Estadística III 3009137, semestre 01 de 2022
Equipo de Trabajo No. 2 Serie No. 3 Curso: Ma – Ju, Horario: 10:00 – 12.00

ANÁLISIS DE SERIES DE TIEMPO. MODELOS DE REGRESIÓN DE ERRORES ESTRUCTURALES ARMA(p, q) ESTACIONARIOS

Sofía Cuartas García[1], Simón Cuartas Rendón[2], Deivid Zhang Figueroa[3]
Fecha de entrega: 03-06-2022

1. INTRODUCCIÓN

Bla, bla, bla.

2. ANÁLISIS DESCRIPTIVO

A continuación se realizará un estudio de la serie de tiempo de ventas nominales que es publicada con periodicidad mensual por el DANE iniciando en enero de 2001 y concluyendo, de momento, en noviembre de 2021, para lo cual se tendrá en cuenta no solo las componentes de la tendencia y la estacionalidad, sino también el error estructural, de tal forma que se logre mejorar el ajuste que se realizaría a partir un modelo global únicamente que ignore tal componente del error estructural, y que por tanto no tenga en consideración a la componente cíclica que, como se verá enseguida, es relevante para esta serie temporal. Dicho esto, es importante tener presente que esta serie cuenta con $N = 251$ registros hasta el momento, y entonces, para poder entender mejor esta serie de tiempo, se debe comenzar observando la ***figura 1***.

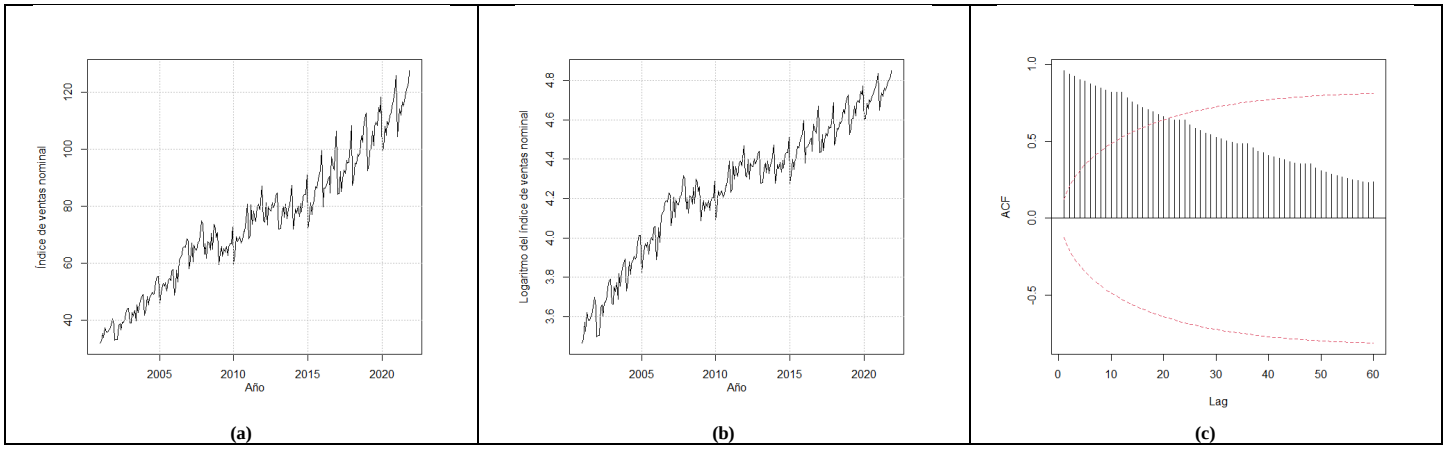

Figura 1. Gráficos descriptivos asociados a la serie temporal del índice de ventas nominales mensuales entre enero de 2001 y noviembre de 2021 calculado y publicado por el *Departamento Administrativo Nacional de Estadística, DANE* de Colombia con modificaciones para los meses abarcados por la pandemia de la *COVID-19*. **(a)** Índice de ventas nominales contra el tiempo. Nótese que la escala del índice nominal de ventas está en escala original, mientras que el tiempo es presentado en años calendario. En esta gráfica se nota que hay un aumento en la variabilidad de este índice conforme avanza el tiempo. **(b)** Índice de ventas nominales contra el tiempo. En este caso el índice de ventas nominales se encuentra en escala logarítmica y puede evidenciarse que se ha estabilizado la variabilidad que se da alrededor de la tendencia en la figura *(a)* con la escala original. **(c)** Gráfica de la función de autocorrelación del *logaritmo* del índice de ventas nominales mensuales, la cual refleja la presencia de la componente estacional en la serie

Con ayuda de la figura 1 se puede realizar un análisis descriptivo de esta serie temporal. Así pues, se debe comenzar en la ***figura 1 (a)***, en la cual se refleja que la tendencia de la serie es creciente en tanto el índice de ventas nominales en Colombia tiende a aumentar con el tiempo, y se tiene también que la serie es multiplicativa, ya que la variabilidad alrededor del valor medio de este índice incrementa con el tiempo; además, es posible advertir la presencia de un comportamiento estacional, dado que hay un patrón repetitivo anualmente.

De igual forma, al observar la ***figura 1 (b)*** que realiza una transformación logarítmica a la escala original del índice de ventas nominales, se confirma que la tendencia es creciente y es claro igualmente que esta es determinística, ya que es posible identificar la presencia de efectos permanentes en las componentes estructurales de la serie temporal, y es destacable el hecho que en el rango histórico de la serie la tendencia aparenta ser global, de manera que si se define a $Y_t^* = log(Y_t)$ como el logaritmo del índice de ventas nominales mensuales de Colombia en un tiempo t, entonces pueda ser modelada en la escala logarítmica como $Y_t^* = T_t^* + S_t^* + E_t$, donde los asteriscos indican que se asocian a una transformación logarítmica, donde T_t^* y S_t^* serán descritos a continuación.

