Centro de Investigación en Matemáticas (CIMAT). Unidad Monterrey

Econometria: Tarea 4

Jorge Luis Ramos Zavaleta

6 de junio de 2019

1. Ejercicio

En los Estados Unidos del país de las maravillas, las tazas de crecimiento para el ingreso (GNP) y la demanda de dinero (M2) y la tasa de interés (IR) estan relacionadas siguiendo un modelo VAR(2):

$$\begin{pmatrix} GNP_t \\ M2_t \\ IR_t \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.7 & 0.1 & 0 \\ 0 & 0.4 & 0.1 \\ 0.9 & 0 & 0.8 \end{pmatrix} \begin{pmatrix} GNP_{t-1} \\ M2_{t-1} \\ IR_{t-1} \end{pmatrix} + \begin{pmatrix} -0.2 & 0 & 0 \\ 0 & 0.1 & 0.1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} GNP_{t-2} \\ M2_{t-2} \\ IR_{t-2} \end{pmatrix} + \begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{pmatrix}$$

donde

$$\Sigma_Z = \begin{pmatrix} 0.26 & 0.03 & 0 \\ 0.03 & 0.09 & 0 \\ 0 & 0 & 0.81 \end{pmatrix} = PP', \quad P = \begin{pmatrix} 0.5 & 0.1 & 0 \\ 0 & 0.3 & 0 \\ 0 & 0 & 0.9 \end{pmatrix}$$

1. Muestra que el proceso $X_t = (GNP_t, M2_t, IR_t)'$ es estable.
2. Determina el vector de medias de X_t.
3. Escribe el proceso de X_t en la forma $\boldsymbol{X_t}$, VAR(1).

1.1. Solución

Para reducirnos el problema iniciaremos estableciendo el modelo en su forma VAR(1), el cual queda de la siguiente manera

$$\boldsymbol{X_t} = \begin{pmatrix} GNP_t \\ M2_t \\ IR_t \\ GNP_{t-1} \\ M2_{t-1} \\ IR_{t-1} \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.7 & 0.1 & 0 & -0.2 & 0 & 0 \\ 0 & 0.4 & 0.1 & 0 & 0.1 & 0.1 \\ 0.9 & 0 & 0.8 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} GNP_{t-1} \\ M2_{t-1} \\ IR_{t-1} \\ GNP_{t-2} \\ M2_{t-2} \\ IR_{t-2} \end{pmatrix} + \begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

La condición de establilidad nos indica que debemos calcular el polinomio característico para la matriz de coeficientes y ver que sus raíces sean mayores que uno en modulo.

$$0 = \begin{vmatrix} 1-z0.7 & -0.1 & 0 & 0.2 & 0 & 0 \\ 0 & 1-z0.4 & -0.1 & 0 & -0.1 & -0.1 \\ -0.9 & 0 & 1-z0.8 & 0 & 0 & 0 \\ -1 & 0 & 0 & -z & 0 & 0 \\ 0 & -1 & 0 & 0 & -z & 0 \\ 0 & 0 & 1 & 0 & 0 & -z \end{vmatrix} = 1-1.19z+1.26z^2-0.323z^3-0.021z^4+0.06z^5$$

El cual tiene raíces

```
raicesPol
[1]  0.4295698+0.9197368i  0.4295698-0.9197368i  2.9726606-1.4857032i
[4]  2.9726606+1.4857032i -5.4919608+0.0000000i
```

y sus modulos son:

```
[1] 1.015109 1.015109 3.323255 3.323255 5.491961
```

Los cuales son todos mayores que uno, por lo que podemos concluir que nuestro proceso es estable. Por último para calcular el vector de medias se tienen las siguiente identidades

$$\boldsymbol{\mu} = E[\tilde{\boldsymbol{X}}_t] = (\boldsymbol{I_{kp}} - \tilde{\boldsymbol{\phi}})^{-1} v$$

$$E[\boldsymbol{X_t}] = \boldsymbol{J}\boldsymbol{\mu},\ con\ \boldsymbol{J} = (\boldsymbol{I_k}, 0, ..., 0)$$

