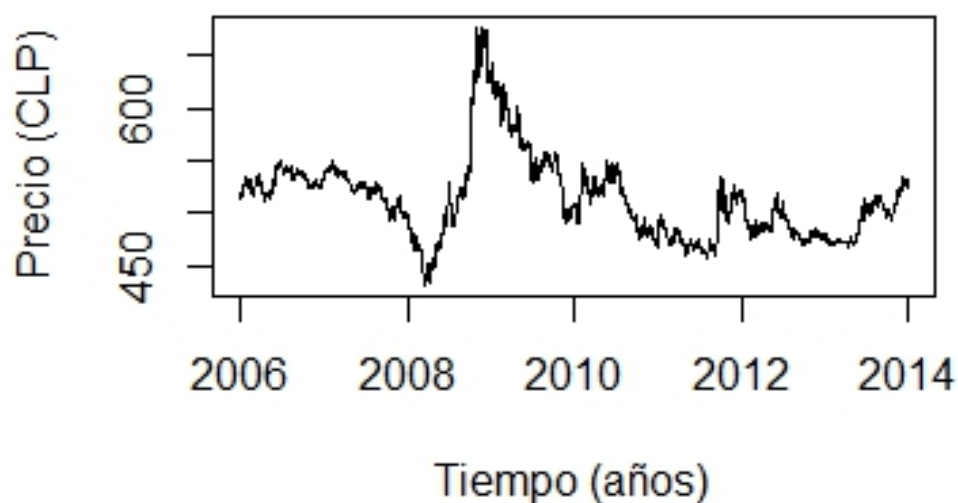


Figura 4.19: Precios del dólar durante los años 2006-2014 en Chile

En la figura anterior se aprecia que el punto mas bajo del precio de dolar fue a principios del año 2008 y que inmediatamente comenzó un alza para alcanzar el punto mas alto en el año 2009. Otro punto a destacar es que a partir de mediados de 2013 el dólar inició un alza en su precio que se mantiene hasta finales del año 2014.

En la Figura 4.20 se presenta el gráfico Q-Q normal, el cual compara los cuantiles de las observaciones con los cuantiles teóricos de la distribución normal, dejando en evidencia que los datos no presentan normalidad. A modo de confirmar la no normalidad de los datos se realiza el *test* de Jarque-Bera, el resultado del p-valor de dicho *test* es muy pequeño, por lo que se rechaza la hipótesis nula (Los datos se aproximan a una distribución normal)

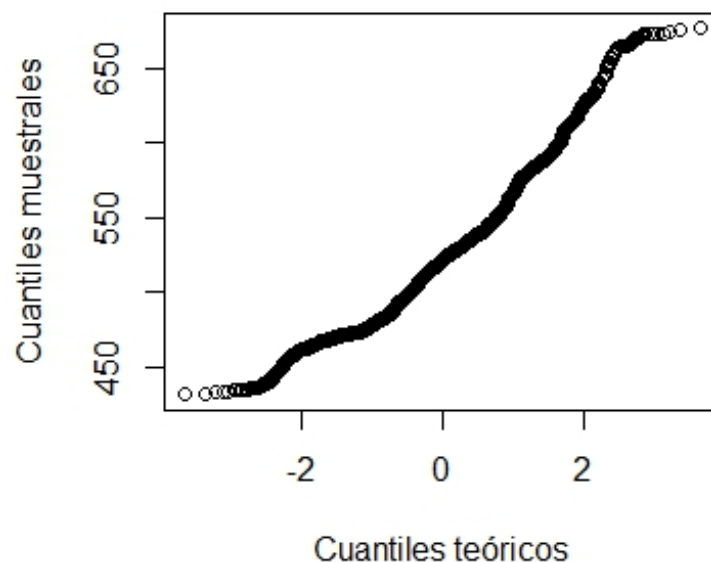


Figura 4.20: Gráfico Q-Q normal de los datos del dólar

Un paso importante en el análisis de series de tiempo es determinar la estacionariedad de los datos, para ello se aplican tres pruebas:

- Test Dickey-Fuller aumentado (ADF)
- Test Kwiatkowski-Phillips-Schmidt-Shin (KPSS)
- Test Phillips-Perron (PP)

Es importante que al aplicar estas pruebas se compruebe que los datos son estacionarios,

Prueba	p-valor	H_0
ADF	0,71	Serie de tiempo no estacionaria
KPSS	0,01	Serie de tiempo estacionaria
PP	0,69	Serie de tiempo no estacionaria

Tabla 4.14: Resultados de las pruebas de raíz unitarias del precio dólar

En la Tabla 4.14 se aprecia que los datos no cumplen con la condición de estacionariedad, entonces se aplica una transformación logarítmica para luego realizar una diferenciación regular. Una vez aplicadas estas transformaciones de los datos se vuelve a aplicar las pruebas de raíces unitarias para comprobar que ahora los datos cumplen con la condición de estacionariedad. Condiciones que se cumplen y se observan en la Tabla 4.15

Prueba	p-valor	H_0
ADF	0,01	Serie de tiempo no estacionaria
KPSS	0,1	Serie de tiempo estacionaria
PP	0,01	Serie de tiempo no estacionaria

Tabla 4.15: Resultados de las pruebas de raíz unitarias a la diferenciación logarítmica del precio dólar

Como ya se cumple esta condición se puede continuar para lograr construir un modelo de predicción para el logaritmo de los precios del dólar. Un punto importante a mencionar es que para la construcción del modelo se trabajara con una menor cantidad de datos, todo esto con el fin de facilitar el trabajo del software R-project y así disminuir tiempos. Se utiliza el 90% de los datos que corresponden a 2959 observaciones (datos de entrenamiento).

4.3.4. Construcción del modelo

Entonces para esta sección de la construcción del modelo es importante observar como quedan los datos luego de que se aplique la diferenciación logarítmica de los precios del dólar, a través de la Figura 4.21 se aprecia el comportamiento de los datos a través del tiempo en el cual se ha eliminado la tendencia y dispersión de las observaciones.

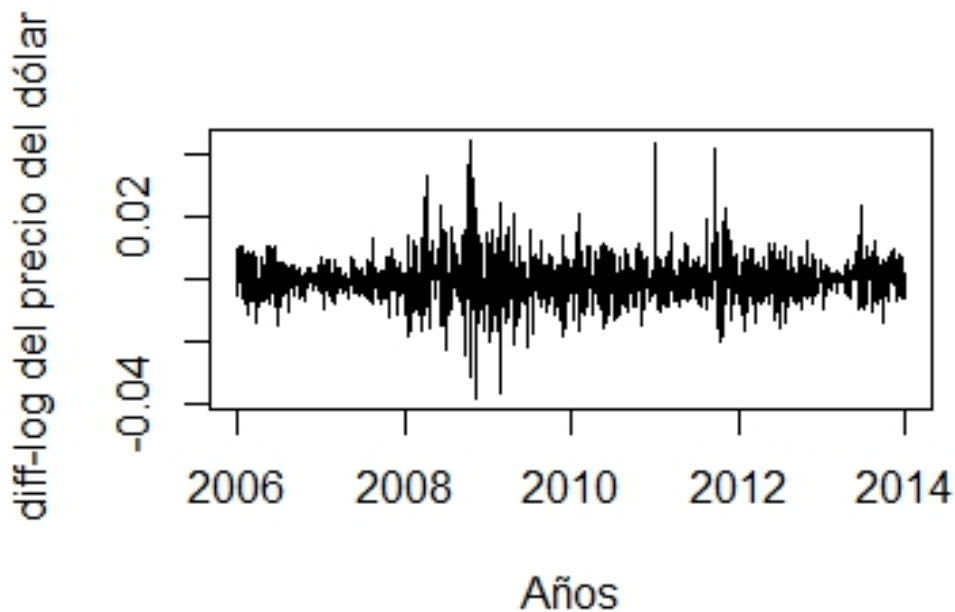


Figura 4.21: Gráfico de dispersión de la primera diferenciación regular del logaritmo del dólar

Para identificar los componentes autorregresivos es necesario analizar la función de autocorrelación parcial de la serie de tiempo por lo tanto se debe identificar aquellos rezagos significativos, esos que están por fuera de la banda de confianza en el correlograma. Por otra parte el componente media móvil de la serie de tiempo se identifica mediante la función de autocorrelación simple también se debe observar los rezagos significativos.

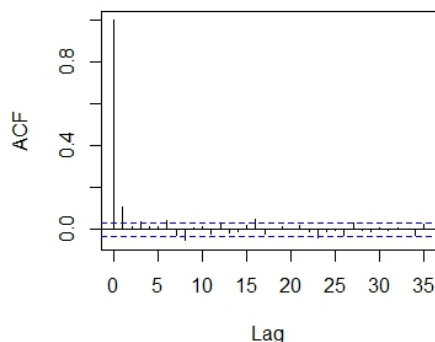


Figura 4.22: Función de autocorrelación simple del logaritmo con diferenciación regular del dólar

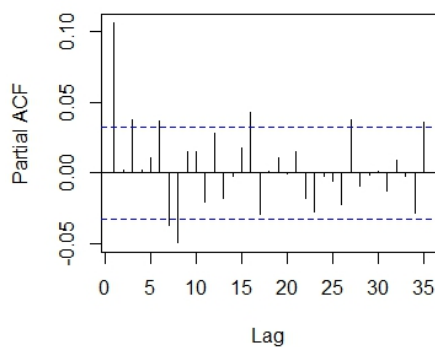


Figura 4.23: Función de autocorrelación parcial del logaritmo con diferenciación regular del dólar

Para tener una mejor estimación de los parámetros del modelo se utiliza el *software* R-project para tener valores mas precisos y que el modelo se ajuste mejor a los datos.

