



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias Naturales Matemáticas

“Comparación de métodos de predicción Univariado y
Multivariados para series de tiempo”

INFORME DE PROYECTO INTEGRADOR

Previa a la obtención del Título de:

INGENIERO EN ESTADISTICA INFORMÁTICA

Ronny Xavier Merchán Ponce

GUAYAQUIL – ECUADOR

AÑO: 2016

Agradecimientos

Mis más sinceros agradecimientos a mi familia especialmente a mi abuela mis amigos, M.Sc. Gaudencio Zurita el cual me dio la oportunidad de formar parte de su plantilla de ayudantes, M.Sc Margarita Martinez por hacerme parte del mejor proyecto en cual trabaje durante mi vida estudiantil, Phd. Francisco Vera y su esposa la M.Sc. Sonia Reyes por su excelente cátedra en distintas materias y su apoyo personal en algunos momentos.

Dedicatoria

El presente proyecto lo dedico a las personas que me apoyaron a lo largo de esta travesía mis tíos Carlos Ponce(+) , Plinio Rodríguez (+) , mi abuela Esther Rodríguez los cuales fueron participe dándome su apoyo en todo momento a mi mama Patricia Ponce a mis amigos Ivan Aveiga , Nathalie Tejena , Michelle Fajardo los cuales fueron mis baluartes en la estadía en la universidad , a mi novia, y a las personas que confiaron en mí.



ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL

Facultad de Ciencias Naturales y Matemáticas

Ingeniería en Estadística Informática

INFORME DEL PROFESOR DE LA MATERIA INTEGRADORA

Habiendo sido nombrado PROFESOR DE LA MATERIA INTEGRADORA del señor

Ronny Xavier

Merchán Ponce

Con el tema del proyecto integrador Comparación de métodos de predicción Univariado y Multivariados para series de tiempo, previa a la obtención del título de Ingeniero en Estadística Informática, me permito informar que he leído el contenido y he revisado el formato del proyecto integrador, luego de lo cual indico que estoy de acuerdo en que el mismo se lo ha desarrollado conforme a los lineamientos de la Unidad de Titulación Especial de la ESPOL.

M.Sc. Sonnia Paola Reyes Ramos

PROFESOR DE LA MATERIA INTEGRADORA

Guayaquil, 25 octubre de 2016

Tribunal de evaluación

.....
Nombre del Profesor

Francisco Xavier Vera Alcívar Ph.D

.....
Nombre del Profesor

M.Sc Edison Isais del Rosario Campusano

Declaración expresa

"La responsabilidad y la autoría del contenido de este Trabajo de Titulación, me corresponde exclusivamente; y doy mi consentimiento para que la ESPOLE realice la comunicación pública de la obra por cualquier medio con el fin de promover la consulta, difusión y uso público de la producción intelectual"

.....
Ronny Xavier Merchán Ponce

Resumen

La planificación que hacen los gobiernos en los países a nivel mundial se ha caracterizado en las últimas décadas en utilizar predicciones estadísticas, para con ello poder obtener una buena aproximación en diferentes variables económicas que son importantes para el desarrollo de un país, y así tener una mejor planificación y control de la economía.

Por ello surge la necesidad de utilizar modelos de predicción para poder estimar con menor error los índices de desarrollo del país para aquello se utilizaran modelos estadísticos Univariantes Modelos Auto regresivos (AR) y Multivariantes Auto regresivos Vectoriales (VAR) para series de tiempo, los cuales mediante una serie de técnicas se podrá obtener criterios los que servirán para proporcionar pronósticos y obtener resultados de tal manera que se compararan y el resultado final será compara los dos métodos estadísticos y obtener el más eficiente.

Por consecuencia las comparaciones se realizaran con respecto a las variables que sean de interés para el presente documento y demostrar resultados de cada uno de los métodos así se escoge el de mayo eficiencia al pronosticar.

Contenido

Agradecimientos	ii
Dedicatoria.....	iii
Tribunal de evaluación.....	v
Declaración expresa	vi
Resumen	vii
<i>Introducción</i>	xi
Capítulo 1	xii
1. DESCRIPCIÓN DEL PROBLEMA.....	xii
1.1 Introducción	xii
1.2 Justificación del proyecto	xii
1.3 Objetivos del proyecto	12
1.3.1 Objetivos Generales	12
1.3.2 Objetivos específicos.....	2
Capítulo 2	3
2 Metodología	3
2.1 Procesos estocásticos en series de datos Univariantes	3
2.2 Modelo Arima (Auto Regresivos Integrados Media Móvil)	4
2.3 Modelos Ar(p)	4
2.4 Series de datos Multivariantes definición.....	5
2.4.1 Estacionaridad	6
2.4.5 Linealidad	7
2.4.6 Invertibilidad en Series de tiempo Multivariantes	8
2.4.7 Cross covarianza y correlación de matrices en series de tiempo multivariantes.....	9
2.4.8 Modelos Var (Vectores Auto- regresivos)	10
2.4.9 Test de causalidad de Granger	11
Capítulo 3	13
3 Análisis estadísticos y resultados	13

3.2 Estadísticas descriptivas y gráficos	13
3.2.1 Masa monetaria (Moneda local actual) MMA(MLA).....	13
3.2.2 Gasto de consumo final del gobierno general (US\$ a precios actuales)	15
3.2.3 Gasto final del consumo de los hogares, etc. (US\$ a precios actuales)	17
3.4 Reconocimiento del Modelos ARIMA	19
3.4.1 Masa monetaria (Moneda local actual) MMA(MLA).....	19
3.4.2 Gasto de consumo final del gobierno general (US\$ a precios actuales)	19
3.4.3 Gasto de consumo final del gobierno general (US\$ a precios actuales)	20
3.5 reconocimiento del modelo VAR	20
3.5.1 Reconocimiento del modelo	20
3.5.2 Test de Causalidad.....	21
3.5.3 Resumen del Modelo Var(2).....	22
3.7 Resultados de la predicción.....	24
3.7.1 Predicción para las series modeladas con ARIMA(1,10).....	24
3.7.2 Predicción en el modelo Multivariante Vectorial Auto Regresivo...	25
3.8 Comparaciones de los pronósticos y datos reales	26
3.8.1 Masa monetaria (moneda local actual) MMA (MLA)	26
3.8.2 Gasto de consumo final del gobierno general (US\$ a precios actuales) GCFGG (PRA)	26
3.8.3 Gasto final del consumo de los hogares, etc. (US\$ a precios actuales) GFCHO (PRA)	26
4 Conclusiones	27
Recomendaciones	27
Referencias.....	28

Índice de tablas

Tabla 1 Tabla de frecuencias MMA(MLA).....	14
Tabla 2 Estadísticos Descriptivos	14
Tabla 3 Tabla de frecuencias (GCFGG)	15
Tabla 4 Estadísticos Descriptivos (GCFGG).....	16
Tabla 5 Tabla de frecuencias (GCH)	17
Tabla 6 Estadísticos Descriptivos (GCH).....	18
Tabla 7 Comparación de predicciones.....	26
Tabla 8 Comparación de predicciones.....	26
Tabla 9 Comparación de predicciones.....	26

Índice de gráficos

Ilustración 1.....	13
Ilustración 3.....	14
Ilustración 2.....	14
Ilustración 4.....	15
Ilustración 6.....	16
Ilustración 5.....	16
Ilustración 7.....	17
Ilustración 8.....	18
Ilustración 9.....	18

Introducción

El ambiente de la Economía global, se ha generado mucho impacto en la última década por los impactos que ha causado la crisis económica en el ámbito de decisiones de gobierno se necesita optimizar los recursos de cada país.

Para ello se basan en pronósticos para poder hacer la planificación anual con respecto a los factores de ingresos como el del PIB (Producto Interno Bruto) entre otros que puedan optimizar los recursos del país, por tales motivos se ha pretendido estudiar que modelos de pronósticos son más robustos al momento de predecir estas variables .

Se trabaja con datos históricos de series de tiempo de índole económica para poder comparar los modelos Univariantes y Multivariantes así poder tomar la decisión de cuál es el más óptimo para la toma de decisiones.

El estudio de modelos de pronósticos tiene como propósito el establecer comparación entre los modelos Univariantes y Multivariantes, con el fin de verificar la validez de las predicciones en función de las variables Masa Monetaria Gasto Gubernamentales, Gasto consumo final de hogar.