Asi

$$\boldsymbol{\mu} = \begin{pmatrix} 0.3 & -0.1 & 0 & 0.2 & 0 & 0 \\ 0 & 0.6 & -0.1 & 0 & -0.1 & -0.1 \\ -0.9 & 0 & 0.2 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 6.8750 \\ 14.3750 \\ 30.9375 \\ 6.8750 \\ 14.3750 \\ 30.9375 \end{pmatrix}$$

Luego $E[\boldsymbol{X_t}] = \boldsymbol{J\mu} = (6.8750, 14.3750, 30.9375)$

2. Ejercicio

Muestra el procedimiento de cómo presentar las funciones de respuesta-impulso de manera ortogonalizada

2.1. Solución

La idea de las funciones de respuessta-impulso ortogonalizadas aparece debido a que la matriz de covarianzas de las ecuaciones a usar en un modelo VAR son positivas, esto implica que no podemos asumir que al establecer el impulso de una de las variables las otras permanezcan constante por lo que debemos establecer una forma de separar los efectos.

Para ello el método de las funciones de respuesta-impulso ortogonalizadas descomponen la matriz de covarianzas estimada $\hat{\Sigma}$ en un matriz triangular inferior. Con esto se puede aislar la respuesta de una de las variables por un impulso en la misma ecuación.

Para estimar las funciones de impulso-respuesta ortogonalizadas, generamos la descomposición de Cholesky de la matriz de covarianza $\boldsymbol{\Sigma = PP'}$ con $\boldsymbol{P}$ una matriz triangular inferior por construcción. Ahora definimos un vector u_t de px1, considerando un modelo VAR(p), tal que $\boldsymbol{Pu_t = \epsilon_t}$ lo que implica que $\boldsymbol{u_t = P^{-1}\epsilon_t}$. Por construcción los errores en $\boldsymbol{u_t}$ no estan correlaciones pues $E[\boldsymbol{u_t}, \boldsymbol{u_t}] = \boldsymbol{P^{-1}}E[\epsilon_t\epsilon_t^{-1}]\boldsymbol{P'^{-1}} = \boldsymbol{I_p}$. Entonces ahora las funciones de respuesta-impulso ortogonalizadas son simplemente

$$\frac{\partial y_{i,t+h}}{\partial u_{j,t}} = \{\boldsymbol{\Phi}_h\boldsymbol{P}_{i,j}\}$$

es decir, el producito de los coeficientes de las matrices del proceso $MA(\infty)$ y la matriz triangular inferior P.

3. Ejercicio

Modifica el experimento Monte Carlo de tal forma que observes el funcionamiento muestral de los diferentes estadısticos que se analizaron en clase. Modifica a tu interés los siguientes parámetros: coeficientes del modelo VAR, rezagos, estructuras de covarianza en los errores y tamaño de muestras y comenta lo hallado. Para cada uno de ellos prueba dos diferentes casos.

3.1. Solución

Primero iniciaremos modificando los coeficientes del modelo VAR, inicialmente se tenian coeficientes generados por una distribución U(-0.5,0.5). Ajustaremos los coeficientes disminuyendo los intervarlos a (-0.25,0.25) y (-0.1,0.1). Considerando los coeficientes en U(-0.25,0.25) obtenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.524 0.110 0.366
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.994  0.954
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
   0.39    0.53
```

Considerando ahora los coeficientes en U(-0.1,0.1) tenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.196 0.006 0.062
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.988  0.978
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
  0.100   0.096
```

Ahora regresaremos a tener los coeficientes en U(-0.5,0.5) y modificaremos el numero de rezagos primero a 3 y luego a 4 para ver su comportamiento. Estableciendo k=3 tenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.940 0.998 0.996
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.994  0.968
```

```
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
  1.000   0.998
```

Con k=4 tenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.954 0.920 0.998
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.996  0.988
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
      1       1
```

Originalmente se consideran un tamaño de muestra de 500, primero lo disminuiremos a 300 y después lo aumentaremos a 1,000 para ver el cambio que se genera. Regresando a k=3 y disminuyendo la muestra a 300 tenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
      AIC        SC        HQ
0.8300000 0.9566667 0.9633333
> colSums(resultados[,c("serial", "arch")] > alpha)/R
   serial      arch
0.9900000 0.9533333
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
  Series1   Series2
0.9966667 0.9933333
```

Ahora aumentando a 1,000 en el tamaño de muestra tenemos

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
```

```
0.580 0.150 0.411
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.992  0.956
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
  0.648   0.226
```