4.3.5. Identificación del modelo y estimación de parámetros

Una vez que se tiene una serie temporal la cual es estacionaria se procede a realizar un ajuste de modelos ARIMA. Por lo tanto en esta sección se ajustan los datos a un modelo ARIMA parsimonioso y que explique de la mejor manera posible los datos.

Identificación del modelo ARIMA

Para lograr una identificación del modelo mas parsimonioso se ocupa el *software* R-project, en complemento de los gráficos presentados anteriormente, como lo son

las funciones de autocorrelación simple, parcial y de dispersión.

Los criterios de la información Akaike y Bayesiano ayudan a encontrar el modelo que mejor represente los datos, es por esto que se prueban una serie de modelos

Modelo	AIC	BIC
Arima(0,1,1)	-25031,06	-25018,87
Arima(1,1,1)	-25029,4	-25011,11
Arima(1,1,2)	-25027,36	-25002,97
Arima(1,1,3)	-25027,35	-24996,87
Arima(1,1,4)	-25025,54	-24988,96
Arima(2,1,0)	-25029,45	-25011,16
Arima(2,1,1)	-25027,45	-25003,06
Arima(2,1,2)	-25027,74	-24997,25
Arima(2,1,3)	-25025,78	-24989,2
Arima(2,1,4)	-25024,06	-24981,37
Arima(3,1,0)	-25029,68	-25005,29
Arima(3,1,1)	-25027,57	-24997,08
Arima(3,1,2)	-25025,63	-24989,04
Arima(3,1,3)	-25029,57	-24996,89
Arima(3,1,4)	-25037,60	-24988,82
Arima(4,1,0)	-25027,72	-24997,23
Arima(4,1,1)	-25025,71	-24989,12
Arima(4,1,2)	-25023,87	-24981,19
Arima(4,1,3)	-25037,07	-24988,99
Arima(4,1,4)	-25021,86	-24966,98

Tabla 4.16: Resultados del ajuste de los modelos ARIMA a los datos del logaritmo del precio del dólar.

En la Tabla 4.16 se aprecia que los modelos ARIMA que presentan un menor valor en sus criterios de la información Akaike y Bayesiano son los modelos ARIMA(3,1,4) y ARIMA(0,1,1) respectivamente, por lo tanto de los veinte modelos planteados inicialmente solo se reducen a dos ya que se buscan los menores valores AIC y BIC.

Entonces los modelos de interés serán dos y por lo tanto ahora se deben calcular los parámetros que se encuentran presentes en cada modelo planteado.

Parámetro	Estimación	Error estándar	z value	Pr(> z)
ar1	-0,138838	0,286929	-0,4839	0,62847
ar2	-0,338205	0,226355	-1,4941	0,03514
ar3	0,650098	0,272928	2,3819	0,01722
ma1	0,254203	0,288388	0,8815	0,37807
ma2	0,363294	0,266651	1,3624	0,17306
ma3	-0,570572	0,306809	-1,8597	0,04293
ma4	-0,066513	0,049431	-1,3456	0,17844

Tabla 4.17: Contraste de significación de los parámetros del modelo ARIMA(3,1,4)

Parámetro	Estimación	Error estándar	z value	Pr(> z)
ma1	0,113671	0,017239	6,5938	4,286e-11

Tabla 4.18: Contraste de significación de los parámetros del modelo ARIMA(0,1,1)

En la Tabla 4.17 y la Tabla 4.18 se dan a conocer los contrastes de significación del modelo ARIMA (3,1,4) y el modelo ARIMA(0,1,1) respectivamente. Para el caso del modelo ARIMA(3,1,4) se aprecia que cuatro de los siete parámetros del modelo no son significativos para el modelo, esto se debe a que tiene un p-valor mucho mayor que un nivel de significación de $\alpha = 0,5$. Para el caso del modelo ARIMA(0,1,1) se tiene que no hay parámetros no significativos, es decir que ningún p-valor es mayor a $\alpha = 0,5$.

Lo que se presenta a continuación es el ajuste de los modelos ARIMA(3,1,4) y ARIMA(0,1,1), para lograr tener una mejor visualización de los datos el ajuste se realizó solo a 100 datos para lograr tener una mejor visualización del ajuste del modelo.

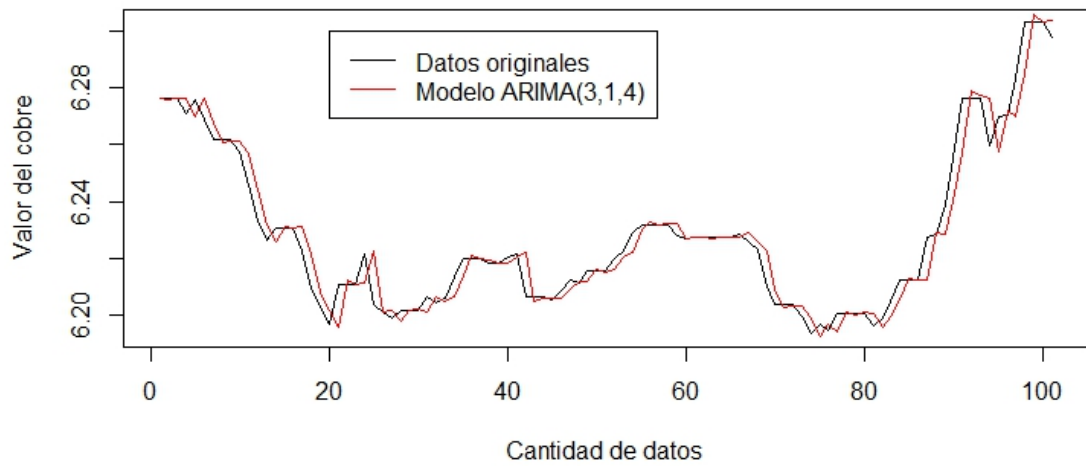


Figura 4.24: Ajuste del modelo ARIMA(3,1,4) a los datos del logaritmo del dólar

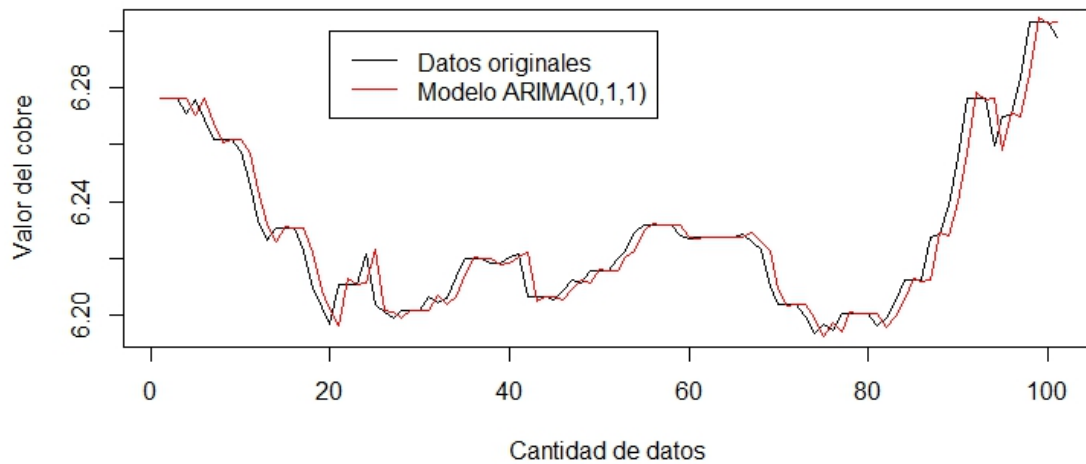


Figura 4.25: Ajuste del modelo ARIMA(1,1,1) a los datos del logaritmo del dólar.

Para determinar si estos modelos son los adecuados para representar a los datos se deben analizar los residuos, análisis que se lleva a cabo a continuación en la sección de diagnóstico del modelo.

4.3.6. Diagnóstico del modelo

Una vez seleccionados los posibles mejores modelos para la representación de los datos, sigue el paso de diagnóstico de los modelos propuestos, para este caso son los modelos $ARIMA(3,1,4)$ y $ARIMA(0,1,1)$. La validación de un modelo corresponde a analizar los residuos del modelo, un modelo es el adecuado cuando se cumple que los residuos se deben comportar como ruido blanco. (Recordar que para que se comporten como ruido blanco la media es igual a cero, una varianza constante y observaciones independientes). Para determinar si la media es cero y la varianza constante se analizara el gráfico de dispersión, por otro lado se observara el correlograma de la función de autocorrelación parcial y simple para la independencia de los residuos y por ultimo mediante la prueba de Ljung-Box se verificará que los residuos sean no autocorrelacionados. Entonces el modelo que cumpla con dichos requisitos se seleccionará como el modelo que represente a los datos de la diferenciación logarítmica del precio del dólar.