Los métodos estadísticos aplicados se fundamentan en los modelos de Box – Jenkins y los de modelos de Christopher Sins, en el ámbito informático se utilizó el software R, la combinación de métodos y fundamentos teóricos permitieran obtener predicciones para las comparaciones de los modelos

Capítulo 1

1. DESCRIPCIÓN DEL PROBLEMA

1.1 Introducción

En este primer capítulo se va a dar a conocer una breve introducción de los modelos Univariantes y Multivariantes, conceptos objetivos y el alcance del proyecto en general; se detallaran los aspectos teóricos relacionados con la metodología a utilizar así con las variables con las cuales se van a trabajar adicionalmente se explicara un poco el desarrollo de los modelos.

1.2 Justificación del proyecto

El proyecto se propone analizar las diferentes formas de trabajar con series de tiempo de tipo económico, a fin de poder analizar estos datos de indicadores económicos del país, estableciendo comparaciones y proponiendo predicciones a base de criterios óptimos y estimaciones para la toma de decisiones idóneas en el contexto financiero y gubernamental.

1.3 Objetivos del proyecto

1.3.1 Objetivos Generales

Determinar un modelo estadístico apropiado para la comparación de predicciones que permitirá escoger que modelo se aproxima a la realidad, así se optimizara recursos económicos en el contexto gubernamental para futuros estudios.

Para ello utilizaremos el software R que nos provee de paquetes específicos para poder hacer este tipo de estudios.

1.3.2 Objetivos específicos

- Proponer modelos estadísticos univariante y multivariante estableciendo comparaciones de pronósticos.
- Analizar los modelos Univariante y Multivariante, para la aplicación de acuerdo al grado y variantes respectivos.
- Establecer comparaciones con los modelos estadísticos ARIMA y VAR, estableciendo pronósticos en función de las aproximaciones a los datos reales.

Capítulo 2

2 Metodología

Las variables fueron obtenidas de la página del banco mundial mediante sus índices por países de ellas se escogió las siguientes para nuestro estudio las cuales son de gran importancia para la dinámica de una nación en ámbitos económicos e sociológicos.

Masa monetaria (moneda local actual) MMA (MLA)

Gasto de consumo final del gobierno general (US\$ a precios actuales) GCFGG (PRA)

Gasto final del consumo de los hogares, etc. (US\$ a precios actuales) GFCHO (PRA)

Los dos tipos de métodos que se plantean son el ARIMA y VAR para comparar las predicciones de cada uno.

2.1 Procesos estocásticos en series de datos Univariantes

Los procesos estocásticos son estacionarios en un sentido dicho estricto o fuerte cuando la distribución de probabilidad conjunta de la serie es invariante con respecto al tiempo.

$$F(x_t, x_{t+1}, \dots, x_{t+k}) = F(x_{t+k+1}, x_{t+k+2}, \dots, x_{t+k+s})$$

Para poder trabajar una serie de datos univariante se tiene que cumplir que se un proceso estocásticos con ruidos, cotidianamente se lo hace para calcular los retardos hasta llegar a esta propiedad para las series de tiempo, los siguientes supuestos nos indican que sea un proceso estocástico.

1. $E(a_t) = \mu_a$ donde $\mu < \infty \quad \forall t$
2. $Var(a_t) = \sigma_a$ donde $\sigma < \infty \quad \forall t$
3. $Cov[(a_t a_{t+k})] = 0 \quad |a| < \infty$

Autocovarianza y autocorrelacion

$$\gamma_k = \begin{cases} \sigma_a^2 & k=0 \\ 0 & k \neq 0 \end{cases}$$

$$\rho_k = \begin{cases} 1 & k=0 \\ 0 & k \neq 0 \end{cases}$$

$$\phi_{kk} = \begin{cases} 1 & k=0 \\ 0 & k \neq 0 \end{cases}$$

2.2 Modelo Arima (Auto Regresivos Integrados Media Móvil)

Una de las metodologías más reconocidas en el ámbito de la modelación matemática de series de tiempo es la ARIMA la cual fue prescrita y formalizada por un matemático y un químico llamados George E. P. Box y Gwilym Jenkins en el año de 1976 por ello la notación de estos modelos se les denomina modelos Box-Jenkins

Estos modelos se fundamentan en la teoría de los procesos estocásticos.

El nombre de estos modelos ARIMA procede de las siglas en inglés de Autorregresive, integrated and moving average. Su significado es Modelos integrados (I) autorregresivos (AR) y de medias móviles (MA).

Este enfoque parte del hecho que la serie temporal que se trata de predecir es generada por un proceso estocástico o aleatorio cuya naturaleza puede ser caracterizada mediante un modelo.

Esta aproximación para que los datos y los modelos sigan un procesos estocásticos para modelos ARIMA.

Notación de Box Jenkins :

ARIMA(p,d,q)

p: Número de parámetros autorregresivos.

d: Número de diferenciaciones para que la serie sea estacionaria.

q: Número de parámetros de medias móviles.

Modelo completo:

$$(1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p)(1 - B)^d y_t = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) a_t$$

2.3 Modelos Ar(p)

Un modelo principal de los ARIMA, es la parte autorregresiva con el cual vamos a trabajar estos tienen una particularidad de proceso en las que sus observaciones en el momento inicial son predecibles a partir de observaciones pasadas más un término de error

El proceso autorregresivo de orden p representado por ARIMA(P,0,0) o simplemente AR(p) se denota de la siguiente manera :

$$AR(p) \equiv \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \dots + \Phi_p y_{t-p} + a_t = y_t$$

Mediante el operador de retardo puede escribirse de la siguiente manera :

$$(1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p) y_t = a_t \quad B^p(y_t) = y_{t-p}$$

Un proceso AR(p) es estacionario si las raíces del polinomio de B $(1 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p)$ caen fuera de la unidad esto equivale a que las raíces de la ecuación $x^p - \Phi_1 x^{p-1} - \Phi_2 x^{p-2} - \dots - \Phi_{p-1} x^{p-k} = 0$ sean todas diferentes en modulo a la unidad.

2.4 Series de datos Multivariantes definición

En el análisis de las series de datos multivariantes se considera series de tiempo de forma simultánea, esta rama de la estadística multivariada que se ocupa específicamente de datos dependientes.

En general este tipo de series se analiza con menos complejidad con respecto a las series de tiempo univariantes, específicamente cuando hay un número importante de observaciones dentro de la serie que se vaya a analizar, por ejemplo los datos de Economía, Ingeniería que se recolectan siempre deben estar en intervalo de tiempos tales como horas semanas meses etc.

El análisis estadístico para esta clase de series de tiempo a nivel multivariado en la vida real a menudo se pueden comprometer otros factores que están interrelacionados con las variables estos factores pueden ser aleatorios por ejemplo una baja inesperada del precio del petróleo, pero además si tuviéramos el mando de estos factores y tuviéramos la comprensión de los mismo, residieramos en obtener predicciones óptimas para la toma de decisiones a niveles gerenciales o multinacionales en caso que se trabajara en cualquier ambiente de incertidumbre. Los objetivos principales de los modelos de series de tiempo multivariadas son:

- El estudio de las relaciones dinámicas entre las variables, estas variables pueden ser contemporáneamente relacionadas, cualquier de estas series puede conducir a las otras en la relación de las otras antecesoras.
- Mejora de la exactitud de predicción. Cuando la información histórica de la serie puede estar contenida en otros tipos de datos, la predicción o el resultado mejora cuando la serie se es modelada conjuntamente.

En el mundo de las estadísticas multivariantes, se habla de k dimensiones, en la parte de series de tiempo con características multivariadas se hace una suposición

que se tiene k variables aleatorias (dimensiones) y_{it} , donde $i = 1, 2, \dots, k$ que son las componentes de un vector y_t .

$$y_t = [y_{1t}, y_{2t}, \dots, y_{kt}]'$$

Para simplicidad de trabajar con este vector de datos tenemos la notación y_t , cuando tratamos las propiedades de y_t lo tratamos como un vector aleatorio si se tiene en cuenta una aplicación se puede mencionar el vector de la realización de y_t , para nuestro caso del proyecto se asume que el y_t es un vector de series de tiempo continuo multivariado y con distribución de probabilidad continua eso indica que el vector aleatorio discreto o categóricos de una serie de tiempo no se tomarían en cuenta, debido a que se trabaja con matrices y vectores aleatorios los cuales se trabajaran a lo largo de este proyecto.