Por último vamos a cambiar la estructura de la distribución de los errores, inicialmente se tiene una distribución normal estándar, ahora usaremos una distribución t con 3 grados de libertad y una distribución t con 5 grados de libertad. Usando 3 grados de libertad

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.846 0.968 0.952
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.988  0.902
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
  0.998   0.970
```

y usando 5 grados de libertad

```
> # observamos los resultados
> sum(resultados[,"maxroot"] < 1)/R
[1] 1
> colSums(resultados[, c("AIC", "SC", "HQ")] == p)/R
  AIC    SC    HQ
0.856 0.988 0.976
> colSums(resultados[,c("serial", "arch")] > alpha)/R
serial   arch
 0.994  0.938
> colSums(resultados[,paste("Series",1:p, sep = "")] < alpha)/R
Series1 Series2
  0.928   0.988
```

Primero cabe observar que aun con cambios la máxima raíz que aparece en todos los casos siempre es la misma. En el caso de los rezagos obtenidos por los criterios de información se puede observar que restringiendo los coeficientes a un intervalo mas pequeño se requieren menores óptimos, de igual manera al aumentar el tamaño de la muestra. Mientras que no se observan cambios significativos cambiando el número de rezagos y ni cambiando la estructura de covarianza de los errores.

Ahora para el caso de los p valores de las pruebas de correlación serial y efecto tenemos que no se observan cambios que pudieran parecer significativos en ninguno de los casos, ya que se trata de variaciones muy pequeñas. Por último para el p valor de las pruebas de causalidad de Granger tenemos una analogia con el número de rezagos usando los criterios de información disminuciones en los p valores con intervalos de coeficientes mas restringidos y con mayor tamaño de muestra, aunque los p valores en estos casos sufren una disminución bastante significativa. Por otro lado modificando el numero de rezagos y la estructura de covarianza no se observan cambios significativos en los p valores de la prueba de causalidad de Granger.

4. Ejercicio

Construye un modelo VAR cuyo objetivo sea analizar y pronosticar la inflación interanual de México para el siguiente periodo no disponible. Elige un periodo de muestra que consideres apropiado y realiza las pruebas que consideres necesarias para verificar que el modelo pronosticará bien a futuro. Puedes utilizar variables relacionadas endógenamente (según la teorıa económica) como oferta de billetas y monedas (M0), tipo de cambio nominal, tasa de desempleo y salarios reales. Alguna variable de tu sugerencia es bienvenida.

4.1. Solución

Para este problema se usaron 3 variables: INPC mensual por niveles con base 2018, la tasa de desocupación y la oferta de dinero (M0). Se usaron series mensuales a partir de enero de 2005, esto debido a que este es un periodo de relativa calma para estas variables macroeconómicas, se habían considerado otras variables como el tipo de cambio interbancario a 24 horas y los salarios reales reportados por el IMSS.

Sin embargo ambas series presentaban problemas específicos, el tipo de cambio tenia un dato no calculado (Agosto 2015) y se eligió hacer una interpolación usando los 5 datos anteriores. Los salarios reales antes de 2015 vienen establecidos como un promedio entre las zonas económicas por lo que lo hace una serie no tan práctica para realizar predicciones cuando se consideran muchos datos antes del 2015. No se consideraron las 2 variables dado que consideraron por separado y en forma conjunta con las otras variables elegidas pero no lograron pasar por mucho la prueba de Portmentau por lo que los errores residuales nos indicaban que el modelo no sería sufcientemente bueno para la predicción.

Para iniciar se consideraron los logaritmos de los datos debido a los cambios de escalas, a continuación a dichas series se les realizó una prueba de estacionariedad y se encontró que las series son estacionarias de primer orden como se puede observar a continuación