En la Figura 4.26 y Figura 4.27 se aprecian los gráficos de dispersión de los residuos de ambos modelos seleccionados y analizados. En ambos gráficos de dispersión es posible apreciar una media y varianza constante, por lo tanto se cumple la primera condición.

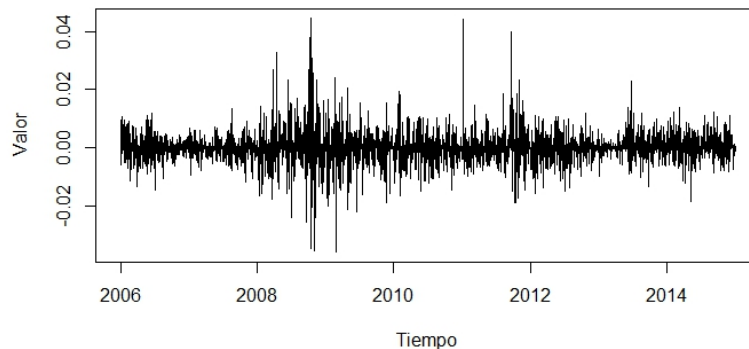


Figura 4.26: Gráfico de dispersión de los residuos del modelo $ARIMA(3,1,4)$

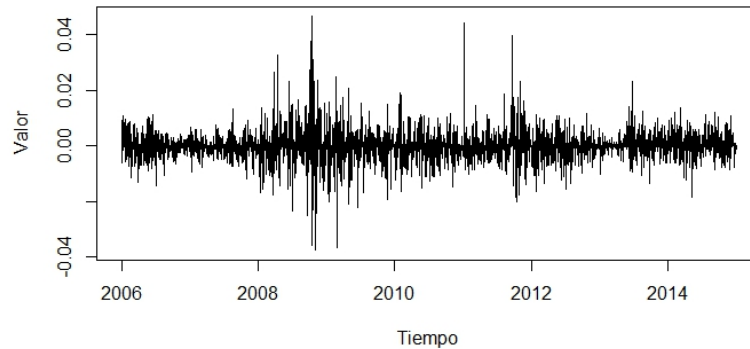
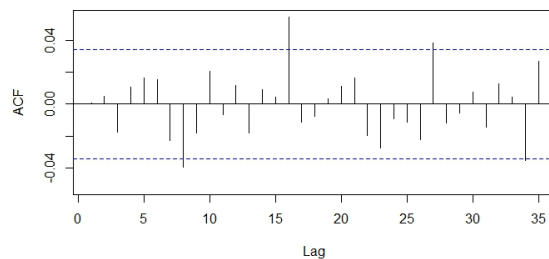
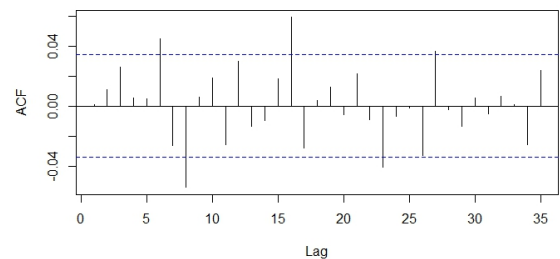


Figura 4.27: Grafico de dispersión de los residuos del modelo ARIMA(0,1,1)

En los gráficos que se presentan a continuación se aprecian las funciones de autocorrelaciones simples de ambos modelos propuestos, es posible apreciar que no se presentan gran cantidad de correlaciones significativas que estén por fuera de las bandas de confianza, por lo tanto los residuos a primera vista son independientes entre sí, es así como no existe autocorrelación en los residuos o dependencia en el tiempo en los residuos.



(a) ACF de los residuos del modelo ARI-MA(3,1,4)



(b) ACF de los residuos del modelo ARI-MA(0,1,1)

Figura 4.28: ACF de los residuos de modelos ARIMA propuestos

La prueba de Ljung-Box asegura de manera estadística la independencia de los residuos, es por esto que se aplica la prueba a los modelos seleccionados. En la Tabla 4.19 se aprecian los resultados, en donde no se rechaza la hipótesis nula de Ljung-Box que dice que la distribución de los datos son independientes. Entonces se confirma que los residuos son independientes.

Modelo	p-valor
ARIMA(3,1,4)	0,9742
ARIMA(0,1,1)	0,9702

Tabla 4.19: Resultado de test Ljung-Box aplicado a los modelos seleccionados

Entonces al cumplir con las condiciones y tener los modelos con el menor AIC y BIC, desde ahora se trabajara con estos dos modelos, específicamente el modelo ARIMA(3,1,4) y ARIMA(0,1,1), ya que cumplen con todos los requisitos de los criterios de selección.

4.3.7. Pronóstico

En la sección 2.4 se habló de pronóstico y errores de predicción, en esta parte del capítulo se realiza el pronóstico de los datos, mencionar que se realizan predicciones sobre los modelos ARIMA seleccionados (ARIMA(0,1,1) y ARIMA(3,1,4)). Entonces se procederá a realizar el pronóstico del 10% de los datos que y se determinará el mejor mediante el análisis de los errores de pronósticos.

En la Figura 4.29 se muestra la imagen donde se aprecia el pronóstico para 327 datos pronosticados del modelo ARIMA(0,1,1) y solo se muestra de la observación 2800 en adelante para apreciar de mejor manera los datos.

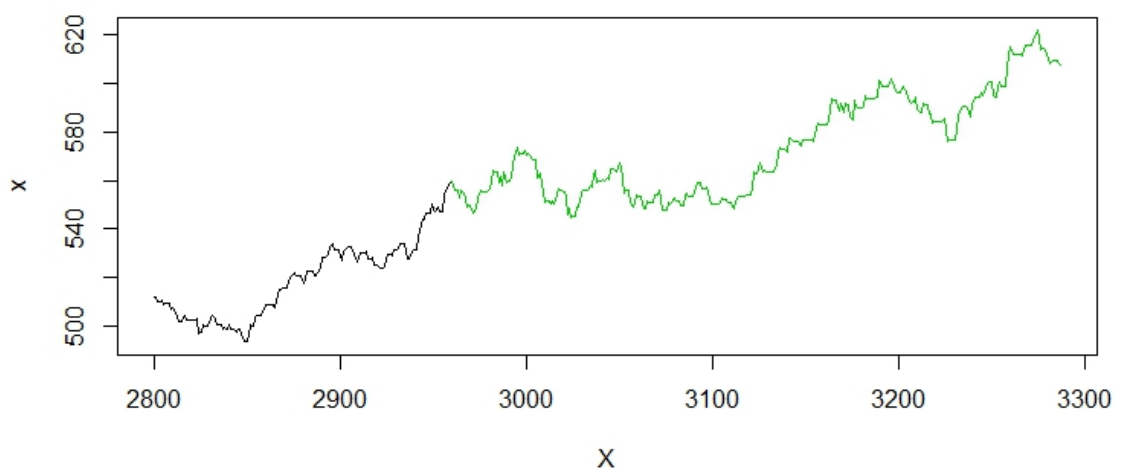


Figura 4.29: Pronóstico del modelo ARIMA(0,1,1)

En la figura que se muestra a continuación se aprecia un contraste entre los datos reales y los datos pronosticados para el precio del dólar.

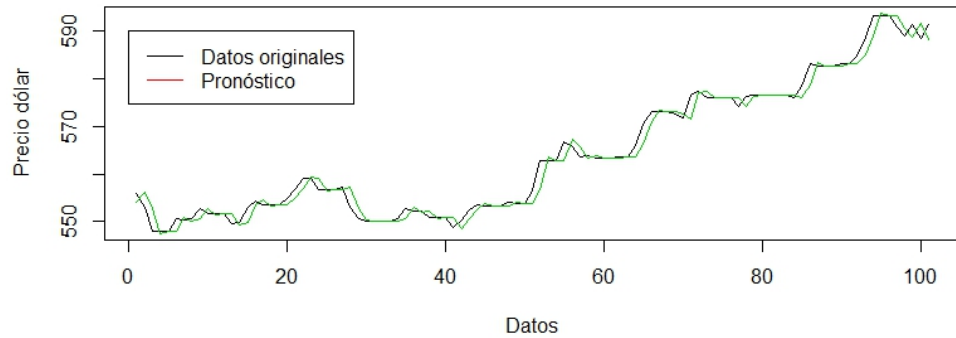


Figura 4.30: Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(0,1,1)

En la Figura 4.31 se muestra la imagen donde se aprecia el pronóstico para 327 datos pronosticados del modelo ARIMA(3,1,4) y solo se muestra de la observación 2800 en adelante para apreciar de mejor manera los datos.