[1] Estudiante de Estadística, Universidad Nacional de Colombia, Sede Medellín
[2] Estudiante de Estadística, Universidad Nacional de Colombia, Sede Medellín
[3] Estudiante de Estadística, Universidad Nacional de Colombia, Sede Medellín

Es así que resulta posible hacer una representación de la tendencia mediante una curva suave de la forma $T_t^* = \beta_0 + \sum_{j=1}^{p} \beta_j t^j$, donde t representa el tiempo de cada uno de los periodos de esta serie (por ejemplo, enero de 2001 está asociado a $t = 1$), T_t^* es el valor del componente de la tendencia en la escala logarítmica del índice de ventas nominales para un tiempo t y p es el grado más alto del polinomio que define a la tendencia; además, a partir de la forma que muestra la tendencia, es razonable plantear que $p \geq 3$ y que es impar, ya que un modelo cuadrático o par mostrará que en algún punto la tendencia comenzará a decrecer, lo cual no se corresponde con lo que se advierte en esta figura. Adicional a lo anterior, es importante tener presente que las cualidades determinística y global de la tendencia son posibles gracias que han sido imputados los índices realmente observados entre marzo de 2020 y noviembre de 2021 por los efectos que tuvo la pandemia de COVID-19 en la economía y por tanto en índices económicos como este; asimismo, esta representación podría tener dificultades por la presencia de ciclos en algunos periodos, fundamentalmente entre los años 2007 y 2015, lo que abre la puerta al planteamiento de modelos locales.

Ahora bien, respecto a la estacionalidad, como se puede ver en las ***figuras 1 (a)*** y ***1 (b)***, existe un comportamiento repetitivo en el índice de ventas nominales dentro de un año calendario, teniendo un aumento progresivo a lo largo del año con algunos picos dentro de cada año, y con esto es posible para los modelos globales considerar funciones trigonométricas, y a través de un periodograma es posible mostrar que esta para la escala logarítmica puede ser modelada como $S_t^* = \sum_{j=1}^{5}[\alpha_j \, sin(2\pi F_j t) + \gamma_j \, cos(2\pi F_j t)]$, para $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$, de manera que en la escala original se obtendría que $S_t = e^{\sum_{j=1}^{5}[\alpha_j \, sin(2\pi F_j t) + \gamma_j \, cos(2\pi F_j t)]}$ para los F_j antes mencionados. Además, al analizar la ***figura 1 (c),*** se corrobora la presencia de la componente estacional, puesto que para observaciones rezagadas $k = sw, s = 12, w = 1,2,3,4,5$ periodos en el tiempo hay una interrupción en el decrecimiento de los valores estimados de la ACF para tener un aumento ligero. Además, nótese que en general se tiene un patrón cola de decaimiento lento, asociado con la presencia de la componente estacional en esta serie temporal, lo cual implica que el índice de ventas nominales no se asocia con un proceso ergódico. Por otro lado, es destacable que $\rho(1) = Corr(Y_t^*, Y_{t+1}^*) > 0$, lo cual se debe a la presencia de ciclos en la serie como se describió antes.

A continuación, es posible determinar si esta serie es estacionaria o no, y rápidamente es posible descartar esta posibilidad, puesto que una serie estacionaria demanda que se tenga varianza y media constantes, y si bien la variabilidad se logra estabilizar con la transformación logarítmica de la escala original del índice de ventas nominales, dado a que esta tiene tendencia, se tiene que la media no es constante. Adicional a esto, teniendo la ACF se pueden realizar *tests* para si el logaritmo del índice de ventas nominales es ruido blanco, para lo que se plantean las siguientes hipótesis para cada $k = 1, 2, \dots, 60$:

$$H_0: \rho(k) = Corr(Y_t^*, Y_{t+1}^*)0 \quad vs. \quad H_1: \rho(k) = Corr(Y_t^*, Y_{t+1}^*) \neq 0$$

Y para los cuales el estadístico de prueba es tal que $\hat{\rho}(k) \sim aprox.\ N(0, {}^{1}/_{n}) \forall k$ y para un $\alpha \approx 5\,\%$ se rechaza H_0 si $|\hat{\rho}(k)| \geq {}^{2}/_{\sqrt{n}}$, y observando la ***figura 1 (c),*** la cual demanda que ninguna de las barras verticales supere a las franjas rojas que demarcan los valores de ${}^{2}/_{\sqrt{n}}$, se concluye que la función de autocorrelación es significativa para $k = 1, 2, 3, \dots, 18, 19, 20$, lo que implica que la serie no corresponda a un ruido blanco.

Luego, teniendo presente que es posible hacer un ajuste global a esta serie temporal, se va a proceder con esto tomando un modelo exponencial polinomial estacional de grado seis estacional con funciones trigonométricas en cinco frecuencias $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$, cuya ecuación es la **(1)**:

$$\begin{aligned} Y_t = exp\Big[&\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) \\ &+ \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\Big] \\ &+ E_t, \ \{E_t\}_{t \in \mathbb{Z}^+} \ un\ R.B. \sim N(0, \sigma^2) \end{aligned} \qquad \textbf{(1)}$$

Para el ajuste de este modelo solo van a ser consideradas las primeras $n = 239$ observaciones, dejando las últimas doce como parte del periodo *ex post* para poder hacer validación cruzada del ajuste realizado, y con ayuda de ***R*** se realiza el ajuste de esta serie, cuyos coeficientes estimados, así como sus errores estándar y los valores del estadístico calculado y el valor p asociados a la prueba de significancia de cada uno de estos se presenta en la ***tabla 1.***

Tabla 1. Parámetros estimados para el modelo global: exponencial polinomial estacional de grado seis estacional con funciones trigonométricas en cinco frecuencias $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$