```
Augmented Dickey-Fuller Test
data:  diff(datos2$M0)
Dickey-Fuller = -10.792, Lag order = 1, p-value = 0.01

data:  diff(datos2$INPC)
Dickey-Fuller = -6.8948, Lag order = 1, p-value = 0.01

data:  diff(datos2$Desocupacion)
Dickey-Fuller = -14.053, Lag order = 1, p-value = 0.01
```

y el algoritmo reporta que se tienen en realidad p valores menores que 0.01. Para continuar se realizaron algunos graficos de ACF y PDF de las series en su versión estacionaria.

Se puede observar cierta tendencia en el dato, asi como cierta estacionalidad. Por lo que se modelara el VAR como un modelo con estacionalidad anual y con tendencia. Graficamos las series en un mismo grafico para ver su comportamiento como se muestra en la figura 4.4.

Después se busco el rezago óptimo para generar nuestro modelo VAR con tendencia, e inicialmente se considero el criterio de información HQ(n) para ello con un número de 3 rezagos, pero debido a que el criterio AIC muestra hasta 25 rezagos se consideraron diversos modelos para los que se realizó la prueba de Portmentau para ver si sus residuales formaban ruido blanco, de acuerdo a este criterio se considero usar 5 rezagos debido a que el p valor reportado de esta prueba es el segundo mas alto con 0.2023 y nos permitiría concluir de mejor manera que no existe correlación serial y a su vez nos permite encontrar relaciones con la causalidad de Granger.

```
AIC(n)  HQ(n)  SC(n) FPE(n)
    25      3      1      4

Portmanteau Test (asymptotic)