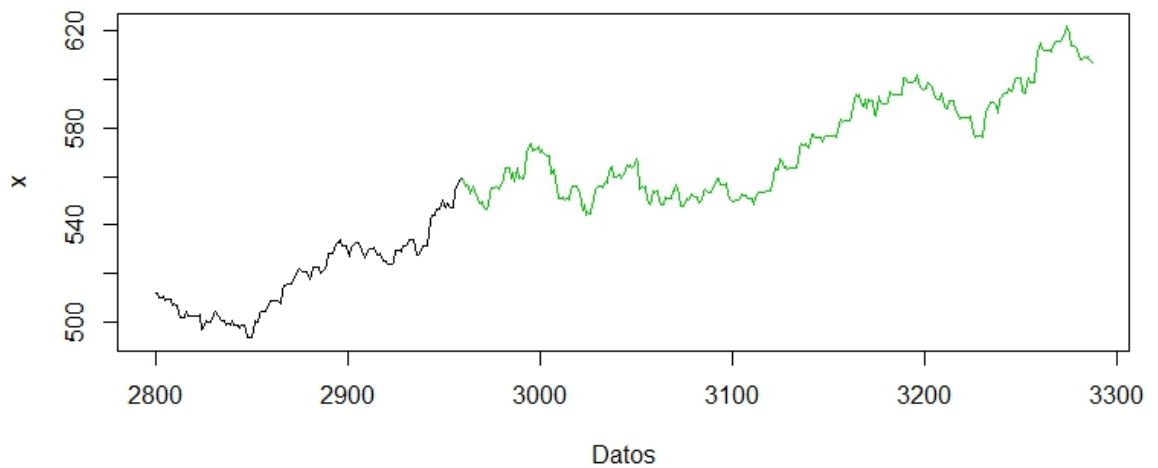


Figura 4.31: Pronóstico del modelo ARIMA(3,1,4)

En la figura que se muestra a continuación se aprecia un contraste entre los datos reales y los datos pronosticados para el precio del dólar.

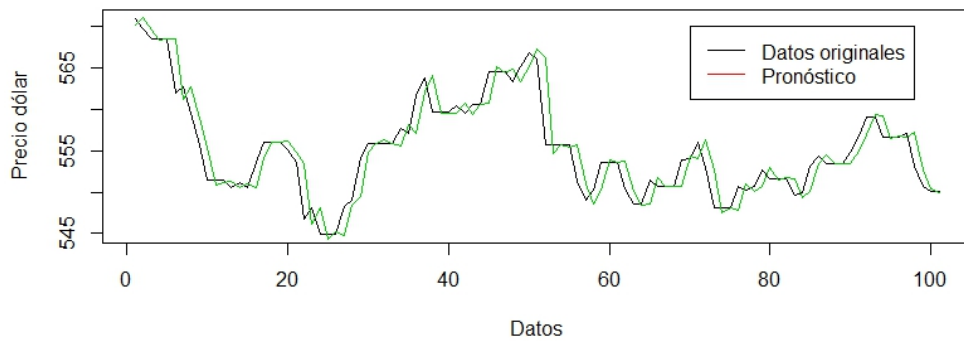


Figura 4.32: Serie de precios del dólar en contraste con el pronóstico estimado por el modelo ARIMA(3,1,4)

Para saber cual modelo pronostica de mejor manera las observaciones se calculan tres medidas de errores de pronósticos. En la Tabla 4.20, se muestra el resultado del error medio (ME), error absoluto medio (MAE) y raíz del error cuadrático medio (RMSE). El modelo con mayor precisión al momento de pronosticar, es aquel que contiene los valores menores. Por ende y como se aprecia el modelo ARIMA(0,1,1) es aquel que contiene los menores valores de las medidas de error. Se concluye que el modelo ARIMA(0,1,1) es seleccionado como adecuado para realizar pronósticos.

Error de pronóstico	ARIMA(0,1,1)	ARIMA(3,1,4)
ME	-0,1288053	-0,1216698
RMSE	2,218435	2,234842
MAE	1,461198	1,493885
MAPE	0,255269	0,2609301

Tabla 4.20: Valores de los errores de pronóstico de los modelo ARIMA para el dólar.

4.4. Análisis de serie de tiempo multivariada del cobre y dólar

4.4.1. Datos del cobre y dólar

Como se mencionó anteriormente los datos del cobre fueron recopilados a través de la bolsa de metales de Londres (<http://lme.com>) y la recopilación de información para la serie de tiempo del dolar se realizó mediante los datos proporcionados por el banco central de Chile (<http://bcentral.cl>)

Para el análisis de la moneda de mayor circulación en el mundo se utilizan 3287 datos, que corresponden al período comprendido entre el 1 de Enero del 2006 y el 31 de Diciembre del 2014. En la siguiente tabla es posible apreciar como se distribuyen los datos a través de los años, es decir, la cantidad de observaciones y datos faltantes.

Año	2006	2007	2008	2009	2010	2011	2012	2013	2014	Total
Cantidad de datos	365	365	366	365	365	365	366	365	365	3287
Datos faltantes	0	0	0	0	0	0	0	0	0	0
Porcentaje (%)	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

Tabla 4.21: Cantidad de datos por año.

4.4.2. Análisis de serie temporal de cobre y dólar

En esta sección del capítulo se abarca un análisis de series de tiempo de manera multivariada con el fin de encontrar el modelo que mejor ajuste los datos y que interprete de mejor forma la información, para así pronosticar observaciones.

Una parte fundamental del análisis es observar como se distribuyen los datos a través del tiempo, es por ello que a continuación se presenta la serie de tiempo conjunta del cobre y dólar entre el período comprendido entre el 1 de Enero del 2006 y el 31 de Diciembre del 2014.

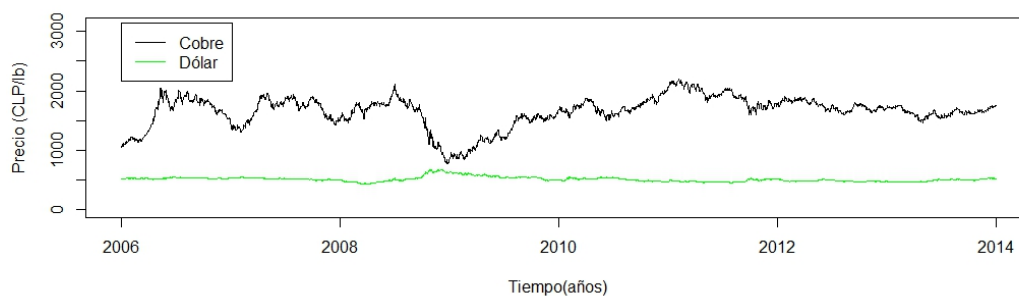


Figura 4.33: Precios del dólar y cobre durante los años 2005-2014 en Chile

Al no tener datos faltantes en los datos es posible realizar un análisis descriptivo de ambas variables. A continuación se presenta el gráfico de correlaciones entre el dólar y cobre:

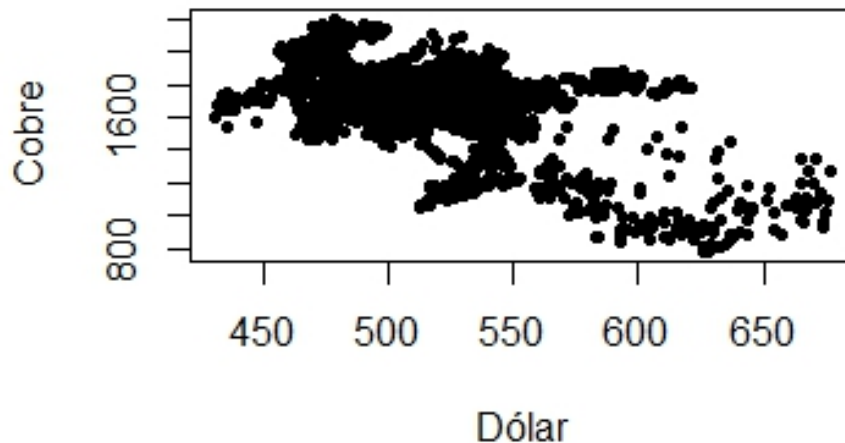


Figura 4.34: Gráfico de correlaciones, correlación =-0,53

Se aprecia en el gráfico anterior una correlación negativa entre los precios de ambas variables, esto se explica debido a que cuando el dolar baja su precio existe una mayor demanda de cobre, por lo tanto el precio del metal va incrementando debido a su demanda.

Como se analizó en secciones pasadas de este mismo capítulo los datos de las series de tiempo del dolar y cobre presentan datos que no son estacionarios, por lo tanto se aplica una transformación logarítmica y luego una diferenciación para eliminar la variabilidad y tendencia de las observaciones respectivamente, entonces después de aplicar la diferenciación logarítmica a los datos si se cumple que los datos sean estacionarios para ambas variables, a continuación se presentan los valores que se tienen el las pruebas de raíz unitaria para ambas variables de interés:

Prueba	p-valor	H_0
ADF	0,01	Serie de tiempo no estacionaria
KPSS	0,1	Serie de tiempo estacionaria
PP	0,01	Serie de tiempo no estacionaria

Tabla 4.22: Resultados de las pruebas de raíz unitarias a la diferenciación logarítmica del precio dólar

Prueba	p-valor	H_0
ADF	0,01	Serie de tiempo no estacionaria
KPSS	0,1	Serie de tiempo estacionaria
PP	0,01	Serie de tiempo no estacionaria

Tabla 4.23: Resultados de las pruebas de raíz unitarias de la diferenciación logarítmica del precio cobre

Por lo tanto a través de lo anterior se prueba que los valores del cobre y dólar, una vez aplicada la diferenciación logarítmica, siguen un proceso estacionario.