En esta orientación, de un enfoque multivariante significaría una mayor aproximación a una realidad, en tanto que se pudiese capturar las interacciones dinámicas entre las series, pero en cualquiera de los casos, los cimientos sobre los que se construye dicho planteamiento descansan en los establecidos para el modelo univariante. Este último se puede generalizar fácilmente al caso multivariante, en el que se analiza un vector $N \times 1$ de variables observables.

$$y_t = [y_{1t}, y_{2t}, \dots, y_{kt}]'$$

2.4.1 Estacionaridad

Un vector de series de tiempo con k – dimensiones y_t se dice que es débilmente estacionario cuando.

- $E(y_t) = \mu$ (vector de constantes con k dimensiones)
- $Cov(y_t) = E[(y_t - \mu)(y_t - \mu)'] = \Sigma_y$ una matriz definida positiva de constantes $k \times k$.

En este caso $E(y_t)$ y $Cov(y_t)$ denotados por una expresión de un prospecto de matriz de varianza y covarianza del vector aleatorio y_t , por lo tanto la media y la matriz de covarianza son débilmente estacionarias ya que los datos de las series de tiempo no dependen del tiempo, que es decir los dos primeros momentos de la serie de tiempo y_t no varían en el tiempo. Implícitamente en la definición, para el caso de las series de tiempos multivariantes nosotros requerimos que la media y la matriz de covarianzas necesariamente deben ser débilmente estacionarios y deben existir.

Una serie de tiempo y_t con k dimensiones es estrictamente estacionario si la distribución conjuntan de la recopilación de m variables $(y_{t1}, y_{t2}, \dots, y_{tm})$ es la

misma que $(y_{t_1+j}, y_{t_2+j}, \dots, y_{t_m+j})'$, (t_1, t_2, \dots, t_m) , donde m, j , son arbitrarios positivos y enteros.

En términos estadísticos, el decir que el vector de series temporales y_t es estrictamente estacionario requiere que se requiere que la distribución de probabilidad de las serie de tiempo multivariante y_t no pueda variar en el tiempo es decir la serie es independiente del tiempo.

Un ejemplo de una serie de tiempo multivariantes estrictamente estacionaria es un vector idénticamente distribuidas con una población que se modelo con una distribución normal estándar multivariada.

Para casos prácticos se podrá observar los dicho en las definiciones anteriores.

2.4.5 Linealidad

Para la concepción del proyecto y la breve introducción de las series de tiempo, toparemos el caso de la linealidad de las series de tiempo, para el caso de estudio hemos dotado de una cierta rigurosidad en las series de tiempo multivariantes hablando en el sentido estricto, pero en realidad las series de tiempo multivariantes no son lineales en ese sentido, pero los modelos lineales pueden a menudo proporcionar una buena información precisa para las aproximaciones de linealidad para poder hacer inferencias. Una serie de tiempo y_t con K dimensiones es lineal si:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i a_{t-i} \quad (1.1)$$

Donde μ es el vector de medias constantes, cuando i es igual 0, $\psi_0 = I_k$ la cual es la matriz identidad de dimensiones $k \times k$, $\psi_i (i > 0)$ es una matriz de constantes de dimensión $k \times k$, y a_t es la secuencias de valores independientes aleatorios e idénticamente distribuidos con media 0 y matriz de covarianza definida positiva Σ_a .

Se requiere para nosotros que Σ_a sea definida positiva, por otra parte que la dimensión de k puede ser reducida con un análisis de componentes principales que es un método multivariantes de reducción de dimensión la condición $\psi_0 = I_k$ se satisface porque nosotros permitimos que Σ_a sea una matriz general definida positiva.

Una buena alternativa aproximada a una expresión lineal en series de tiempo requiere que ψ_0 sea una matriz triangular superior y que los elementos en la diagonal sean 1 y Σ_a sea una matriz diagonal.

Esto se puede realizar mediante la descomposición de Cholesky para Σ_a la descomposición de la matriz de varianzas y covarianzas $\Sigma_a = LGL'$, donde G es una matriz diagonal y L sea una matriz triangular superior y que los elementos en la diagonal sean 1.

Si $b_t = L^{-1}a_t$ entonces $a_t = Lb_t$ y:

$$\text{Cov}(b_t) = \text{Cov}(L^{-1}a_t) = L^{-1}\Sigma_a(L^{-1})' = L^{-1}(LGL')(L^{-1})' = G$$

Con la secuencia b_t la primera ecuación se puede reescribir

$$y_t = \mu + \sum_{i=0}^{\infty} (\psi_i L) b_{t-i} = \mu + \sum_{i=0}^{\infty} \psi_i^* b_{t-i}$$

Donde $\psi_0^* = L$, la cual es una matriz triangular superior, $\psi_i^* = \psi_i L$ para los $i > 0$ y la matriz de covarianza de b_t es una matriz diagonal.

Para un proceso estacionario puro de y_t , la descomposición de *Wold* afirma que se puede escribir como una combinación lineal de una secuencia de procesos incorrelados e_t .

Esto es cercano pero no idéntico a la ecuación 1.1 porque e_t no debe tener necesariamente una distribución.

2.4.6 Invertibilidad en Series de tiempo Multivariantes

En muchas situaciones por ejemplo para el pronóstico, se es cotidiano expresar las series de tiempo y_t como una función para sus valores rezagados y_{t-i} para $i > 0$ para más información del tiempo t .

Una serie de tiempo y_t se dice que es invertible si se puede escribir de la siguiente manera :

$$y_t = c + a_t + \sum_{i=0}^{\infty} \pi_j y_{t-j}$$

Donde c es un vector de constantes con k dimensiones, a_t se definió anteriormente en ecuación 1.1 y π_j una matriz de dimensión $k \times k$.

Es obvio decir que un ejemplo de invertibilidad de las series de tiempo se aplica en los modelos Var(1) (Vector Auto regresivos de orden 1), es decir $y_t = c + \pi_j y_{t-j} + a_t$, se discute las condiciones de invertibilidad después, para el estudio de las series de tiempo multivariantes se mencionara que es suficiente decir que es y_t es invertible si $\pi_j \rightarrow 0$ y $j \rightarrow \infty$

2.4.7 Cross covarianza y correlación de matrices en series de tiempo multivariantes.

Al medir la linealidad dinámica e dependencia de una serie de tiempo estacionaria definimos ℓ como retrasos de la matriz de cross covarianza como:

$$\Gamma_\ell = Cov(y_t, y_{t-\ell}) = E[(y_t - \mu)(y_{t-\ell} - \mu)']$$

$$\Gamma_\ell = \begin{bmatrix} E(\hat{y}_{1t}\hat{y}_{1,t-\ell}) & E(\hat{y}_{1t}\hat{y}_{2,t-\ell}) & \dots & E(\hat{y}_{1t}\hat{y}_{k,t-\ell}) \\ \vdots & \vdots & & \vdots \\ E(\hat{y}_{1t}\hat{y}_{1,t-\ell}) & E(\hat{y}_{kt}\hat{y}_{2,t-\ell}) & \dots & E(\hat{y}_{kt}\hat{y}_{k,t-\ell}) \end{bmatrix}$$

Donde se dice que $\mu = E(y)$ es el vector de media de y_t y $\hat{y}_t = (\hat{y}_{1t}, \dots, \hat{y}_{kt})' \equiv y_t - \mu$ es la media ajusta de las series de tiempo.

Esta matriz de cross covarianza esta en función ℓ no del tiempo, esto se da porque y_t es estacionario. Para un valor de $\ell = 0$ se obtiene la matriz Γ_0 de y_t , en algunos casos la notación cambia a Σ_y que se denota como la matriz de covarianzas de y_t entonces $\Sigma_y = \Gamma_0$.

2.4.8 Modelos Var (Vectores Auto-regresivos)

Descripción del modelo Vectorial Auto Regresivo

Los modelos a utilizar será los VAR (Vectores Auto-Regresivos) los cuales son una generalización de los modelos AR (Arima con parte auto regresiva) de series de tiempo univariantes para el desarrollo de este tipo de series de tiempo multivariantes para así poder trabajar con mayor cantidad de datos y optimizar el tiempo para los análisis de los mismo y una ventaja es mejorar la calidad de la predicción en el capítulo de los modelos Var analizaremos profundamente esta herramienta.