data:  Residuals of VAR object var1
Chi-squared = 195.58, df = 180, p-value = 0.2023
```

Sin embargo algo que no se menciono es que las series se recortaron hasta el mes de agosto debido a que las otras variables no contaban con este dato, debido a esto al ajustar varios modelos el que mejor predicción daba para la inflación interanual de septiembre es uno que contaba con 12 rezagos. Por lo que se continuo el análisis con este nuevo modelo. Un problema que se encontró es que se tenia un p valor muy bajo pero suficiente para pasar

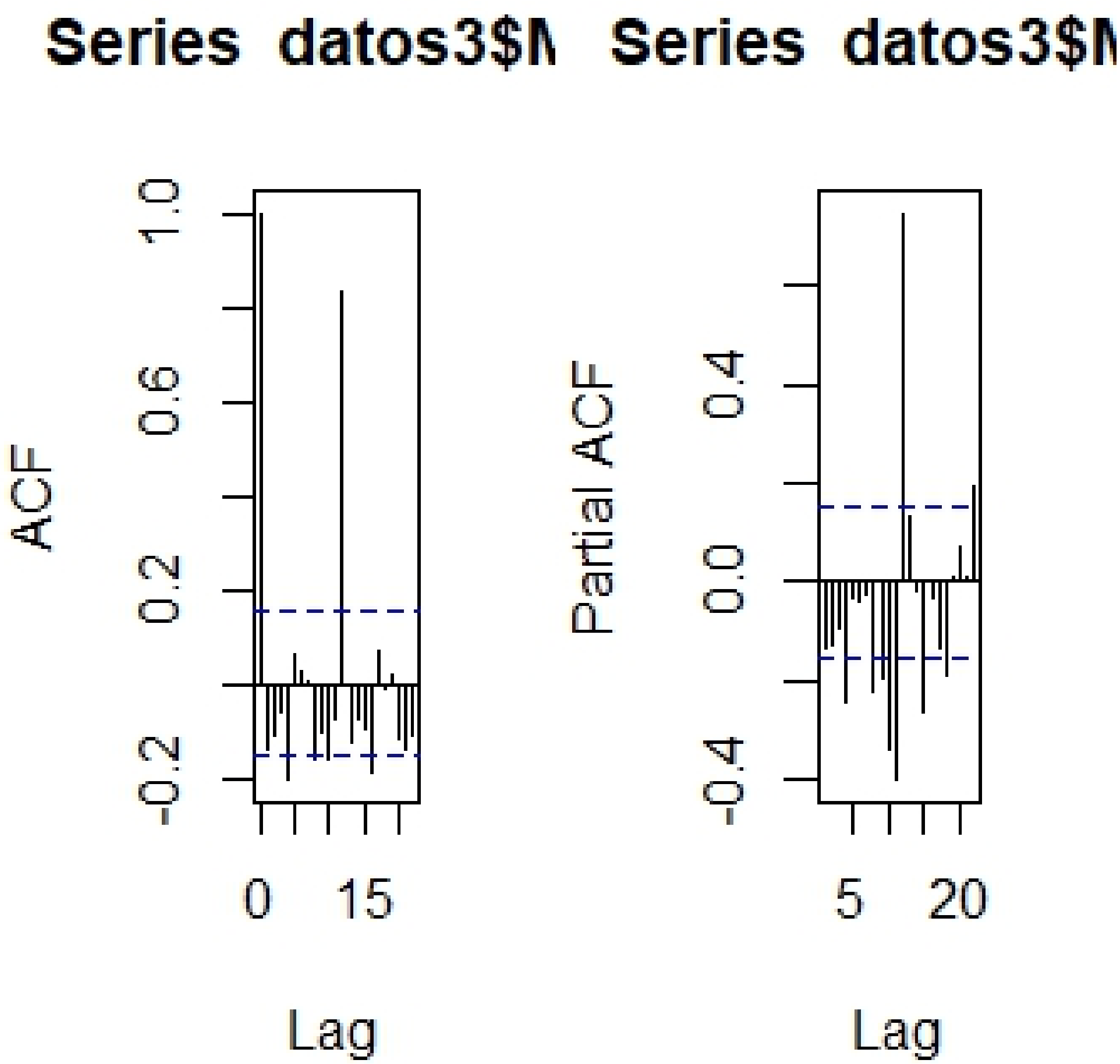

Figura 4.1: ACF y PACF de las serie desestacionalizada de M0

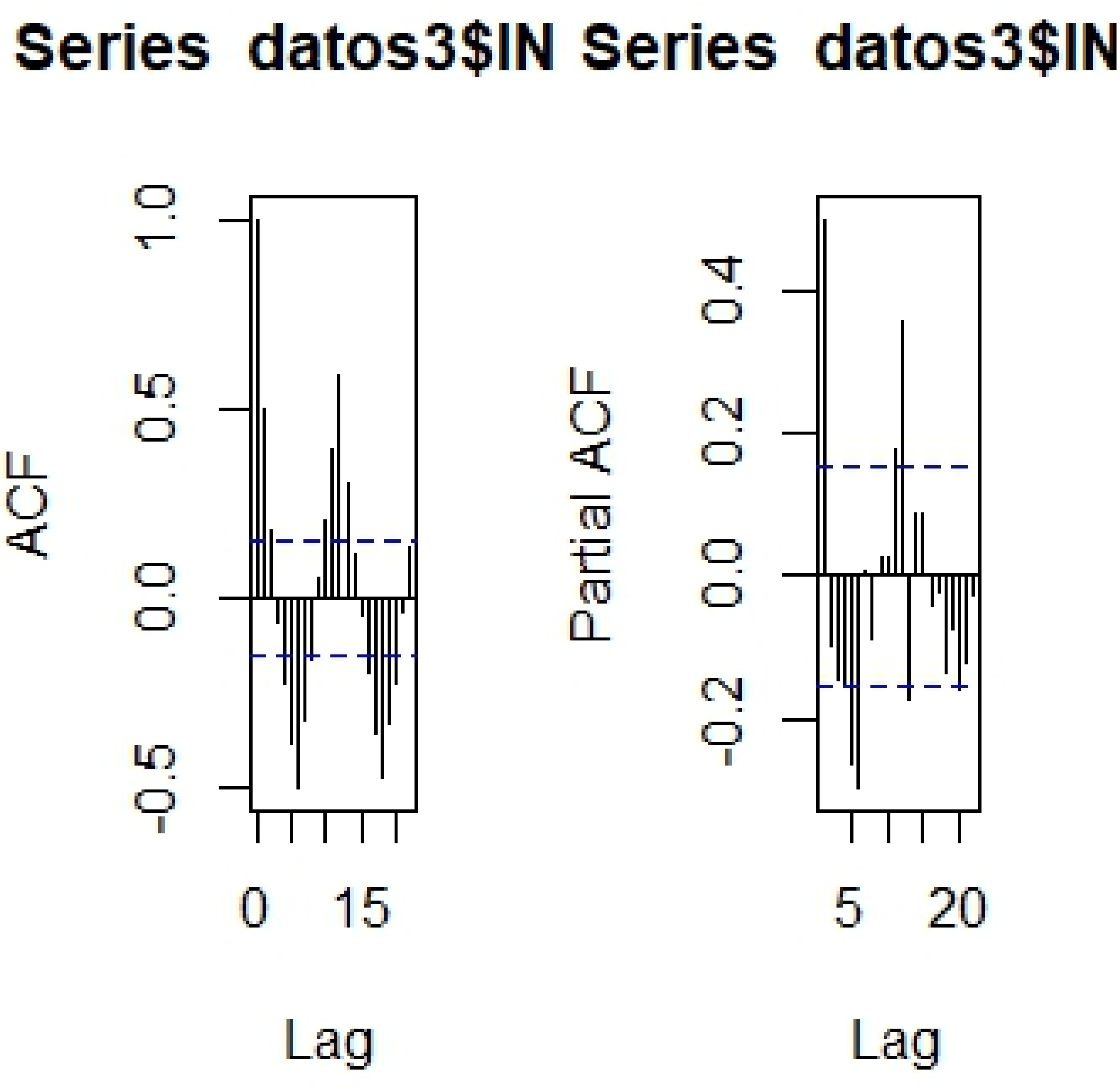