Es necesario mencionar que si se construye un modelo sin aplicar la diferenciación regular es posible que los resultados sean espurios, pero para el caso de los procesos VARMA es posible que los datos sean convergentes para tamaños de muestras grandes, a través de la causalidad de Granger es posible determinar si las variables consiguen un equilibrio en el largo plazo, además de definir si una variable explica a otra o al revés, al aplicar la prueba de causalidad de Granger se obtiene que la variable del dólar explica e influye en el precio del cobre.

4.4.3. Construcción del modelo

En la Figura 4.35 se presentan las correlaciones cruzadas, en la teoría se dice que a través de esta representación gráfica es posible determinar los parámetros del modelo, aquí se aprecia que se dificulta bastante identificar los parámetros del modelo es por esto que para establecer los parámetros más adecuados del modelo se utiliza el software R-project.

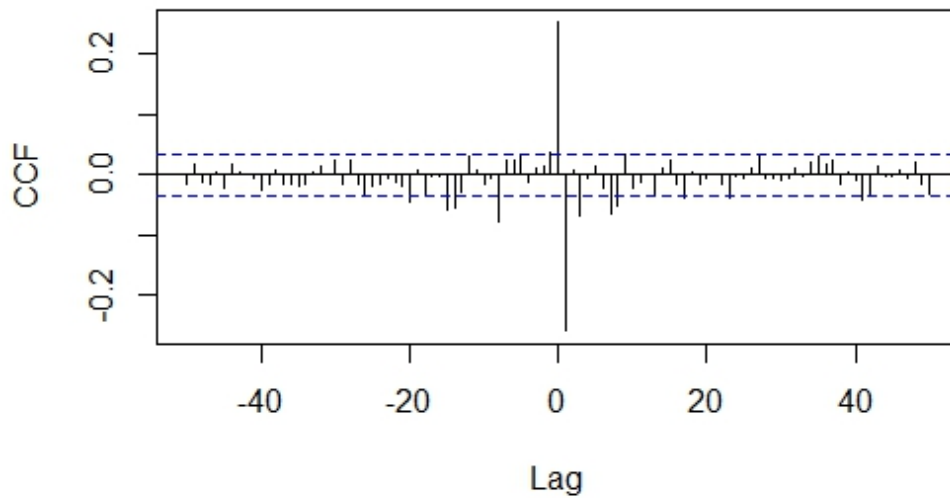


Figura 4.35: Gráfico de correlaciones cruzadas

A través de R-project se prueban diversos modelos con distintos parámetros, para luego proceder a seleccionar los que contengan el menor valor en el criterio de la información Akaike y Bayesiano (descritos en el capítulo 2).

4.4.4. Identificación del modelo y estimación de parámetros

Los datos cumplen con las condiciones necesarias para proceder a la identificación del modelo y estimación de parámetros. En esta sección se ajustará un modelo multivariante de series de tiempo a los datos a través de R-project, en la Tabla 4.24 se ajustan diversos modelos y se considera el que tiene un menor valor de AIC y BIC

MODELO	AIC	BIC
VARMA(1,1)	-17,31765	-17,30127
VARMA(1,2)	-18,67134	-18,64678
VARMA(2,1)	-18,34429	-18,31972
VARMA(2,2)	-18,32918	-18,29642
VARMA(3,1)	-18,66398	-18,63122
VARMA(1,3)	-18,39868	-18,36593
VARMA(2,3)	-18,40591	-18,36497
VARMA(3,2)	-18,57408	-18,53314
VARMA(3,3)	-18,33140	-18,28227
VARMA(4,1)	-18,62737	-18,58642
VARMA(4,2)	-18,56671	-18,51758
VARMA(4,3)	-18,56090	-18,50358
VARMA(4,4)	-18,44316	-18,37765

Tabla 4.24: Resultados del ajuste de los modelos VARMA a los datos del logaritmo del precio del cobre y dólar

El modelo que minimiza los criterios de la información son el VARMA(1,2), por lo tanto es el modelo seleccionado que postula a ser el mas parsimonioso y que interprete de mejor manera la información. Entonces el modelo de interés sera uno y en consecuencia se deben calcular los parámetros que se encuentran presentes en el modelo planteado

Parámetro	Estimación	Error estándar	z value	Pr(> z)
AR(1)[1,1]	0.9980480	0.0011435	872.777	< 2e-16
AR(1)[1,2]	0.0023329	0.0013527	-8.152	0,0846
AR(1)[2,1]	0.0006840	0.0004925	1.725	0,1649
AR(1)[1,2]	0.9992428	0.0005884	1.389	< 2e-16
MA(1)[1,1]	-0.1403323	0.0172145	1698.350	4,44e-16
MA(1)[1,2]	0.2340531	0.0208153	11.244	< 2e-16
MA(1)[1,2]	0.0275210	0.0160420	1.716	0,0862
MA(1)[2,2]	-0.0007605	0.0181843	-0.042	0,9666
MA(2)[1,1]	-0.1001983	0.0030947	-32.378	< 2e-16
MA(2)[1,2]	0.1878532	0.0052113	36.047	< 2e-16
MA(2)[1,2]	-0.0008441	0.0062001	-0.136	0,8917
MA(2)[2,2]	-0.0164960	0.0070125	-2.352	0,0187

Tabla 4.25: Contraste de significación de los parámetros del modelo VAR-MA(1,1)

En la Tabla 4.25 se dan a conocer los contrastes de significación del modelo

VARMA(1,2). Para el caso del modelo VARMA(1,2) se observa que seis de los ocho parámetros no son significativos para el modelo, esto se debe a que tienen un p-valor mayor que un nivel de significación de $\alpha = 0,5$.

Lo que se presenta a continuación es el ajuste del modelo VARMA(1,2), para lograr una mejor visualización de los datos el ajuste se realizó solo a 100 datos para lograr una mejor visualización de dicho ajuste.

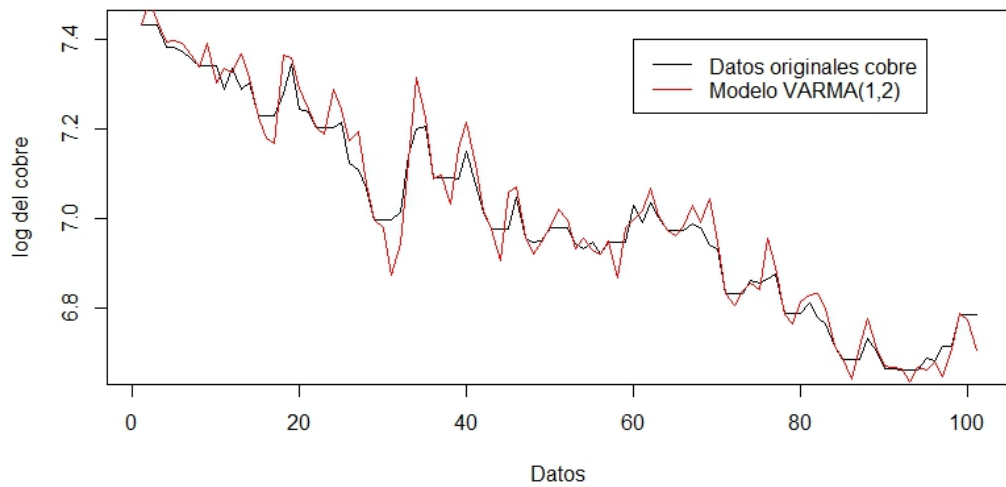


Figura 4.36: Ajuste del modelo VARMA(1,2) a los datos del logaritmo del precio del cobre

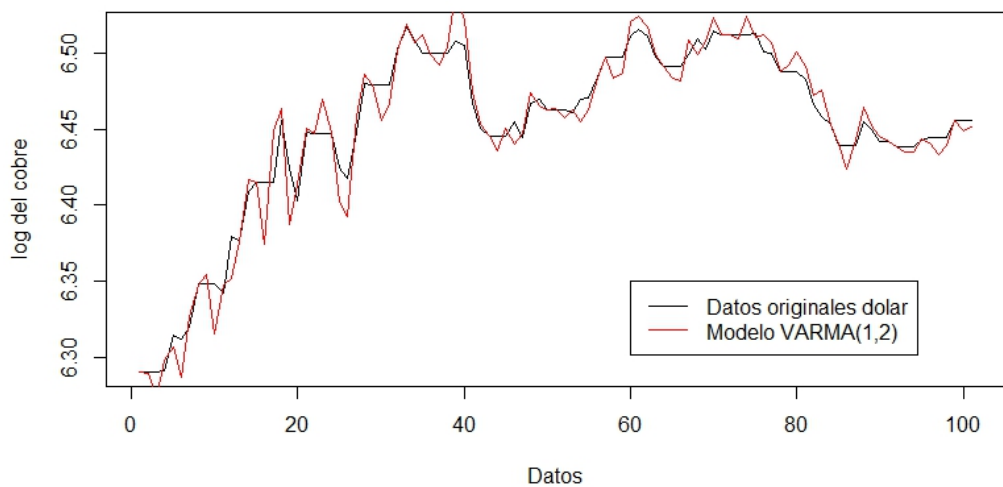


Figura 4.37: Ajuste del modelo VARMA(1,2) a los datos del logaritmo del precio del dólar

Para determinar si el modelo es el adecuados para representar a los datos se deben analizar los residuos, análisis que se lleva a cabo a continuación en la sección de diagnostico del modelo.