En las series de datos multivariantes siendo $y_t = [y_{1t}, y_{2t}, \dots, y_{kt}]'$ el vector de datos en series de tiempo estos datos se pueden modelar con el modelo de Vector auto regresivo de orden p o mejor representado mediante Var(p) se tienen la siguiente notación:

$$y_t = \phi_0 + \sum_{i=0}^{\infty} \phi_i y_{t-i} + a_t \quad [1]$$

Dónde:

ϕ_0 es un vector de constantes de k dimensiones.

ϕ_i es una matriz de kxk dimensiones de coeficientes auto regresivos para valores de $i > 0$, $\phi_i \neq 0$.

a_t es el vector aleatorio que es idénticamente distribuido con media 0 y Σ_a , esta matriz de varianzas y covarianzas es definida positiva como se definió anteriormente.

$$Var(p) = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + a_t \quad [2]$$

A esto llamamos un modelo Var generalizado con algo de efectos contemporáneos entre las variables.

$$By_t = \phi_0^* + \sum_{i=0}^{\infty} \phi_i^* y_{t-i} + a_t^* \quad [3]$$

Es condición suficiente que B sea no singular para que pueda transformarse en la ecuación inicial se puede multiplicar B^{-1} el modelo es llamado modelo Var estructural [3] y el principal [1] se dice que está en forma reducida.

Si se utiliza los operadores de retardo como en los modelos univariantes L , la ecuación [1] se puede escribir como:

$$(L + \phi_1 L^1 - \phi_2 L^2 - \dots - \phi_p L^p) y_t = c + a_t$$

$$\Phi(L) y_t = c + a_t$$

Para cuestiones de aprendizaje se empieza por un ejemplo el modelo bivariado Var (1)

$$y_t = \phi_0 + \phi_1 y_{t-1} + a_t$$

Este modelo se escribe de forma explícita:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \phi_{10} \\ \phi_{20} \end{bmatrix} + \begin{bmatrix} \phi_{1,11} & \phi_{1,12} \\ \phi_{1,21} & \phi_{1,22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} a_{1t} \\ a_{2t} \end{bmatrix}$$

O es equivalente al sistema de ecuaciones

$$y_{1t} = \phi_{10} + \phi_{1,11} y_{1,t-1} + \phi_{1,12} y_{2,t-1} + a_{1t}$$

$$y_{2t} = \phi_{20} + \phi_{1,21} y_{1,t-1} + \phi_{1,22} y_{2,t-1} + a_{2t}$$

Estos elementos que contienen la secuencias de ecuaciones con respecto a los coeficientes de ϕ_1 tal como $\phi_{1,12}$ se presenta como la dependencia lineal de y_{1t} sobre $y_{2,t-1}$ con la presencia de $y_{1,t-1}$, de igual el elemento de $\phi_{1,21}$ la medida lineal representada por la relación entre y_{2t} y $y_{1,t-1}$ en la presencia de $y_{2,t-1}$ los otros parámetros de ϕ_1 se pueden definir de igual manera.

2.4.9 Test de causalidad de Granger

Si se tiene dos variables Y_t X_t que representan dos series en el tiempo t y Y_{t-p} X_{t-p} representa el tiempo de los p retardos se puede decir que dos variables están cointegradas según granger si una combinación lineal de estas variables que sea:

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

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

Entonces, se dice que las variables Y , X están cointegradas si el mismo error cuadrático medio de la predicción X_{t-p} , Y_{t-p} si se trabaja con funciones lineales significaría que :

$$Mse[E(x_{t+s} | x_t \dots x_{t-1})] = Mse[E(x_{t+s} | x_t \dots x_{t-1} \dots, y_t \dots y_{t-1})]$$

Para mejor comprensión esto indica que :

x_t es exógena fuerte respecto a los parámetros de la condicional y_t

Por tanto la incógnita , ¿ x_t puede predecir y_t ?

Es un punto clave si el VAR es un enfoque el cual permite ver si en la predicción por cada variable de los rezagos de las son útiles, por oposición a los modelos univariados.

Capítulo 3

3 Análisis estadísticos y resultados

3.1 Introducción

En este capítulo se desarrollara la estadística descriptiva y los métodos gráficos de los datos de panel para las series de tiempo que se han tomado como caso de estudio para la aplicabilidad de los modelos Arima y Autoregresivos.

Para realizar los gráficos y las estadísticas descriptivas se necesitó ayuda de los paquetes ggplot2, vars, forecast, que se encuentran incluidos en el software R V 3.3.2 con el cual se trabajara el proyecto.

3.2 Estadísticas descriptivas y gráficos

3.2.1 Masa monetaria (Moneda local actual) MMA(MLA)

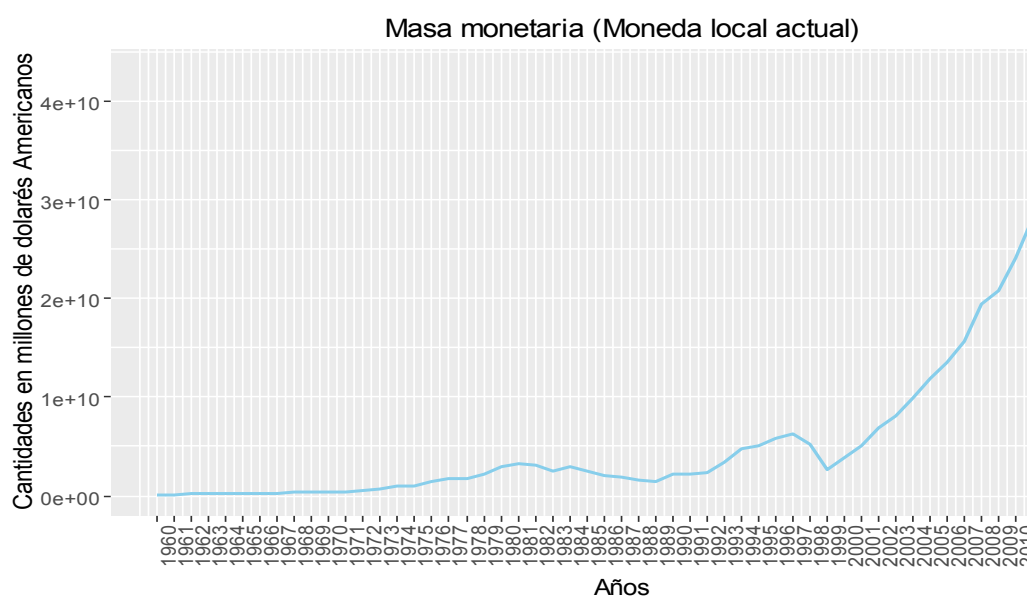


Ilustración 1

Figura 3.2.1.1 Diagrama de series de tiempo MMA(MLA).

En este primer grafico se ilustra la serie de tiempo de la masa monetaria, desde el año 1960 hasta el 2010, donde se puede observar un pico desde el año 2000, fecha en que inició la dolarización que estuvo ligada a un salto económico muy grande para la economía del país.

Intervalos	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
(1.11e+08,4.94e+09]	37	37	0.72549020	0.4509804
(4.94e+09,9.75e+09]	7	44	0.13725490	0.7843137
(9.75e+09,1.46e+10]	3	47	0.05882353	0.8823529
(1.46e+10,1.94e+10]	1	48	0.01960784	0.9411765
(1.94e+10,2.42e+10]	3	51	0.05882353	10.000.000

Tabla 1 Tabla de frecuencias MMA(MLA)

Mínimo	772564021
Máximo	4,4012E+10
Cuartil1	2446544276
Cuartil2	1,0204E+10
Cuartil3	1,7694E+10
Media	1,2251E+10
Varianza	1,19E+26
Desviación	1,0907E+10

Tabla 2 Estadísticos Descriptivos

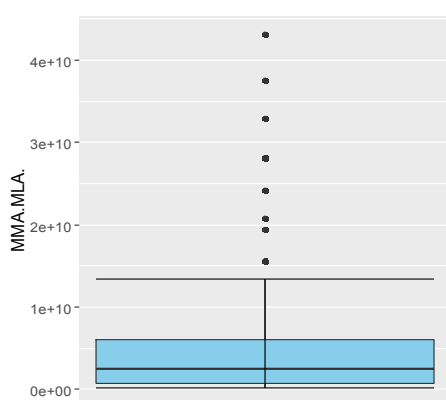


Ilustración 2

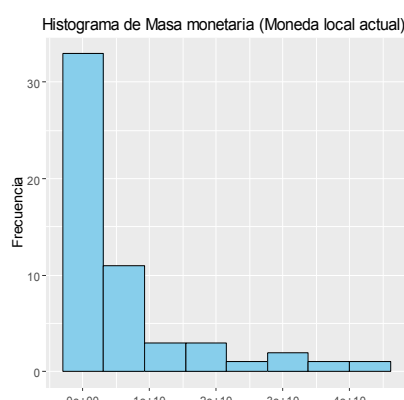


Ilustración 3

Figura 3.2.1.2 Diagrama de Cajas

Figura3.2.1.3 Histograma

Este cuadro resume estadísticas básicas de Masa Monetaria en dólares que demuestran un cambio durante 54 años, evidenciando que el promedio fue de $1,22 \cdot E+10$ millones de dólares con un pico máximo $4,4012 \cdot E+10$. Se observa que la variabilidad de los datos con respecto a la media es muy grande el valor es $1,19E+26$.