Parámetros	Estimación	Error estándar	T_0	$P(\lvert t_{222}\rvert > \lvert T_0\rvert)$
β_0	3.55	3.637×10^{-2}	97.608	$< 2\times10^{-16}$
β_1	2.182×10^{-3}	3.478×10^{-3}	0.627	0.531009
β_2	-2.801×10^{-4}	1.107×10^{-4}	2.530	0.012113
β_3	-4.481×10^{-6}	1.576×10^{-6}	-2.844	0.004876
β_4	2.919×10^{-8}	1.111×10^{-8}	2.628	0.009183

β_5	-8.674×10^{-11}	3.794×10^{-11}	-2.286	0.023192
β_6	9.781×10^{-14}	5.011×10^{-14}	1.952	0.052203
α_1	-4.125×10^{-2}	3.581×10^{-3}	-11.519	$< 2\times10^{-16}$
γ_1	1.473×10^{-2}	3.527×10^{-3}	4.176	4.27×10^{-5}
α_2	-2.968×10^{-2}	3.581×10^{-3}	-8.289	1.09×10^{-14}
γ_2	1.357×10^{-2}	3.506×10^{-3}	3.871	0.000143
α_3	-1.925×10^{-2}	3.533×10^{-3}	-5.448	1.35×10^{-7}
γ_3	2.035×10^{-2}	3.550×10^{-3}	5.732	3.22×10^{-8}
α_4	-1.534×10^{-2}	3.572×10^{-3}	-4.295	2.61×10^{-5}
γ_4	2.346×10^{-2}	3.508×10^{-3}	6.688	1.81×10^{-10}
α_5	4.155×10^{-3}	3.542×10^{-3}	1.173	0.242005
γ_5	2.378×10^{-2}	3.530×10^{-3}	6.735	1.38×10^{-10}

$\sqrt{MSE}$=2.899 (escala log.); $exp(C_n^*(p))$: *AIC=9.9001224, BIC=11.52640*

Y así se llega a que la ecuación de ajuste para este modelo es la siguiente:

$$\begin{aligned}\widehat{Y}_t = exp\Big[&3.55 + 2.182\times10^{-3}t - 2.801\times10^{-4}t^2 - 4.481\times10^{-6}t^3 + 2.919\times10^{-8}t^4 - 8.674\times10^{-11}t^5 + 9.781\times10^{-14}t^6 \\ &- 4.125\times10^{-2}sin\left(\frac{1}{6}\pi t\right) + 1.473\times10^{-2}cos\left(\frac{1}{6}\pi t\right) - 2.968\times10^{-2}sin\left(\frac{1}{3}\pi t\right) + 1.357\times10^{-2}cos\left(\frac{1}{3}\pi t\right) \\ &- 1.925\times10^{-2}sin\left(\frac{1}{2}\pi t\right) + 2.035\times10^{-2}cos\left(\frac{1}{2}\pi t\right) - 1.534\times10^{-2}sin\left(\frac{2}{3}\pi t\right) + 2.346\times10^{-2}cos\left(\frac{2}{3}\pi t\right) \\ &+ 4.155\times10^{-3}sin\left(\frac{5}{6}\pi t\right) + 2.378\times10^{-2}cos\left(\frac{5}{6}\pi t\right)\Big]\end{aligned}$$

En la ***figura 2*** se presenta un gráfico en el que se contrasta la serie ajustada, en rojo, con la serie original, en negro.

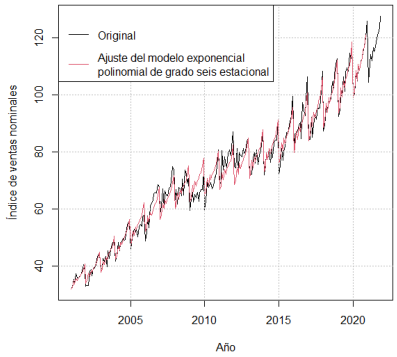

Figura 2. Contraste de la serie de tiempo del índice de ventas nominales mensuales de Colombia entre enero de 2001 y noviembre de 2021 en negro y la serie ajustada con un modelo exponencial polinomial de grado seis estacional con funciones trigonométricas en las frecuencias $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$ en rojo.

De la ***figura 2*** es claro que el modelo logra captar de forma adecuada la tendencia y la estacionalidad de la serie; empero, esta no logra ajustar los ciclos de esta, lo cual se evidencia en una disparidad entre la serie real y la serie ajustada en algunos periodos, siendo esto especialmente evidente en los periodos de los años 2008 y 2009, donde la serie ajustada sigue la tendencia creciente, mientras que los datos reales reflejan una contracción (coincidiendo con la crisis económica mundial de dichos años), si bien es esperable que esto sucediese teniendo presente que la componente cíclica no fue modelada en el modelo global ajustado. Por último, respecto al ajuste es reseñable que la raíz cuadrada del error cuadrático medio es 2.899, el criterio de información de Akaike equivale a 9.9001224 y el criterio de información bayesiano es igual a 11.52640 aproximadamente.

Ahora bien, a la hora de hacer pronósticos luego del periodo $n = 239$ se apela a la siguiente ecuación:

$$\begin{aligned}\widehat{Y}_t = exp\Big[&3.55 + 2.182\times10^{-3}(239+L) - 2.801\times10^{-4}(239+L)^2 - 4.481\times10^{-6}(239+L)^3 + 2.919\times10^{-8}(239+L)^4 - 8.674\times10^{-11}(239+L)^5 \\ &+ 9.781\times10^{-14}(239+L)^6 - 4.125\times10^{-2}sin\left(\frac{1}{6}\pi(239+L)\right) + 1.473\times10^{-2}cos\left(\frac{1}{6}\pi(239+L)\right) \\ &- 2.968\times10^{-2}sin\left(\frac{1}{3}\pi(239+L)\right) + 1.357\times10^{-2}cos\left(\frac{1}{3}\pi(239+L)\right) - 1.925\times10^{-2}sin\left(\frac{1}{2}\pi(239+L)\right) \\ &+ 2.035\times10^{-2}cos\left(\frac{1}{2}\pi(239+L)\right) - 1.534\times10^{-2}sin\left(\frac{2}{3}\pi(239+L)\right) + 2.346\times10^{-2}cos\left(\frac{2}{3}\pi(239+L)\right) \\ &+ 4.155\times10^{-3}sin\left(\frac{5}{6}\pi(239+L)\right) + 2.378\times10^{-2}cos\left(\frac{5}{6}\pi(239+L)\right)\Big]\end{aligned}$$

Y a partir de esta ecuación se puede construir la tabla de pronósticos para los meses del periodo *ex post*, que se observa en la ***tabla 2.***

Tabla 2. Pronósticos para el periodo *ex post* del modelo exponencial polinomial de grado seis.