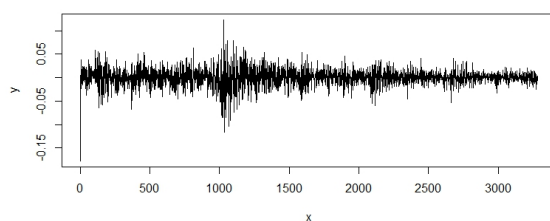
4.4.5. Diagnóstico del modelo

Una vez seleccionado el posible mejor modelo para la representación de los datos, sigue el paso de diagnostico de los modelos propuestos, para este caso son los modelos VARMA(1,2). La validación de un modelo corresponde a analizar los residuos del modelo, un modelo es el adecuado cuando se cumple que los residuos se deben comportar como ruido blanco. Para determinar si la media es cero y la varianza constante se analizara el gráfico de dispersión, por otro lado se observara el correlograma de la función de autocorrelación parcial y simple para la independencia de los residuos y por ultimo mediante la prueba de Ljung-Box se verificará que los residuos sean no autocorrelacionados. Entonces el modelo que cumpla con dichos requisitos se seleccionará como el modelo que represente a los datos.

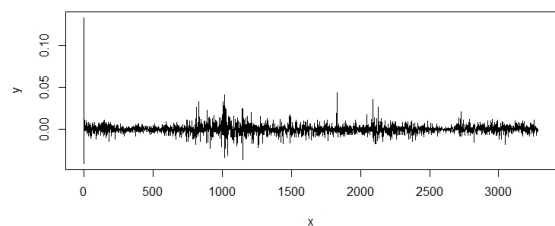
Para el análisis de los residuos que se evalúan mediante el modelo VARMA se procede a separar los residuos de ambas variables, cobre y dólar, entonces se analizan los resultados de ambos residuos de manera separada para cada uno de los modelos planteados.

Análisis de residuos modelo VARMA(1,2)

En el Figura 4.39 se aprecian los gráficos de dispersión de los residuos para el modelo VARMA(1,2) tanto en los residuos que se obtienen para el cobre como el dólar. En ambos gráficos de dispersión se aprecia una media y varianza constate, por lo tanto se cumple dicha condición.



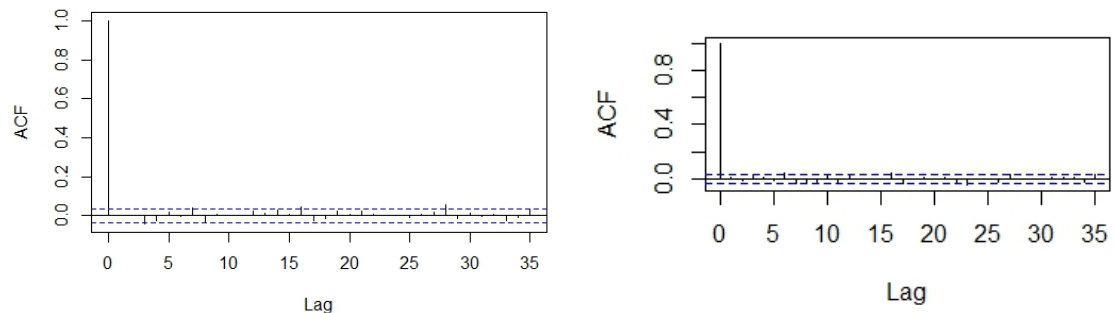
(a) Gráfico de dispersión de los residuos del cobre para modelo VARMA(1,2)



(b) Gráfico de dispersión de los residuos del dólar para modelo VARMA(1,2)

Figura 4.38: Gráfico de dispersión de residuos modelo VARMA(1,2)

En los gráficos que se presentan a continuación se aprecian las funciones de autocorrelaciones simples de ambos modelos propuestos, es posible apreciar que no se presentan correlaciones significativas que estén por fuera de las bandas de confianza, por ende los residuos a primera vista son independientes entre sí, es así como no existe autocorrelación en los residuos.



(a) ACF de los residuos del cobre para modelo VARMA(1,2)

(b) ACF de los residuos del dólar para modelo VARMA(1,2)

Figura 4.39: ACF de residuos modelo VARMA(1,2)

Para tener una real certeza que los residuos son independientes existe la prueba de Ljung-Box que asegura de manera estadística la independencia de los residuos. En la Tabla 4.26 se observa que no se rechaza la hipótesis nula de Ljung-Box que dice que la distribución de los datos son independientes. Por lo tanto se confirma que los residuos son independientes.

Modelo	p-valor
Residuos del cobre	0,7904
Residuos del dólar	0,8205

Tabla 4.26: Resultado de test Ljung-Box aplicado a los residuos del modelo VARMA(1,1)

Entonces el modelo VARMA(1,2) cumple con las condiciones necesarias para realizar predicciones.

4.4.6. Pronóstico

En esta sección del capítulo se realizan pronósticos y medición de errores de predicción, todo ello será realizado sobre el modelo VARMA(1,2). Entonces se procederá a realizar el pronóstico del 10% de los datos y se determinarán sus errores de predicción.

Es necesario mencionar que las predicciones de un modelo VARMA arrojan resultados de pronóstico tanto para los valores del cobre como el dólar, por ende se comparan las predicciones de ambas variables con las observaciones del cobre y dólar. En la Figura 4.40 se aprecia el pronóstico para 327 datos pronosticados del modelo VARMA(1,2) para los precios del cobre y solo se muestra de la observación 2800 en adelante para apreciar de mejor manera los datos.

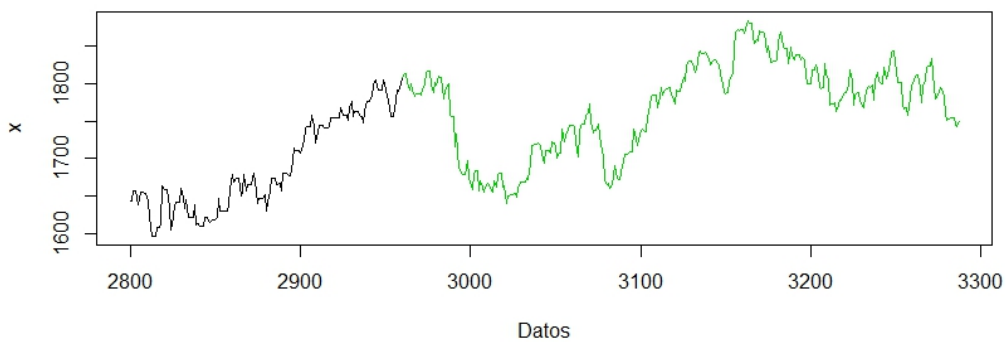


Figura 4.40: Pronóstico del modelo VARMA(1,2) para el cobre

En la Figura 4.41 se observa el pronóstico estimado por el modelo VARMA (1,2), en dicho gráfico se aprecia un pronóstico similar pero no igual

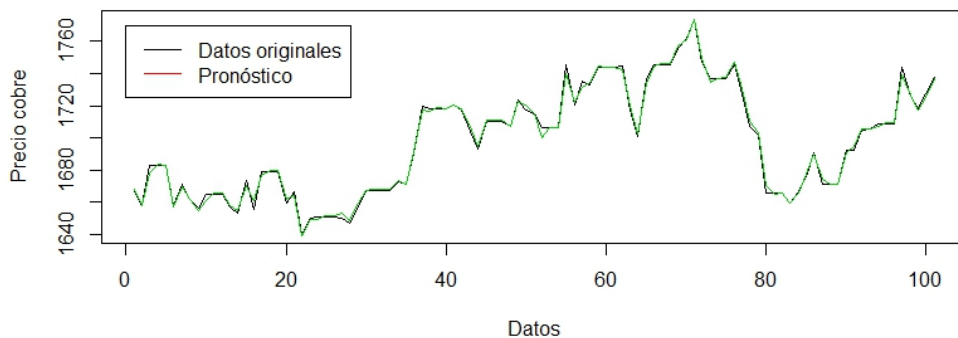


Figura 4.41: Serie de precios del cobre en contraste con el pronóstico estimado por el modelo VARMA (1,2)

En la Figura 4.42 se aprecia el pronóstico para 327 datos pronosticados del modelo VARMA(1,2) en los precios del dólar

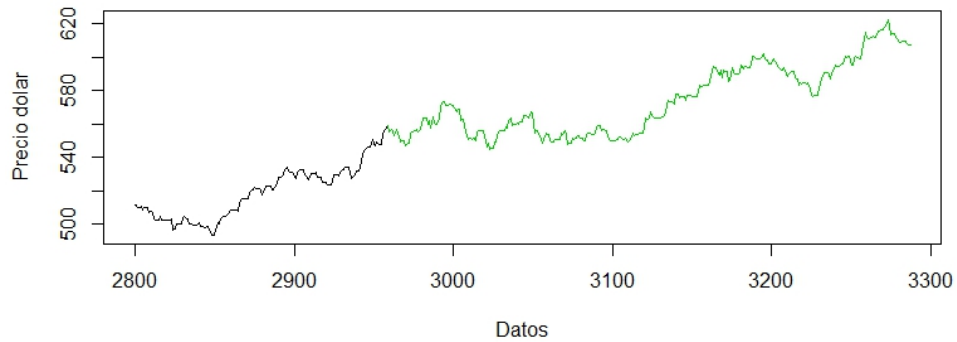


Figura 4.42: Pronóstico del modelo VARMA(1,2) para el dólar

En la Figura 4.43 se observa el pronóstico estimado por el modelo VARMA (1,2), en dicho gráfico se aprecia un pronóstico similar pero no igual, todo ello con respecto a los valores del dólar.