3.2.2 Gasto de consumo final del gobierno general (US\$ a precios actuales)

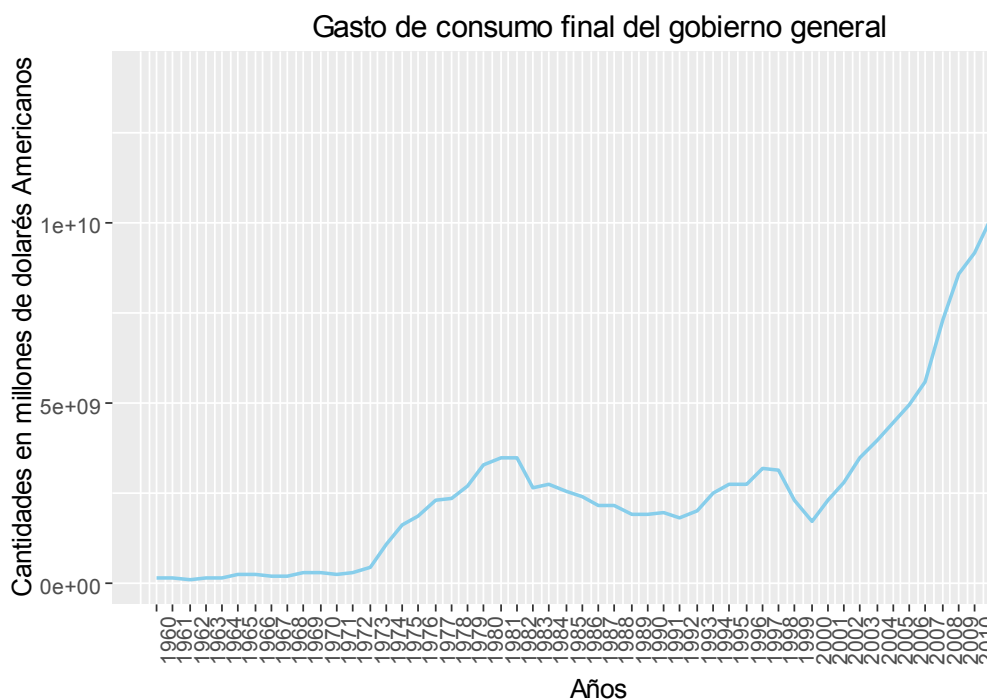


Ilustración 4

Figura 3.2.2.1 Diagrama de series de tiempo.

El grafico se ilustra la serie de tiempo del Gasto de consumo final del Gobierno, desde el año 1960 hasta el 2010. A partir del 2006 hasta el 2010 se observa un incremento considerable en base a esta variable, esto corresponde a la administración del actual gobierno y el incremento se puede percibir que es exponencial, esto se observa en la gráfica.

Intervalos	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
(8.81e+07,1.91e+09]	21	21	0.41176471	0.4117647
(1.91e+09,3.73e+09]	23	44	0.45098039	0.8627451
(3.73e+09,5.55e+09]	3	47	0.05882353	0.9215686
(5.55e+09,7.36e+09]	2	49	0.03921569	0.9607843
(7.36e+09,9.19e+09]	2	51	0.03921569	10.000.000

Tabla 3 Tabla de frecuencias (GCFGG)

Mínimo	9,72E+07
Máximo	9,18E+09
Cuartil1	3,57E+08
Cuartil2	2,31E+09
Cuartil3	2,96E+09
Media	2,35E+09
Varianza	4,19E+24
Desviación	2,05E+09

Tabla 4 Estadísticos Descriptivos (GCFGG)

Este cuadro de estadísticas básicas del consumo final del Gobierno describe el cambio que se ha dado durante los diferentes mandatos, ya que su promedio en dólares americanos fue de $2,35 \times 10^9$ millones de dólares con un pico máximo $9,811 \times 10^9$ y un mínimo de $9,72 \times 10^7$. Se observa que la varianza es muy grande, lo que indica que hay mucha dispersión con media su valor es $4,19 \times 10^{24}$.

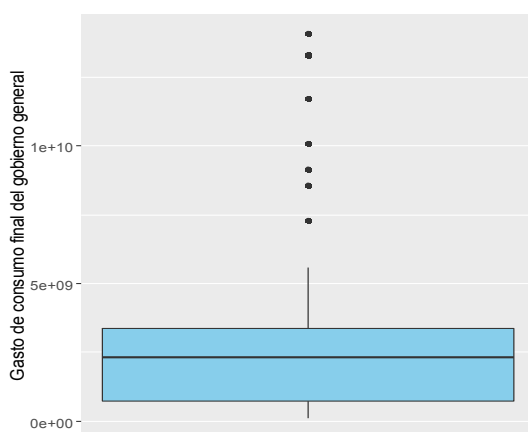


Ilustración 6

Figura 3.2.2.2 Diagrama de Cajas

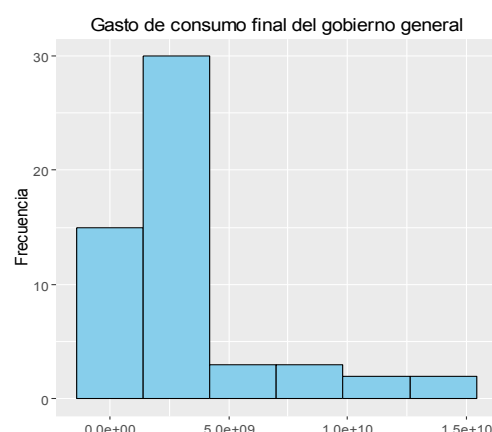


Ilustración 5

Figura 3.2.2.3 Histograma

3.2.3 Gasto final del consumo de los hogares, etc. (US\$ a precios actuales)

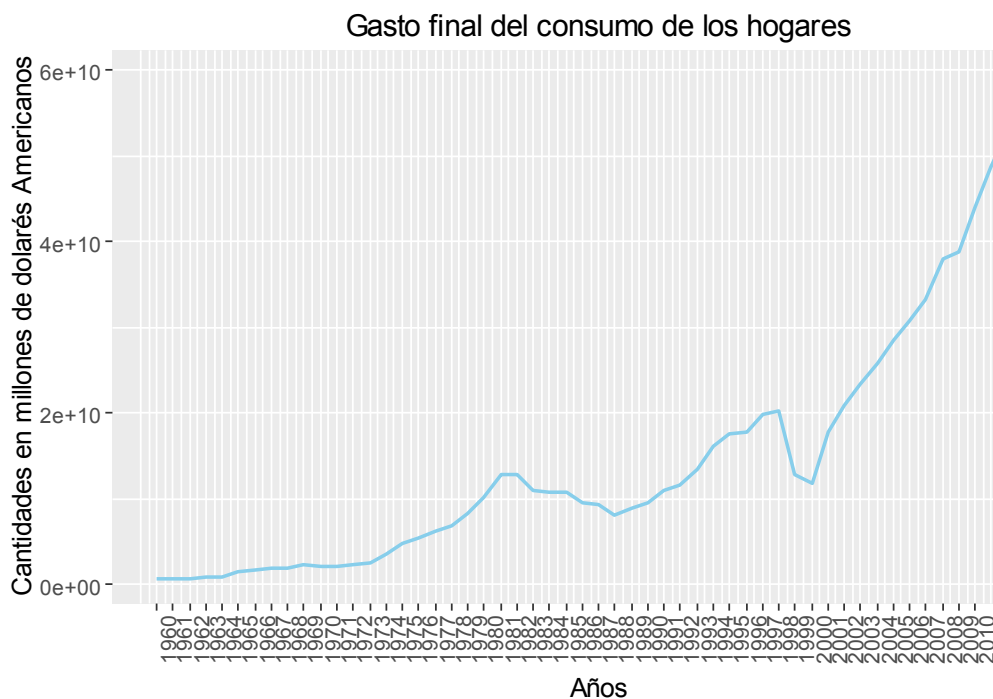


Ilustración 7

Figura 3.2.3.1 Diagrama de series de tiempo.