Figura 4.2: ACF y PACF de las serie desestacionalizada de INPC

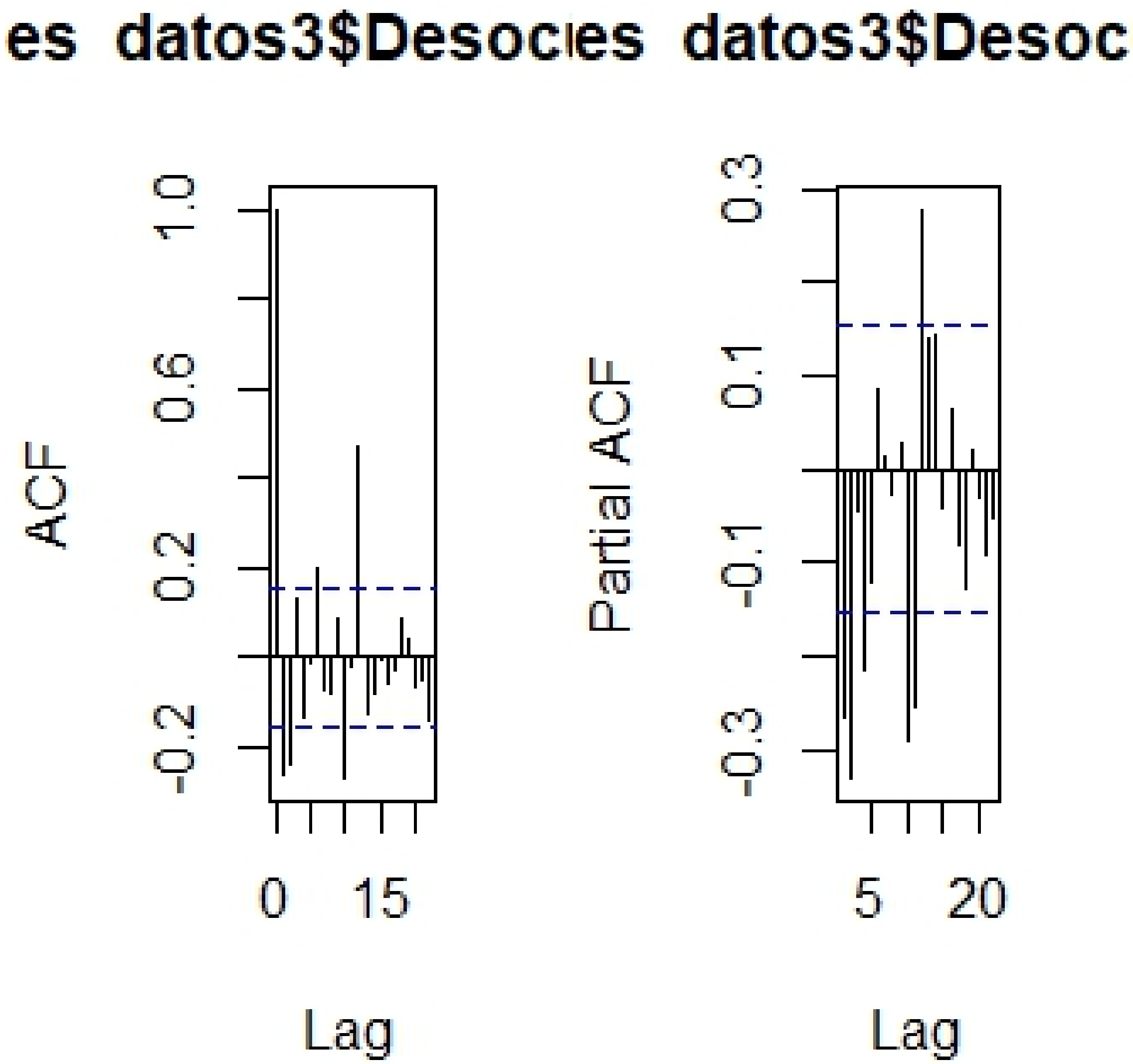

Figura 4.3: ACF y PACF de las serie desestacionalizada de tasa de desocupación

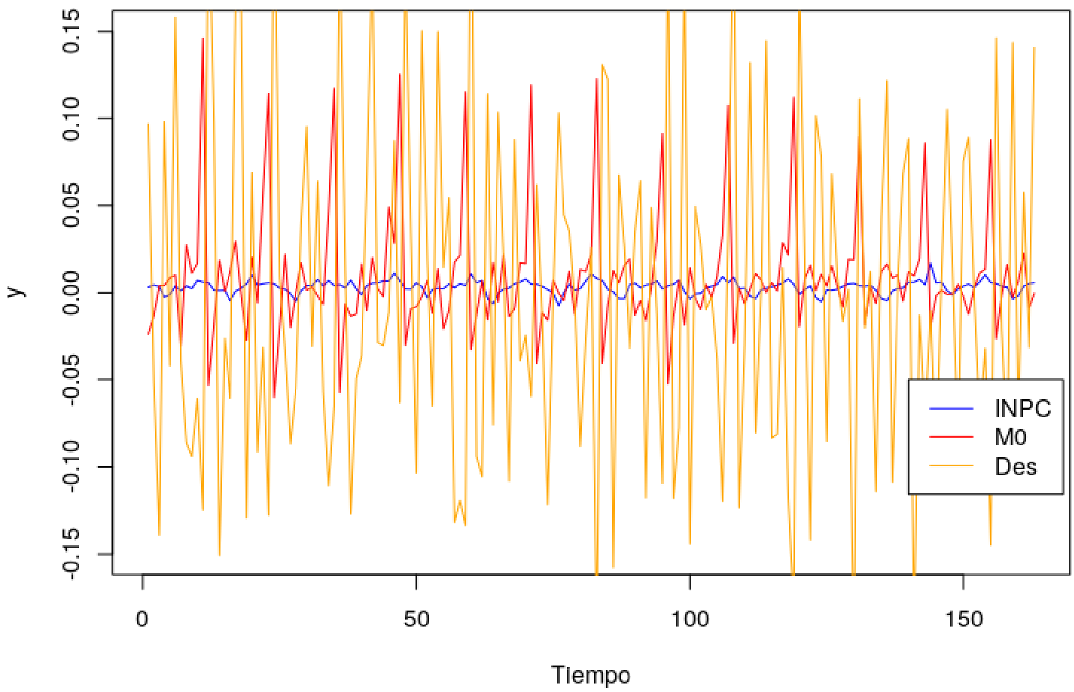

Figura 4.4: Series estacionarizadas en su versión logaritmica

la prueba considerando una confianza de 99 %.