Periodo	L	Real	Pronóstico
2020 Dic	1	125.95	125.6811
2021 Ene	2	104.49	104.9223
2021 Feb	3	107.87	107.3833
2021 Mar	4	114.16	114.9873
2021 Abr	5	112.05	111.4442
2021 May	6	116.74	117.5733
2021 Jun	7	115.32	155.1623
2021 Jul	8	116.86	118.5237
2021 Ago	9	120.13	120.0877
2021 Sep	10	121.77	122.9627
2021 Oct	11	123.64	124.1830
2021 Nov	12	127.70	128.9216

Nótese que por ser el modelo global exponencial, entonces no es posible obtener para las estimaciones intervalos de predicción. Con esto presente, se tiene por ejemplo que en agosto de 2021, $L = 9$, el índice de ventas nominales pronosticado fue de 120.0877 puntos, toda vez que en este periodo se tuvo que el índice real fue de 120.13 puntos. De forma adicional, como se conocen los valores reales de los índices de ventas nominales para los meses del periodo *ex post*, entonces se pueden calcular medidas de error como el MAE, el MAPE y el RMSE, las cuales son presentadas en la ***tabla 3.***

Tabla 3. Precisión de los pronósticos puntuales.

Medida	Valor
RMSE *	0.8279634
MAE *	0.6862845
MAPE (%)	0.5830712
* Unidades en puntos del índice de ventas nominales.	

De la ***tabla tres*** se concluye que el modelo global se equivocó en promedio en cada pronóstico del periodo *ex post* en 0.8279634 puntos del índice de ventas nominales, mientras que el MAE señala una equivocación en promedio de 0.6862845 puntos; por último, del MAPE se concluye que el modelo global se ha equivocado en promedio para cada pronóstico un 0.4439505 % respecto a cada valor real. De estas métricas pues se puede concluir que se está logrando un ajuste bueno de la serie en tanto los errores cometidos, de acuerdo con estos valores, son pequeños, y como se pudo evidenciar en la figura ***figura 2*** que estos errores se dan fundamentalmente no haber incorporado la componente cíclica en este modelo. Con esto, vale la pena finalizar esta sección con la ***figura 3.***

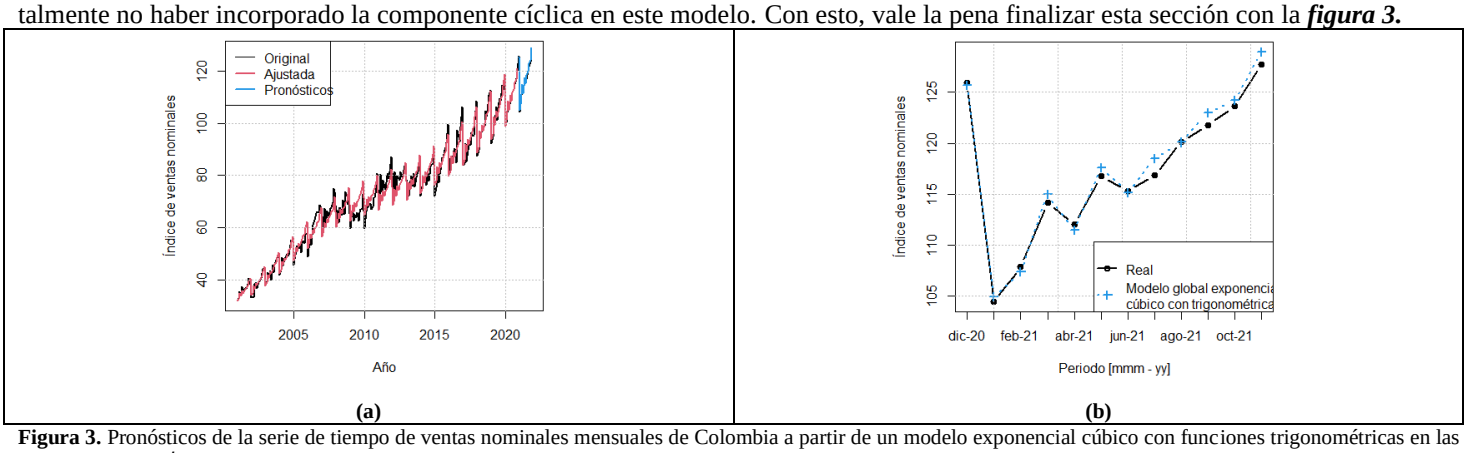

Figura 3. Pronósticos de la serie de tiempo de ventas nominales mensuales de Colombia a partir de un modelo exponencial cúbico con funciones trigonométricas en las frecuencias $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$. **(a).** Modificación de la ***figura 2*** para mostrar también el pronóstico realizado. **(b)** Enfoque en los periodos *ex post* para validar gráficamente la calidad los pronósticos realizados, en azul, en contraste con los valores realmente observados, en negro.

Entonces, se puede validar gráficamente con ayuda de ***figura 3*** y con los valores de la ***tabla 2*** que este modelo global realiza un ajuste un ajuste adecuado para esta serie, aunque se debe tener en consideración que esto está siendo posible gracias a que en los meses del periodo *ex post* se está siguiendo la tendencia y la estacionalidad de la serie sin afectaciones por comportamientos cíclicos. Sin embargo, si algunos de estos periodos fuesen abarcados algún ciclo, se evidenciaría como los pronósticos no serían muy buenos.