En este primer grafico se ilustra la serie de tiempo del Gasto de consumo de los hogares desde el año de 1960 hasta el 2014 el cual se puede observar un incremento a finales de los años 80 y un pico descendiente en los años 1999 y 2000.

Intervalos	Frecuencia Relativa	Frecuencia Acumulada	Frecuencia Relativa	Frecuencia Relativa Acumulada
(7.29e+08,9.42e+09]	23	23	0.45098039	0.4509804
(9.42e+09,1.81e+10]	17	40	0.33333333	0.7843137
(1.81e+10,2.67e+10]	5	45	0.09803922	0.8823529
(2.67e+10,3.54e+10]	3	48	0.05882353	0.9411765
(3.54e+10,4.41e+10]	3	51	0.05882353	10.000.000

Tabla 5 Tabla de frecuencias (GCH)

Mínimo	7,73E+08
Máximo	4,40E+10
Cuartil1	2,45E+09
Cuartil2	1,02E+10
Cuartil3	1,77E+10
Media	1,23E+10
Varianza	1,19E+26
Desviación	1,09E+10

Tabla 6 Estadísticos Descriptivos (GCH)

Este cuadro nos resume las estadísticas básicas, las cuales nos puede describir cómo hubo cambio del Gasto de consumo final del Gasto final de hogares a lo largo del tiempo, su promedio en dólares americanos fue de $1,23 \times 10^{10}$ millones de dólares con un pico máximo $4,40 \times 10^{10}$ se observa la varianza es muy grande lo cual indica que hay mucha dispersión con respecto a la media.

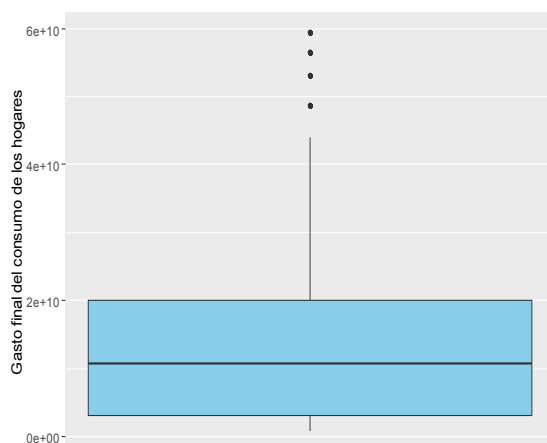


Ilustración 8

Figura 3.2.3.2 Diagrama de Cajas

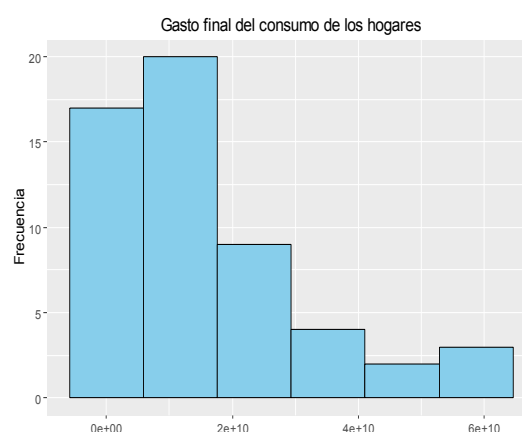


Ilustración 9

Figura 3.2.3.3 Histograma

3.4 Reconocimiento del Modelos ARIMA

En el caso de estudio de nuestras las series de tiempo utilizaremos los criterios básicos para elegir un modelo óptimo para trabajar, comenzando con los métodos gráficos convencionales, en este caso el correspondiente a los de Auto correlación simple y parcial.

3.4.1 Masa monetaria (Moneda local actual) MMA(MLA)

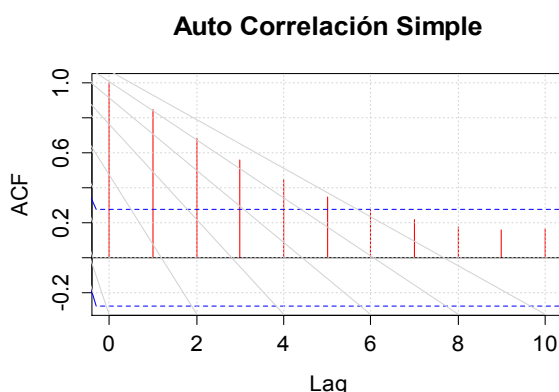


Figura 3.4.1 Auto correlación Simple

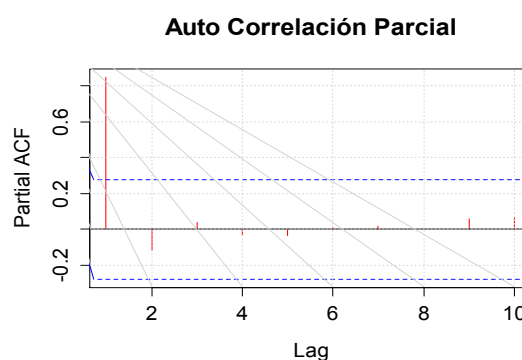


Figura3.4.2 Auto correlación Parcial

Mediante el análisis grafico de las correlaciones simples y parciales se puede observar que tienen una caída exponencial descendiente, de lo cual se deduce que la serie se pueda modelar con un $ARIMA(1,1,0)$.

3.4.2 Gasto de consumo final del gobierno general (US\$ a precios actuales)

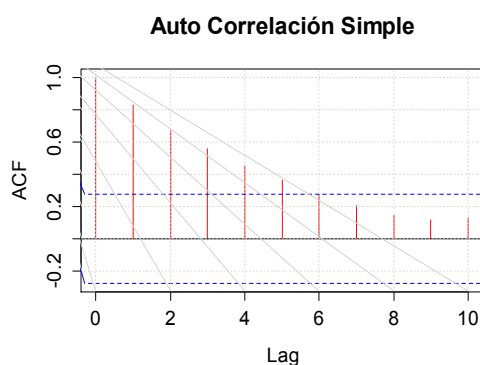


Figura 3.4.2.1 Auto correlación Simple

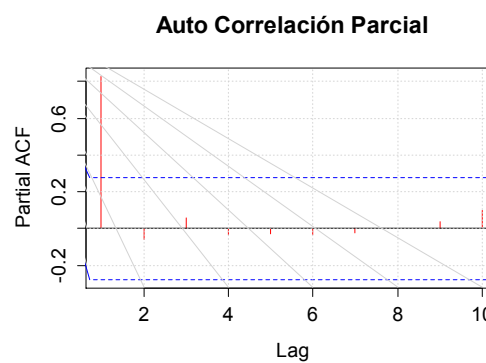


Figura3.2.2.1Auto correlación Parcial

Mediante el análisis grafico de las correlaciones simples y parciales se puede observar que tienen una caída exponencial descendiente, de lo cual se deduce que la serie se pueda modelar con un $ARIMA(1,1,0)$.