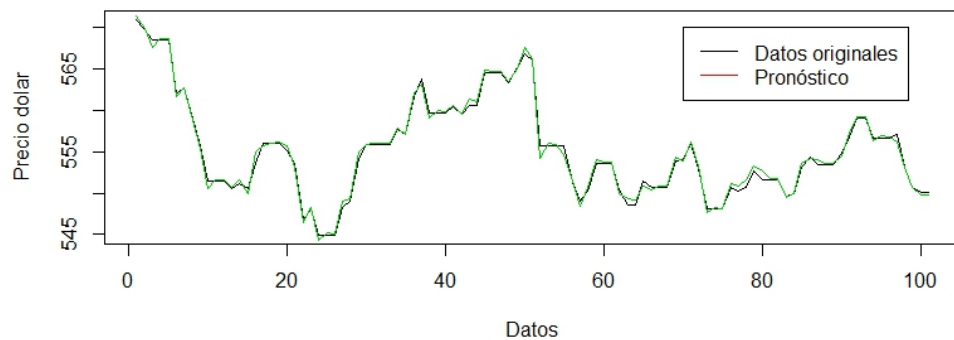


Figura 4.43: Serie de precios del dólar en contraste con el pronóstico estimado por el modelo VARMA (1,2)

En la Tabla 4.27 y Tabla 4.28 se muestran los resultados de los errores de pronóstico para el dólar y cobre a través de un proceso VARMA(1,1), en dichas tablas se muestran el resultado del error medio (ME), error absoluto medio (MAE), y raíz del error cuadrático medio (RMSE). Con estos valores se puede comparar con los modelos univariados de series de tiempo descritos en secciones anteriores de este capítulo y así realizar las conclusiones necesarias que se abordan en el siguiente capítulo.

Error de pronóstico	VARMA(1,2)
ME	0,4164418
RMSE	2,038931
MAE	1,433269
MAPE	0,08115307

Tabla 4.27: Valores de los errores de pronóstico de los modelo VARMA para el cobre.

Error de pronóstico	VARMA(1,2)
ME	0,1693471
RMSE	0,5056794
MAE	0,3633562
MAPE	0,06329184

Tabla 4.28: Valores de los errores de pronóstico de los modelo VARMA para el dólar.

Capítulo 5

Conclusiones

Para este trabajo de titulación se desarrolló un análisis de los precios del cobre y dólar, donde se logra determinar que a medida que el cobre eleva sus precios ocurre lo contrario con el dólar, esto se debe a que cuando baja el precio de la divisa norteamericana existe una real demanda mayor por adquirir el metal y en variadas ocasiones se produce una escasez de cobre, es por eso que sube.

Para este trabajo de titulación se presentaron y probaron diversos modelos a los datos que se aplicaron diversas transformaciones y pruebas para cumplir con los requisitos. Cabe recordar que el objetivo general de este trabajo es diseñar un modelo de pronóstico para el precio del cobre y dólar en Chile, por ende construir un modelo que represente de mejor manera la información de los datos. En la teoría de series de tiempo existen diversos modelos y para el caso de las observaciones presentadas se ajustan modelos de series de tiempo, modelos que siguen un proceso ARIMA y VARMA. Al realizar los respectivos ajustes se determinó considerar los modelos que minimizaran el criterio AIC y BIC, es por ello que para el caso univariado del análisis se selecciona un modelos ARIMA para cada variable y para el caso multivariado o vectorial se selecciona un modelo que minimiza estos criterios.

El modelo que se ajusta al análisis univariado de series de tiempo para los datos del cobre es un modelo ARIMA(1,1,1), entonces este modelo es el que representa mejor la información, dicho modelo es el que minimiza los errores de pronóstico.

El modelo que mejor ajusta los datos para los precios de dólar en el análisis univariado de series temporales es el modelo ARIMA(0,1,1), ya que es el que minimiza los errores de pronóstico.

Para el caso vectorial el modelo que mejor ajusta los datos es el modelo VARMA(1,2), para este caso se probó solo ese modelo ya que coincide en que presenta los menores valores de AIC y BIC.

Al contrastar los modelos univariados para la predicción de precios para ambas variables con los modelos multivariados de series de tiempo, se llega a la conclusión de que el modelo multivariado explica mejor la información y realiza mejores pronósticos

para ambas variables en estudio.

Como trabajo futuro se pretenden ajustar modelos de series de tiempo VARIMA, incluyendo las mismas variables e integrando poco a poco mas variables de interés

Bibliografía

- [1] Peña, D. (2010). Análisis de series temporales. Madrid: Alianza editorial S.A.
- [2] Chatfield, C... (1995). The analysis of time series, fifth edition. United Kingdom: Chapman & Hall/CRC.
- [3] Shumway, R., & Stoffer, D. (2006). Time series analysis and its applications. United States: Springer.
- [4] Gujarati, D. & Porter, D. (2009). Econometría. México: Mc Graw Hill.
- [5] Wei, W. (2006). Time series analysis: univariate and multivariate methods. United States: Pearson education Inc.
- [6] Sarmiento, E. (2013). Predicción con series de tiempo y regresión. Panorama, 2, pp. 36-58.
- [7] Mauricio, J. (2007). Introducción al análisis de series temporales. Madrid, España: Universidad Complutense de Madrid.
- [8] Monsalve, A. & Harmath, P. (2015). Introducción al análisis de series de tiempo con aplicaciones a la econometría y finanzas. Mérida - Venezuela: Ivic.
- [9] González, M. (2009). Análisis de series temporales: Modelos ARIMA. Bilbao - España: Sarriko - on ISBN: 978-84-692-3814-1.
- [10] Hamilton, J. (1994). Time series analysis. New Jersey, United States: Princeton university press.
- [11] Reinsel, G. (1993). Elements of multivariate time series analysis. New York, United States: Springer.
- [12] Lütkepohl, H. (1991). Introduction to multiple time series analysis. German: Springer.
- [13] Lutkepohl, H. (2004). Forecasting with VARMA Models. European university. Sitio web: <http://www.iue.it/personal/Lutkepohl>

- [14] De Gregorio, J., Tokman, A. & Valdes, R. (2005). Tipo de cambio flexible con metas de inflación en Chile: Experiencia y temas de interés. Banco Central de Chile. Sitio web: <http://si2.bcentral.cl/public/pdf/documentos-politica-economica/pdf/dpe14>
- [15] Box, G. & Jenkins, G. (1976). *Time Series Analysis Forecasting and Control*. San Francisco: Holden-Day.
- [16] Sims, C. (1980). *Macroeconomics and reality*. The econometric society, 48, pp.1-48
- [17] De Gregorio, J., Gonzalez, H. & Jaque, F. (2005). Fluctuaciones del dólar, precio del cobre y términos de intercambio. Banco Central de Chile Sitio web: <http://www.bcentral.cl/esp/estpub/estudios/dtbc>
- [18] Foix, C. & Weber, R. (2007). Pronostico del precio del cobre mediante redes neuronales. *Ingeniería de sistemas*, XXI, pp.63-86. Engel, E. & Valdes, R. (2001). Prediciendo el precio del cobre: Más allá del camino aleatorio?. Ministerio de Hacienda. Sitio web: <http://www.minhda.cl>
- [19] Engle, R. F. y Granger, C. W. J. (1989): Cointegration and Error Correction: Representation, Estimation and Testing, *Econometrica*, volumen 55, páginas 251-276.
- [20] Gujarati, Damodar N. (1997). *Econometría*, Editorial McGraw-Hill Interamericana, SA , Santa Fe de Bogotá, Colombia, Cap. 21, pp.693-715
- [21] Ballabriga, Fernando C. y Luis Julián Álvarez González y Javier Jareño Morago (1998) Un Modelo Macroeconómico BVAR para la economía española: metodología y resultados <http://www.bde.es/informes/be/sazul/azul64.pdf>
- [22] Johansen, S., (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* 12, 231–254.
- [23] Asteriuo, Dimitrios. (2001). Notas sobre Análisis de Series de Tiempo: Estacionariedad, Integración y Cointegración (H.L. Mata, Trans.). <http://www.personal.rdg.ac.uk/~less00da/lecture3.htm>.
- [24] Stock, J.H. y Watson, M.W. (2003). *Introduction to Econometrics*. Addison Wesley.
- [25] Vargas, M. (1999). “Modelización de Series Temporales Estacionarias en el Espacio de Estados”. Documento de trabajo 2/1999/4. Fac. CC.EE. y EE. Albacete, U.C.L.M.