3. VALIDACIÓN DE SUPUESTOS SOBRE EL ERROR ESTRUCTURAL EN EL MODELO GLOBAL

En la ***ecuación (1)*** se observa al final que el modelo global planteado incluye un supuesto muy importante sobre los errores estructurales y es que estos son un ruido blanco, lo cual resulta importante ya que con esto se construyen las herramientas que permiten realizar inferencia con el modelo construido. Así pues, se tiene que la suposición de que los errores son ruido blanco implica que

estos se distribuyan idénticamente como una normal con media cero y varianza constante para cualquier tiempo t, y que hay incorrelación entre cualquier par de observaciones sin importar su rezago en el tiempo, lo que en otras palabras significa que los datos del pasado no proporcionan información sobre los errores del futuro.

De este modo, para poder hacer esta evaluación, se van a considerar los residuales del modelo y se va a iniciar chequeando que estos tengan media cero, varianza constante e independencia, lo cual es posible los gráficos de estos residuales que se ilustran en la ***figura 4.***

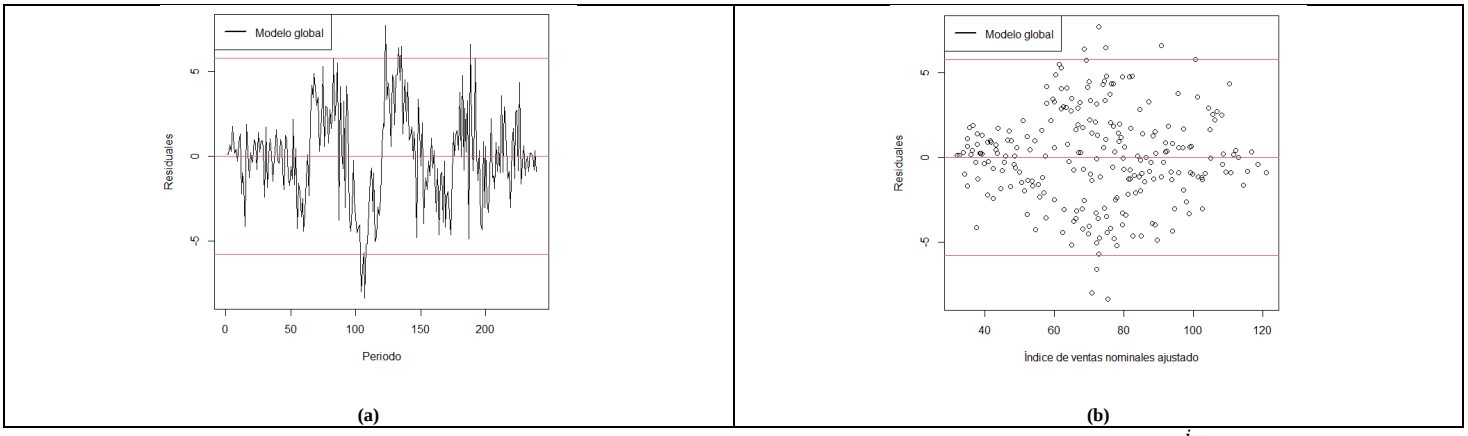

Figura 4. Gráficos para los residuales del modelo exponencial polinomial de grado seis con trigonométricas en en las frecuencias $F_j = {}^{j}/_{12}, j = 1,2,3,4,5$. **(a)** Serie de los residuales del ajuste del modelo. **(b)** Gráfico de dispersión de los residuales de ajuste contra los valores ajustados del modelo.

A partir de las ***figuras 4 (a) y 4 (b)*** se puede determinar que no existe evidencia en contra de que la media de los errores sea diferente de cero, pues la serie de los residuales se da alrededor de cero y los residuales contra el índice de ventas nominales ajustado están dispersos alrededor de cero. Ahora bien, con la varianza se comienzan a tener inconvenientes y esto es especialmente evidente en el gráfico de dispersión de la ***figura 4 (b)*** dado que se observa que los residuales cuentan con mayor varianza hacia índices ajustados alrededor de ochenta puntos, y menor varianza hacia los menores y los mayores índices de ventas nominales ajustados, por lo que no resulta razonable plantear que los errores poseen varianza constante, y por tanto los errores no se distribuyen de manera idéntica gracias a su varianza. Además, en la ***figura 4 (a)*** se evidencia varios ciclos en los residuales, fundamentalmente a partir del periodo cincuenta y hasta el periodo 175 aproximadamente, lo cual implica que existe una correlación positiva entre errores rezagados un periodo en el tiempo; esto es, $\rho(1) = Corr(E_t, E_{t+1}) > 0$, por lo que no se cumple el supuesto de independencia de los errores, y por tanto no se puede evaluar su normalidad. De esta manera, se concluye que los errores del modelo no son un proceso de ruido blanco, y a continuación se procede a verificar esta conclusión con ayuda de las pruebas de incorrelación de *Ljung-Box*, *Durbin-Watson* y los gráficos de las funciones de autocorrelación y autocorrelación parcial con bandas de Bartlett. Entonces, comenzando con los análisis gráficos se tiene a la ***figura 5.***