3.4.3 Gasto de consumo final del gobierno general (US\$ a precios actuales)

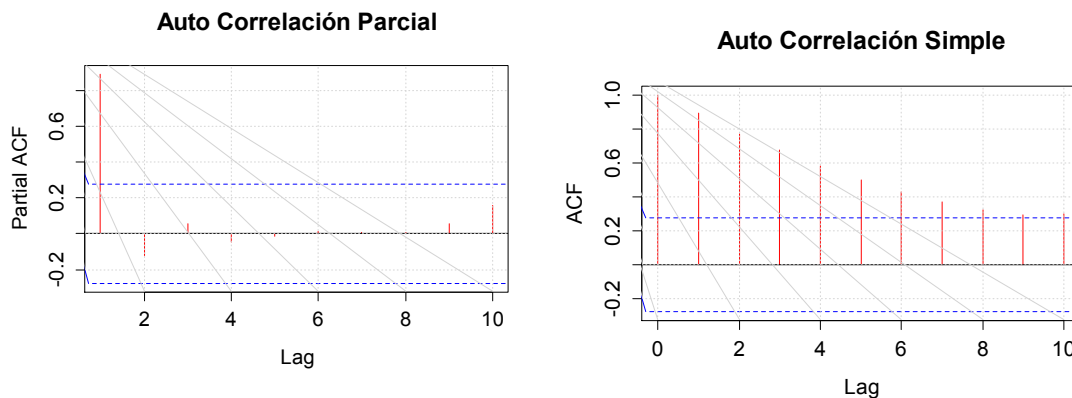


Figura 3.4.3.1 Auto correlación Simple Figura3.4.3.2Auto correlación Parcial

Mediante el análisis gráfico de las correlaciones simples y parciales se puede observar que tienen una caída exponencial descendiente, de lo cual se deduce que la serie se pueda modelar con un ARIMA(1,1,0) .

3.5 reconocimiento del modelo VAR

3.5.1 Reconocimiento del modelo

Los criterios que se utilizará para elegir los retardos óptimos del modelo Vectorial Autoregresivo son:

- Criterio de información de Akaike (AIC)
- Hannan y Quinn (HQ)
- Bayesiana de Schwarz (SC)
- Error de predicción final de Akaike (FPE)

Estos criterios son los más importantes utilizados para este tipo de modelos.

Con la ayuda del software R y el paquete vars() se calcula de manera más rápida los criterios, ya que este incluye un tipo de técnicas de elección de modelos; para aquello se propone un número de lags o retardos, sin embargo el programa calcula el que ajusta a nuestro modelo, el granger test o test de causalidad se aplica para saber si es factible trabajar con las variables propuestas en el proyecto.

```

VARselect(diff(series1),lag.max=8,type="const")
$selection
AIC(n)  HQ(n)  SC(n) FPE(n)
      8      2      2      2

$criteria
      1      2      3      4
5      6
AIC(n) 1.215760e+02 1.211679e+02 1.214132e+02 1.214354e+02 1.214700e
+02 1.209689e+02
HQ(n)  1.217586e+02 1.214875e+02 1.218697e+02 1.220290e+02 1.222006e
+02 1.218364e+02
SC(n)  1.220775e+02 1.220456e+02 1.226670e+02 1.230654e+02 1.234762e
+02 1.233512e+02
FPE(n) 6.318317e+52 4.235880e+52 5.522487e+52 5.864458e+52 6.466830e
+52 4.322774e+52
      7      8
AIC(n) 1.210734e+02 1.207194e+02
HQ(n)  1.220779e+02 1.218609e+02
SC(n)  1.238319e+02 1.238540e+02
FPE(n) 5.559959e+52 4.843566e+52

```

Se puede observar que los criterios son unificados excepto el AIC el cual indica que el óptimo es 8, estas pruebas estadísticas nos dan una sola dirección que es de tomar como máximo retardo 2, para cuestiones de estudio el SC y el HQ proporcionan estimaciones muy consistentes del verdadero orden del lag o retardo mientras que el FPE y AIC sobreestiman el orden del retardo con probabilidad positiva. Analizando todo esto observaremos que el retardo óptimo es 2.

3.5.2 Test de Causalidad

Granger test

```

grangertest(diff(MMA.MLA.) ~ diff(GCFGG.PRA.),order=1,data=DatosEcón
omicos1)
Granger causality test

Model 1: diff(MMA.MLA.) ~ Lags(diff(MMA.MLA.), 1:1) + Lags(diff(GCFG
G.PRA.), 1:1)
Model 2: diff(MMA.MLA.) ~ Lags(diff(MMA.MLA.), 1:1)
  Res.Df Df      F Pr(>F)
1     50
2     51 -1  0.5522 0.4609

```



```

grangertest(diff(diff(MMA.MLA.)) ~ diff(diff(GFCH.PRA.)),order=2,data=DatosEconomicos1)
Granger causality test

Model 1: diff(diff(MMA.MLA.)) ~ Lags(diff(diff(MMA.MLA.)), 1:2) + Lags(diff(diff(GFCH.PRA.)), 1:2)
Model 2: diff(diff(MMA.MLA.)) ~ Lags(diff(diff(MMA.MLA.)), 1:2)
  Res.Df Df      F Pr(>F)
1      46   0.000000 0.999999
2      48  -2  2.1997 0.1224

```

Aplicamos el Test de Granger de causalidad para formular la hipótesis, que consiste en indicar si la variable (a) tiene un impacto en la variable (b) a lo largo del tiempo. El valor p de la prueba indica que la hipótesis es aceptada, lo que implica que hay una incidencia de la variable (a) con respecto a la (b) a lo largo del tiempo; en consecuencia podemos afirmar que este par de variables son óptimas para trabajar con el modelo Vectorial Autoregresivo Var (p).

3.5.3 Resumen del Modelo Var(2)

```

summary(Var2)

VAR Estimation Results:
=====
Endogenous variables: MMA.MLA., GCFGG.PRA., GFCH.PRA.
Deterministic variables: const
Sample size: 44
Log Likelihood: 84.039
Roots of the characteristic polynomial:
0.9548 0.8276 0.8276 0.2546 0.1181 0.1181
Call:
VAR(y = series0, p = 2, type = "const")

Estimation results for equation MMA.MLA.:
=====
MMA.MLA. = MMA.MLA..l1 + GCFGG.PRA..l1 + GFCH.PRA..l1 + MMA.MLA..l2
+ GCFGG.PRA..l2 + GFCH.PRA..l2 + const

              Estimate Std. Error t value Pr(>|t|)
MMA.MLA..l1  1.175183   0.222567   5.280 5.92e-06 ***
GCFGG.PRA..l1 -0.003446   0.254883  -0.014  0.989
GFCH.PRA..l1  0.037770   0.401630   0.094  0.926
MMA.MLA..l2  -0.315753   0.257840  -1.225  0.228
GCFGG.PRA..l2 -0.010005   0.251649  -0.040  0.969
GFCH.PRA..l2  0.113341   0.425557   0.266  0.791
const        -0.086808   1.606829  -0.054  0.957

```

```

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2192 on 37 degrees of freedom
Multiple R-squared: 0.9725,    Adjusted R-squared: 0.9681
F-statistic: 218.4 on 6 and 37 DF,  p-value: < 2.2e-16

Estimation results for equation GCFGG.PRA.:
=====
GCFGG.PRA. = MMA.MLA..l1 + GCFGG.PRA..l1 + GFCH.PRA..l1 + MMA.MLA..l
2 + GCFGG.PRA..l2 + GFCH.PRA..l2 + const

              Estimate Std. Error t value Pr(>|t|)
MMA.MLA..l1   0.80304    0.16884   4.756 2.99e-05 ***
GCFGG.PRA..l1 1.11551    0.19336   5.769 1.29e-06 ***
GFCH.PRA..l1 -0.50190    0.30468  -1.647  0.1080
MMA.MLA..l2  -0.45194    0.19560  -2.311  0.0265 *
GCFGG.PRA..l2 -0.16960    0.19090  -0.888  0.3801
GFCH.PRA..l2  0.09721    0.32283   0.301  0.7650
const         2.89389    1.21895   2.374  0.0229 *
---Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1663 on 37 degrees of freedom
Multiple R-squared: 0.9824,    Adjusted R-squared: 0.9796
F-statistic: 344.6 on 6 and 37 DF,  p-value: < 2.2e-16

Estimation results for equation GFCH.PRA.:
=====
GFCH.PRA. = MMA.MLA..l1 + GCFGG.PRA..l1 + GFCH.PRA..l1 + MMA.MLA..l2
+ GCFGG.PRA..l2 + GFCH.PRA..l2 + const

              Estimate Std. Error t value Pr(>|t|)
MMA.MLA..l1   0.547369   0.128626   4.256 0.000137 ***
GCFGG.PRA..l1 0.055512   0.147302   0.377 0.708431
GFCH.PRA..l1  0.652302   0.232110   2.810 0.007865 **
MMA.MLA..l2  -0.420605   0.149011  -2.823 0.007621 **
GCFGG.PRA..l2 -0.007256   0.145433  -0.050 0.960475
GFCH.PRA..l2  0.106112   0.245938   0.431 0.668637
const         1.822604   0.928617   1.963 0.057232 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1267 on 37 degrees of freedom
Multiple R-squared: 0.9863,    Adjusted R-squared: 0.984
F-statistic: 442.5 on 6 and 37 DF,  p-value: < 2.2e-16

Covariance matrix of residuals:

```

	MMA.MLA.	GCFGG.PRA.	GFCH.PRA.
MMA.MLA.	0.04804	0.01638	0.01884
GCFGG.PRA.	0.01638	0.02765	0.01249
GFCH.PRA.	0.01884	0.01249	0.01604

Correlation matrix of residuals:

	MMA.MLA.	GCFGG.PRA.	GFCH.PRA.
MMA.MLA.	1.0000	0.4495	0.6787
GCFGG.PRA.	0.4495	1.0000	0.5932
GFCH.PRA.	0.6787	0.5932	1.0000

Es importante destacar en el resumen las ecuaciones estructurales para el Var(2) se observa como en cada una de las posibles ecuaciones causales se encuentran las significancia de cada de variable en cada modelo para nuestro caso elegiremos la que tiene por variables endógena Masa Monetaria que es la primera ecuación estructural. Se observa como la matriz de correlaciones o cross - covarianza de los residuos para este tipo de variables se encuentran muy bien correlacionados lo que en la teoría indican que tienen una muy buena linealidad con respecto al tiempo.

3.7 Resultados de la predicción

En este capítulo se presentarán los resultados de la aplicación de los modelos Univariantes y Multivariantes para hacer las comparaciones respectivas de las metodologías.

Por otra parte para estos modelos se utilizó transformaciones logarítmicas para suavizar las series de tiempo para poder estimar mejor la predicción.

Los resultados de la predicción de las variables utilizadas para el presente proyecto son presentadas a continuación:

3.7.1 Predicción para las series modeladas con ARIMA(1,10)

Masa monetaria (moneda local actual) MMA (MLA)

PredicciónMMA					
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011	23.26499	22.98827	23.54171	22.84179	23.68820
2012	23.28698	22.82190	23.75205	22.57571	23.99825
2013	23.29469	22.67612	23.91326	22.34866	24.24071
2014	23.29739	22.54983	24.04496	22.15409	24.44070

Gasto de consumo final del gobierno general (US\$ a precios actuales) GCFGG (PRA)

PredicciónGFGGP

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011		22.26668	22.00279	22.53058	21.86309	22.67028
2012		22.29002	21.82317	22.75688	21.57603	23.00402
2013		22.30075	21.65868	22.94281	21.31879	23.28271
2014		22.30567	21.51210	23.09924	21.09201	23.51933

Gasto final del consumo de los hogares, etc. (US\$ a precios actuales) GFCHO (PRA)

PredicciónGFCH.PRA.

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011		24.11481	23.91746	24.31216	23.81299	24.41664
2012		24.13452	23.78707	24.48196	23.60315	24.66589
2013		24.14336	23.66711	24.61962	23.41500	24.87173
2014		24.14734	23.56008	24.73459	23.24921	25.04547

3.7.2 Predicción en el modelo Multivariante Vectorial Auto Regresivo.**ModeloVar2****MMA.MLA.**

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011		23.23896	23.95807	23.51985	24.80938	23.66855
2012		24.33820	24.80182	23.67457	24.57082	24.90558
2013		24.23407	24.69074	23.77740	24.40312	25.06502
2014		24.23102	24.60669	23.85536	24.27619	25.18586

GCFGG.PRA.

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011		23.40328	23.19020	23.61636	23.07740	23.72916
2012		23.50846	23.10884	23.90808	21.89729	23.11963
2013		22.57459	22.02709	23.12209	21.73726	23.41193
2014		22.62378	21.96510	23.28246	21.61642	23.63115

GFCH.PRA.

	Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2011		24.15720	23.99487	24.31953	23.90894	24.40546
2012		24.17837	23.88416	24.47259	23.72841	24.62833
2013		24.18999	23.80258	24.57740	23.59750	24.78248
2014		24.20078	23.74301	24.65855	23.50068	24.90088

3.8 Comparaciones de los pronósticos y datos reales

3.8.1 Masa monetaria (moneda local actual) MMA (MLA)

	ORIGINAL	VAR(2)	ARIMA(1,1,0)
2011	2,82E+10	1,24E+10	1,27E+10
2012	3,30E+10	3,71E+10	1,30E+10
2013	3,76E+10	3,70E+10	1,31E+10
2014	4,32E+10	3,34E+10	1,31E+10

Tabla 7 Comparación de predicciones

3.8.2 Gasto de consumo final del gobierno general (US\$ a precios actuales) GCFGG (PRA)

	ORIGINAL	VAR(2)	ARIMA(1,1,0)
2011	1,01E+10	1,4585E+10	1,4393E+10
2012	1,17E+10	1,6203E+10	1,3973E+10
2013	1,33E+10	1,7311E+10	1,3164E+10
2014	1,41E+10	1,8183E+10	1,3229E+10

Tabla 8 Comparación de predicciones

3.8.3 Gasto final del consumo de los hogares, etc. (US\$ a precios actuales) GFCHO (PRA)

	ORIGINAL	VAR(2)	ARIMA(1,1,0)
2011	4,87E+10	3,0998E+10	2,9712E+10
2012	5,30E+10	3,1662E+10	3,0303E+10
2013	5,64E+10	3,2032E+10	3,0572E+10
2014	5,95E+10	3,2379E+10	3,0694E+10

Tabla 9 Comparación de predicciones

4 Conclusiones

La selección de las variables constituyó un importante sustento para la validez de los métodos estadísticos, ya que a partir de este insumo se determinó que los datos se ajustan con precisión al modelo VAR(2), cabe recalcar que esta prueba es muy importante para poder aplicar este tipo de modelo estadístico multivariado. El test de granger para series de tiempo multivariante es fundamental para este tipo de estudio.

Una vez desarrolladas las actividades propuestas para cada objetivo y los resultados que los modelos arrojaron se puede establecer que:

1. Se constató que el modelo VAR tuvo mayor precisión con respecto al uso de variables exógenas que permitieron la aproximación a los datos reales.
2. Con relación al modelo ARIMA se constató que los pronósticos entre años no presentan mayor variabilidad. Se confirmó que este modelo es de tendencia y tolera ajustes.
3. Cabe señalar que el modelo VAR obtuvo mejores pronósticos con relación al ARIMA, tal como se demuestra en las tablas citadas.

Recomendaciones

Tomar las variables validadas y el modelo VAR para distintos tipos de escenarios como el demográfico y seguridad social.

Analizar detenidamente el modelo VAR para obtener más información acerca del comportamiento de las series de tiempo y los saltos económicos con más técnicas que pueden tener estos modelos.

Divulgar la trascendencia de aplicar los modelos estadísticos para apreciar la realidad de los datos.

Referencias

- [1] Vector Autoregressive Models for Multivariate Time Series Modeling Financial Time Series with S-PLUS® 2006 Erick Zivot Jiahui Wang
- [2] Multivariate Time Series Analysis With R and Financial Applications RUEY S. TSAY Booth School of Business University of Chicago Chicago, ILEditorial Editorial Wiley
- [3] <http://www.banrep.gov.co/es/tags/modelos-var>
- [4] http://www.inegi.org.mx/eventos/2013/RDE_07/RDE_07_Art2.html
- [5] http://www.scielo.org.co/scielo.php?pid=S0121-68052013000200003&script=sci_arttext
- [6] <http://revistas.ucr.ac.cr/index.php/matematica/issue/view/2040>
- [7] <http://robjhyndman.com/uwafiles/Class10.R>
- [8] http://rum.prf.jcu.cz/public/mecirova/eng_ekonomka/William_H_Greene-Econometric_Analysis-Prentice.pdf
- [9] http://dinamicasistemas.otalca.cl/Revista/Vol2Num1/Alvarez_Crespo_et_al_SVAR.pdf
- [10] http://www.carloscapistran.com/uploads/2/9/5/7/29578885/series_de_tiem po_estacionarias_multivariadas_190810.pdf