Apéndice A

Códigos R-project

Para una correcta realización del trabajo de titulación presentado anteriormente se utilizó la herramienta estadística R-Project

- Librerías que se ocupan para el proyecto de título

```
> library(tseries)
> library(TTR)
> library(forecast)
> library(lmtest)
> library(MTS)
> library(moments)
> library(vars)
> library(xlsx)
```

- Análisis descriptivo del cobre y dólar (cuadro 4.3 y cuadro 4.13)

```
> summary(variable)
> var(variable)
> kurtosis(variable)
> skewness(variable)
```

- Histogramas del cobre y dólar (figura 4.3 y figura 4.18)

```
> hist(Variable,main = "Título", xlab = "Nombre eje X", ylab = "Nombre
eje Y")
```

- Transformación de los datos a serie de tiempo

```
> variable.ts <- ts(variable, start=2006, freq=365)
```

- Representación gráfica del cobre y dólar (figura 4.4 y figura 4.19)

```
plot(variable.ts, type="l", main="Título del gráfico", xlab = "Nombre eje X",
      ylab = "Nombre eje Y")
```

- test de raíces unitarias del cobre y dólar (cuadro 4.4, cuadro 4.5, cuadro 4.14 y cuadro 4.15)

```
> adf.test(variable)
> pp.test(variable)
> kpss.test(variable)
```

- Box-Cox (Aplicación del logaritmo a los datos)

```
> log.variable.ts <- BoxCox(variable.ts, 0)
```

- Funciones de autocorrelación simple y parcial del cobre y dólar (figura 4.7, figura 4.8, figura 4.22 y figura 4.23)

```
> Acf(variable, lag.max = Retardos, main = "Título")
> Pacf(variable, lag.max = Retardos, main = "Título")
```

- Ajustes modelos ARIMA del cobre y dólar (cuadro 4.6 y cuadro 4.16)

```
xx <- log.variable.ts
```

```
arima01 <- Arima(xx, order = c(0, 1, 1))
arima02 <- Arima(xx, order = c(1, 1, 1))
arima03 <- Arima(xx, order = c(1, 1, 2))
arima04 <- Arima(xx, order = c(1, 1, 3))
arima05 <- Arima(xx, order = c(1, 1, 4))
```

```

arima06 <- Arima(xx,order = c(2, 1 ,0))
arima07 <- Arima(xx,order = c(2, 1 ,1))
arima08 <- Arima(xx,order = c(2, 1 ,2))
arima09 <- Arima(xx,order = c(2, 1 ,3))
arima10 <- Arima(xx,order = c(2, 1 ,4))
arima11 <- Arima(xx,order = c(3, 1 ,0))
arima12 <- Arima(xx,order = c(3, 1 ,1))
arima13 <- Arima(xx,order = c(3, 1 ,2))
arima14 <- Arima(xx,order = c(3, 1 ,3))
arima15 <- Arima(xx,order = c(3, 1 ,4))
arima16 <- Arima(xx,order = c(4, 1 ,0))
arima17 <- Arima(xx,order = c(4, 1 ,1))
arima18 <- Arima(xx,order = c(4, 1 ,2))
arima19 <- Arima(xx,order = c(4, 1 ,3))
arima20 <- Arima(xx,order = c(4, 1 ,4))

```

- Contraste de significación de los parámetros de modelos ARIMA del cobre y dólar (cuadro 4.7, cuadro 4.8, cuadro 4.17 y cuadro 4.18)

```
> coeftest(Modelo seleccionado)
```

- Ajuste de modelos ARIMA del cobre y dólar (figura 4.9, figura 4.10, figura 4.29 y figura 4.31)

```

> plot(log.variable.ts[1400:1500],type="l",xlab= " Eje X ",ylab = "Eje Y")
> lines(fitted(Modelo seleccionado)[1400:1500],col=red")
> legend(20,6.3,c("Datos originales", "Modelo ARIMA"),lty=c(1,1),lwd=c(0,1, 0,1),
col=c("black",red"))

```

- Gráfico de dispersión de los residuos modelos ARIMA del cobre y dólar (Figura 4.11, figura 4.12, figura 4.26 y figura 4.27)

```
> plot(modelo seleccionado$ residuals,xlab="Eje X",ylab="Eje Y")
```

- ACF y PACF de los residuos modelos ARIMA del cobre y dólar (Figura 4.13, figura 4.28)

```
> Acf(modelo seleccionado$ residuals,lag.max = Retardos,main="Nombre del gráfico")
```

```
> Pacf(modelo seleccionado$ residuals,lag.max = Retardos,main="Nombre del gráfico")
```

- Ljung-Box test (Cuadro 4.9, cuadro 4.19)

```
> Box.test(modelo seleccionado$ residuals,type="Ljung-Box")
```

- Pronósticos 1 a 1

```
> a< -log.variable.ts
```

```
> xx< -a[1:2959]
```

```
> arima1 <- Arima(xx ,order = c(1, 1, 1))
```

```
> a1< -predict(arima02,a.nhead="1")
```

```
> xx< -a[1:2960]
```

```
> arima1 <- Arima(xx,order = c(1, 1, 1))
```

```
> a2< -predict(arima02,a.nhead="1")
```

```
> xx< -a[1:2961]
```

```
> arima1 <- Arima(xx,order = c(1, 1, 1))
```

```
> a3< -predict(arima02,a.nhead="1")
```

- Invertir la transformación de Box-Cox

```
> pronosticos< -InvBoxCox(variable,0)
```

- Diagnóstico del modelo (Cuadro 4.10 y cuadro 4.20)

```
> accuracy(datos originales,pronósticos)
```

- Series de tiempo de ambas variables (Figura 4.33)

```
>plot(cobre.ts,type="l",main=,ylim=c(0,3100),xlab="Tiempo(años)",ylab="Precio (CLP/lb)")
```

```
>lines(dolar.ts,col="green")
```

```
>legend(2006,3150,c("Cobre","Dólar"),lty=c(1,1),lwd=c(0.1,0.1),
```

```
col=c("black","green"))
```

```
>grid()
```

- Causalidad de Granger

```
> GrangerTest(a,p=1,include.mean = F)
```

- Ajustes modelos VARMA

```

> A< - log.cobre.ts
> B< -log.dolar.ts
> a=cbind(A,B)
> AA=ts(a,start=2006,freq=365)
> varma1=VARMA(AA,p=1,q=1,include.mean = FALSE)
> varma2=VARMA(AA,p=2,q=1,include.mean = FALSE)
> varma3=VARMA(AA,p=1,q=2,include.mean = FALSE)
> varma4=VARMA(AA,p=2,q=2,include.mean = FALSE)
> varma5=VARMA(AA,p=3,q=1,include.mean = FALSE)
> varma6=VARMA(AA,p=1,q=3,include.mean = FALSE)
> varma7=VARMA(AA,p=2,q=3,include.mean = FALSE)
> varma8=VARMA(AA,p=3,q=2,include.mean = FALSE)
> varma9=VARMA(AA,p=3,q=3,include.mean = FALSE)
> varma10=VARMA(AA,p=4,q=1,include.mean = FALSE)
> varma11=VARMA(AA,p=4,q=2,include.mean = FALSE)
> varma12=VARMA(AA,p=4,q=3,include.mean = FALSE)
> varma13=VARMA(AA,p=4,q=4,include.mean = FALSE)

```

- Pronósticos

```

D1< - log.cobre.ts[1:2960]
C1< - log.dolar.ts[1:2960]
s1=(cbind(D1,C1))
q1=ts(s1,start=2006,freq=365)
varm1=VARMA(q1,p=1,q=2,include.mean = FALSE)
z1< -VARMApred(varm1,h=1)
D1< - dcobre[1:2961]
C1< - ddolar[1:2961]
s1=(cbind(D1,C1))
q1=ts(s1,start=2006,freq=365)
varm1=VARMA(q1,p=1,q=2,include.mean = FALSE)
z2< -VARMApred(varm1,h=1)
D1< - dcobre[1:2962]
C1< - ddolar[1:2962]
s1=(cbind(D1,C1))
q1=ts(s1,start=2006,freq=365)
varm1=VARMA(q1,p=1,q=2,include.mean = FALSE)
z3< -VARMApred(varm1,h=1)
D1< - dcobre[1:2963]

```

```
C1<- ddolar[1:2963]
s1=(cbind(D1,C1))
q1=ts(s1,start=2006,freq=365)
varm1=VARMA(q1,p=1,q=2,include.mean = FALSE)
```

- Comparación de datos reales y pronósticos (Figura 4.14.)

```
>plot(pronosreal[2800:3287,],type="l",xlab="Datos",ylab="Precio dolar")
>lines(pronosreal[2959:3287,],type = "l",col=3)
>legend(70,570,c("Datos originales", "Pronóstico"),lty=c(1,1),
lwd=c(0.1,0.1),col=c("black","red"))
```