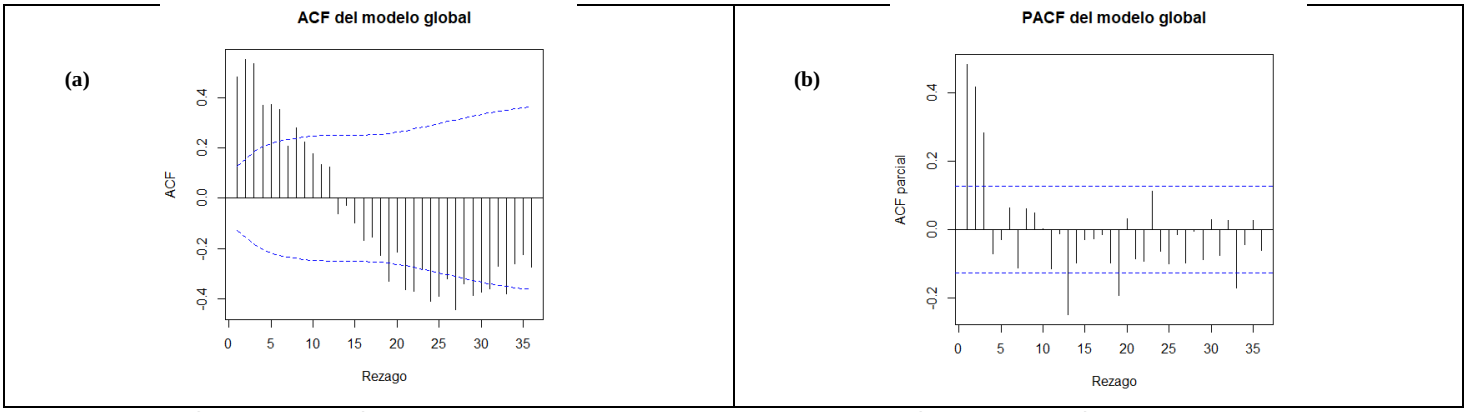

Figura 5. (a). Función de autocorrelación (ACF) muestral con los residuos del modelo global. **(b)** Función de autocorrelación parcial (PACF) muestral con los residuos del modelo global.

Estas gráficas van a ser útiles para evaluar si el error estructural es un ruido blanco, y en particular, con la **ACF** es posible contrastar siguientes hipótesis:

$$H_0: \rho(k) = Corr(E_t, E_{t+k}) = 0 \quad \forall k = 1, 2, \dots, 36$$
$$vs.$$
$$H_1: \exists k: \ \rho(k) = Corr(E_t, E_{t+k}) \neq 0, k = 1, 2, \dots, 36$$

La cual tiene como estadístico de prueba a $\widehat{\rho(k)} = \widehat{Corr}(E_t, E_{t+1}) = \frac{\sum_{t=1}^{239-k} \widehat{E_t E_{t+k}}}{\sum_{t=1}^{239} \widehat{E_t^2}} \sim aprox.N\left(0, \frac{1}{239}\right)$ y que con una significancia de aproximadamente $\alpha \approx 0.05$ rechaza la hipótesis nula si $\left|\widehat{\rho(k)}\right| \geq 2/\sqrt{239}$. Y a partir de la ***figura 5 (a)*** anterior se evidencia que este test se rechaza para $k = 1, 2, 3, 4, 5, 6, 8, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33$, por lo que se rechaza la hipótesis nula y se concluye que los errores estructurales no son ruido blanco. De igual forma, para la **PACF** se tienen las siguientes hipótesis:

$$H_0: \phi_{kk} = Corr(E_t, E_{t+k}|E_{t+1}, \dots, E_{t+k-1}) = 0 \quad \forall k = 1,2, \dots ,36$$
$$vs.$$
$$H_1: \exists k: \ \phi_{kk} = Corr(E_t, E_{t+k}|E_{t+1}, \dots, E_{t+k-1}) \neq 0, k = 1, 2, \dots, 36$$

Y en este caso el estadístico de prueba es $\widehat{\phi_{kk}} = \widehat{Corr}(E_t, E_{t+k}|E_{t+1}, \dots, E_{t+k-1}) \sim aprox.N\left(0, \frac{1}{239}\right)$ y que con una significancia de aproximadamente $\alpha \approx 0.05$ rechaza la hipótesis nula si $\left|\widehat{\rho(k)}\right| \geq 2/\sqrt{239}$. Luego, con ayuda de la ***figura 5 (b)*** se rechaza la hipótesis nula ya que se cumple que la función de autocorrelación parcial es mayor a cero de forma significativa para $k = 1, 2, 3, 13, 19, 33$, lo cual corrobora que los errores estructurales del modelo de regresión global no son un ruido blanco.

XXXX. ¿Para los valores altos de k también se debe decir que hay rechazo?

Ahora se va a analizar qué resulta de los *tests* matemáticos que, a diferencia de los *tests* previos que realizan varias pruebas individuales, en este caso lo hacen de forma conjunta. De esta manera, comenzando con el test *Ljung-Box* se tiene que las hipótesis son:

$$H_0: \rho(1) = \rho(2) = \cdots \rho(m) = 0$$
$$vs.$$
$$H_1: \exists k: \rho(k) \neq 0, k = 1,2, \dots, m$$

Y se tiene que el estadístico de prueba es $Q_{LB} = 239 \times 241 \sum_{k=1}^{m} \frac{[\hat{\rho}(k)]^2}{239-k} \sim \chi_m^2$ y que tiene como criterio de rechazo que el valor p $V_p = P(\chi_m^2 \geq Q_{LB})$ sea pequeño. Con esto claro, se debe tener presente que se va a realizar seis veces este test conjunto para $m = 6, 12, 18, 24, 30, 36$, y con ayuda de ***R*** se obtienen los resultados para este test que se presentan en la ***tabla 4.***

Tabla 4. Test de *Ljung-Box* para los errores estructurales E_t del modelo global.

m	Q_{LB}	Grados de libertad	$P(\chi_m^2 \geq Q_{LB})$
6	300.3857	6.00	0
12	359.7171	12.00	0
18	390.5907	18.00	0
24	570.0018	24.00	0
30	803.6548	30.00	0
36	956.0069	36.00	0

Y como se puede observar, se obtiene que para los seis *tests* de *Ljung-Box* realizados se obtiene un valor p pequeño, lo cual implica que en todos ellos debe ser rechazada la hipótesis nula, lo que implica que existe evidencia muestral suficiente para sugerir que los errores estructurales no están incorrelacionados, y por tanto se llegando una vez más a la conclusión de que estos errores estructurales no son un ruido blanco.