```
Modelo con 12 rezagos
Portmanteau Test (asymptotic)

data:  Residuals of VAR object var1
Chi-squared = 145.48, df = 117, p-value = 0.03819
```

A partir del resultado anterior se prosigió a realizar una prueba de heterocedasticidad para los residuales consiguiendo un p valor de 0.7956 con lo que podría asegurarse que no existe problema de heterocedasticidad para nuestro modelo. De a acuerdo a la prueba de causalidad de Granger no se presenta causalidad en ninguna de las variables, aunque la tasa de desocupación apenas y pasa la prueba por lo que no se debe descartar una posible causalidad para la desocupación.

```
ARCH (multivariate)

data:  Residuals of VAR object var1
Chi-squared = 164.15, df = 180, p-value = 0.7956

> causality(var1, cause="M0")$Granger$p.value
[1,] 0.3811499
> causality(var1, cause="INPC")$Granger$p.value
[1,] 0.7991836
> causality(var1, cause="Desocupacion")$Granger$p.value
[1,] 0.05588753
```

A partir de aqui se considera que nuestro modelo cumple las pruebas de manera que se prosigio a realizar la predicción de la inflación interanual para el mes de octubre. Como se menciono anteriormente las series se recortaron hasta el mes de agosto por lo que se considero el modelo que predecia de mejor manera la inflación interanual para el mes de septiembre.

Ya con esto se produjo la predicción para la inflación interanual para el mes de octubre recordando que se aplico una diferencia a las series y se les aplico una tranformación logaritmo, por lo que al valor predecido se le sumo el último valor de la serie diferenciada, y a esta suma se le aplico una transformación exponencial para generar nuestro valor pronosticado para la inflación interanual, que para el mes de octubre es de **5.037789** que parece ser un resultado bastante confiable considerando que solo se usaron solo otras 2 variables explicativas aparte del propio INPC.

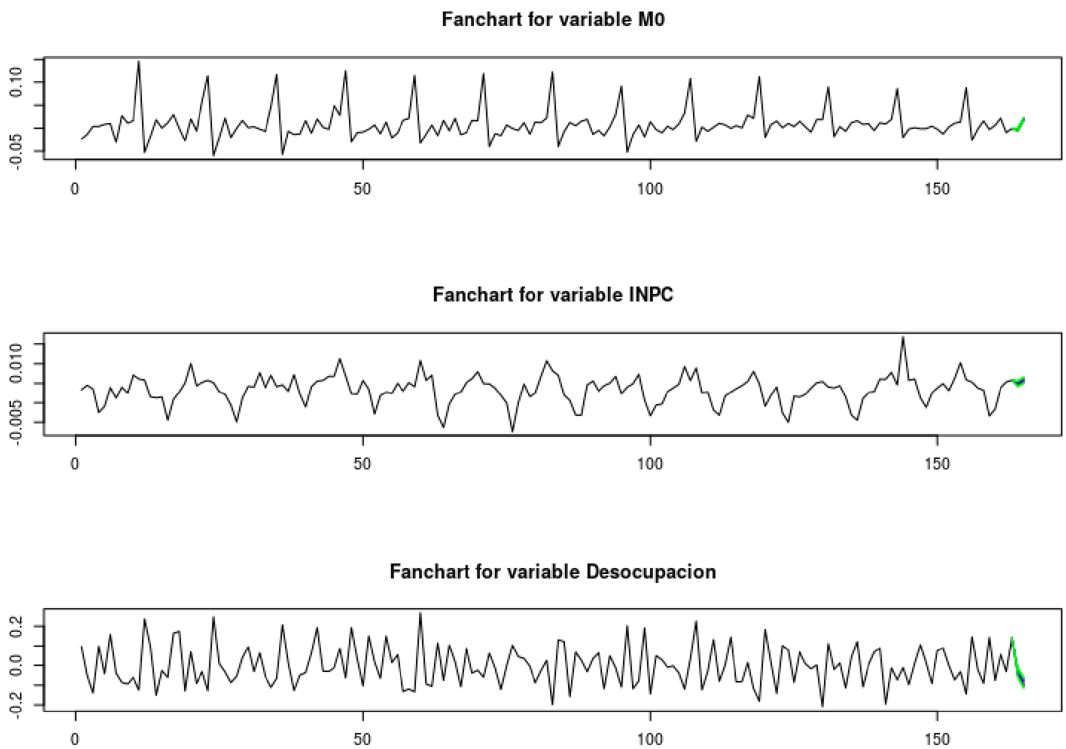

Figura 4.5: Predicción usando nuestro modelo VAR para las 3 variables