Ahora bien, en cuanto al test de *Durbin-Watson*, que de forma similar del test de *Ljung-Box* lleva a cabo un solo test conjunto, este solo puede ser aplicado en modelos que son lineales en sus parámetros, y teniendo en cuenta que el modelo global considerado es exponencial, no es posible usar este test. Pero, de todos modos, se debe notar que las pruebas anteriores son consistentes al señalar que el error estructural no es un ruido blanco.

Por último, se concluye que estos errores estacionarios no son un proceso estacionario, ya que como se vio con los residuales en los gráficos de la ***figura 4,*** si bien no existe evidencia en contra de que la media de los errores es nula, sí existe evidencia en contra de que los errores tienen varianza constante. Además, a partir de la gráfica de la ACF en la figura ***figura 5 (a)*** se podría pensar que estos errores estructurales no son ergódicos ya que parece no haber una convergencia rápida a cero.

Luego, se debe notar que los errores de este modelo no cumplen con el supuesto de ruido blanco, por lo que se procede con el planteamiento de modelos ***ARMA*** con el objeto de satisfacer estos supuestos y así poder realizar inferencia con este modelo de regresión.

Así, en primer lugar, se debe llamar a las gráficas ACF y PACF de la ***figura 5,*** en los que se debe notar que la ACF parece tener un patrón cola exponencial sinusoidal, mientras que la PACF muestra un patrón tipo corte con $p = 19$ o $p = 33$, por lo que se podría plantear un modelo ***AR(19)*** o un modelo ***AR(33)***, donde resulta razonable darle prioridad al primero por ser más parsimonioso y porque la ACF pierde potencia a medida que aumentar el valor de *k*, de tal suerte que aumenta la probabilidad de cometer un error tipo I al evaluar la significancia estadística de $\rho(33)$ para los errores estructurales. De acuerdo con esto, se tiene que el modelo de regresión estaría dado por::

$$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t, \text{ donde}$$

$$E_t = \sum_{j=1}^{p} \phi_j E_{t-j} + a_t,$$

$$\text{con } \{a_t\}_{t\in\mathbb{Z}^+} \; un\; R.B. \sim N(0, \sigma^2).$$

Donde $p = 19$ si se trata de un ***AR(19)***, o bien, $p = 33$ si se trata de un ***AR(33).***

Luego, con la EACF, cuyo código y salida de ***R*** se puede observar en la ***figura 6,*** se obtiene que el modelo más adecuado es un ***ARMA(7, 11).***

```
> eacf <- eacf(residuals(modelo_global),ar.max = 36, ma.max = 36)
AR/MA
   0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
0  x x x x x x x x x x x  o  o  o  o  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x  x
1  x o x x o x x x o o o  o  x  o  o  o  o  o  x  x  x  o  x  x  o  o  x  o  o  o  o  o  x  o  o  x  o
2  x x x x o o x x o o o  o  o  o  o  o  o  o  x  x  x  x  o  x  x  o  o  x  o  o  o  o  o  x  o  o  x  x
3  x o x o o o o x o o o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o
4  x x o o o o o o o o o  x  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o
5  x o o o o o o o o o o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o
6  x o x x o o o o o o o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o
7  x x x o o o o o o o o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
8  x x x o o o o o o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
9  o x x o o x o o o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
10 o x x x x x o o o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
11 o x o x x x x x o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
12 o x o x x x x o o x x  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
13 x o o o x x x x x o o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
14 x x o o o x x x x x o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
15 x x o o o x o o o x x  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
16 x o o o o x o o o x o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
17 o x o o o x o x o x o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
18 x x x o x x x x x x o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
19 x x x x o x x x o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
20 x x x x o o x x o o o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
21 x x x o x o x x o o o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
22 x x x o x x o x o o o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
23 x x x o x o o x o o o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
24 x x x x x x x x x o o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
25 o x o o o o o o x o o  o  o  x  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
26 o x x o o o o o o x o  x  o  x  x  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o
27 o x x x o o o o o x o  x  o  x  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
28 o x x x o o o o o x o  x  o  x  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
29 x x o o o o o o o o x  o  o  o  o  o  x  o  x  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
30 x x o o o x o o o o o  o  o  o  o  o  x  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
31 x x o o o x o o o x o  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
32 x x o x x x x x x x o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
33 x x x x x x o x o o x  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
34 x o o o x x o x o o o  x  o  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o
35 x x o o x o o x o o x  x  o  o  o  o  o  o  o  o  o  o  x  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o
36 x x o o x x o o o o x  x  o  o  o  o  o  o  o  o  o  o  x  o  o  o  o  o  o  o  o  o  o  o  o  o  o  o
```

Figura 6. Código y salida en ***R*** del EACF para el modelo de regresión global planteado. Nótese que a partir de este se sugiere tomar a los errores estructurales como un ***ARMA(7, 11)***.

Y se tendría entonces que el modelo de regresión estaría dado por:

$$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t, \text{ donde}$$

$$E_t = \sum_{j=1}^{7} \phi_j E_{t-j} + a_t + \sum_{i=1}^{11} \theta_i a_{t-i},$$

$$\text{con } \{a_t\}_{t\in\mathbb{Z}^+} \; un\; R.B. \sim N(0, \sigma^2).$$

Luego, con ayuda de la función `SelectModelo()` de la librería `FitAR`, con la cual se pueden encontrar modelos ***AR(p),*** se encuentra que según el criterio AIC, cuyo código y salida en ***R*** se puede ver en la ***figura 7 (a)*** es un ***AR(22)***, ya que si bien obtiene el mayor AIC exacto y aproximado, no tiene una diferencia considerable con los otros valores y resulta siente el más parsimonioso. Por otro lado, usando al criterio de información bayesiano, BIC, para el cual se presenta la ***figura 7 (b)***, se llega a que se puede ajustar un modelo ***AR(3)***.

```
> SelectModel(residuals(modelo_global),lag.max=36,
Criterion="AIC",ARModel="AR")
   p AIC-Exact AIC-Approx
1 23  364.3485  -129.8790
2 24  365.2689  -130.6130
3 22  366.2007  -130.6947
```

(a)

```
> SelectModel(residuals(modelo_global),lag.max=36,
Criterion="BIC",ARModel="AR")
  p BIC-Exact BIC-Approx
1 3  384.7547 -102.99179
2 4  389.0837  -97.70645
3 5  394.3735  -93.19015
```

(b)

Figura 7. Código y salidas ***R*** de la función `SelectModelo()` de la librería `FitAR` para hallar el orden ***p*** adecuado para modelos ***AR(p)*** usando los criterios de información: **(a)** de Akaike (AIC) y **(b)** bayesiano (BIC).

Lo que implica que la ecuación del modelo de regresión que debería considerar según el criterio de información de *Akaike* es:

$$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t, \text{ donde}$$

$$E_t = \sum_{j=1}^{22} \phi_j E_{t-j} + a_t,$$

con $\{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0,\sigma^2)$.

Mientras que, según el criterio de información bayesiano, debería ser:

$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t$, donde

$$E_t = \phi_1 E_{t-1} + \phi_2 E_{t-2} + \phi_3 E_{t-3} + a_t,$$

con $\{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0,\sigma^2)$.

Asimismo, se va a realizar la tarea de identificar modelos ***ARMA*** pero usando ahora la función `auto.arima()` de la librería `forecast`, cuyos códigos y salidas se observan en la ***figura 8.***

```
> auto.arima(serie_et,ic="aic")
Series: serie_et
ARIMA(1,0,2)(0,0,2)[12] with zero mean

Coefficients:
         ar1      ma1     ma2    sma1     sma2
      0.8204  -0.6404  0.2642  0.1241  -0.1534
s.e.  0.0605   0.0801  0.0724  0.0651   0.0714

sigma^2 = 4.519:  log likelihood = -517.61
AIC=1047.23   AICc=1047.59   BIC=1068.09
```

(a)

```
> auto.arima(serie_et,ic="bic")
Series: serie_et
ARIMA(1,0,2) with zero mean

Coefficients:
         ar1      ma1     ma2
      0.8469  -0.6822  0.2879
s.e.  0.0502   0.0725  0.0708

sigma^2 = 4.666:  log likelihood = -522.1
AIC=1052.21   AICc=1052.38   BIC=1066.11
```

(b)

Figura 8. Código y salidas ***R*** de la función `auto.arima()` de la librería `forecast` para hallar los órdenes ***p*** y ***q*** adecuados para modelos ***ARMA(p, q)*** usando los criterios de información: **(a)** de Akaike (AIC) y **(b)** bayesiano (BIC).

De esta forma, se tendría que el modelo de regresión a plantear de acuerdo con el criterio de información de Akaike es:

$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t$, donde

$$E_t = \phi_1 E_{t-1} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \Theta_1 a_{t-12} + \theta_1\Theta_1 a_{t-13} + \theta_1\Theta_2 a_{t-25} + \theta_2 a_{t-2} + \theta_2\Theta_2 a_{t-14} + \theta_2\Theta_2 a_{t-26},$$

con $\{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0,\sigma^2)$.

Mientras que el modelo que se debería considerar de conformidad con el modelo de regresión bayesiano es:

$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t$, donde

$$E_t = \phi_1 E_{t-1} + a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2},$$

con $\{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0,\sigma^2)$.

Y se puede observar pues que usando el criterio de información de Akaike (AIC) en la ***figura 8 (a)*** que un modelo adecuado para este caso es ARIMA(1, 0, 2)(0, 0, 2)[12], lo cual equivale a un proceso estacionario ***ARMA(1, 2)(0, 2)[12]***, el cual no tiene mucho sentido teniendo en cuenta que, como se vio en la ***figura 5 (a)*** la ACF tiene un patrón cola exponencial sinusoidal, lo cual obliga a que $p \geq 2$. Por otro lado, con el criterio de información bayesiano (BIC), cuyo código y salida en ***R*** se refleja en la ***figura 8***, se obtiene que un modelo adecuado es ARIMA(1, 0, 2), el cual equivale a un proceso estacionario ***ARMA(1, 2)***, el cual presenta el mismo inconveniente que el modelo sugerido por `auto.arima()` usando el AIC, pues se tiene que $p = 1 < 2$, lo cual no tiene sentido a la luz de lo identificado para la ACF.

Por último, dentro de la colección de funciones de identificación de modelos **ARMA** dispuestos por R, se va a proceder ahora con la función `armasubsets()` del paquete `TSA`. En la ***figura 9*** se observa el diagrama que resulta de esta función.

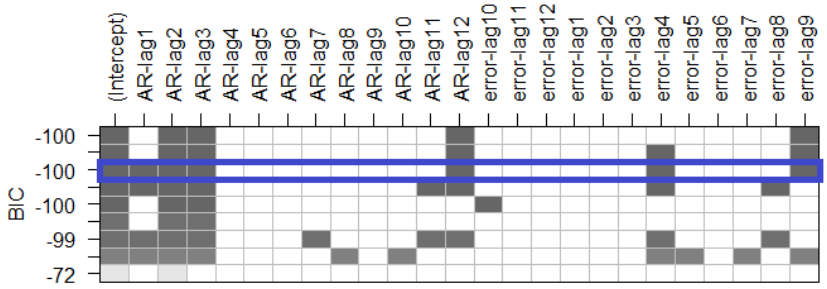

Figura 9. Resultado de la función armasubsets() de la librería TSA sobre los residuales $\widehat{E_t}$ y usando como ***p*** y ***q*** máximos a doce para ambos. La línea de código empleada para obtener este resultado es: `plot(armasubsets(residuals(modelo_global), nar = 12, nma = 12, y.name = 'AR', ar.method = 'ml'))`

A partir del resultado que se visualiza en el primer renglón, se tiene que el modelo a plantear debe ser, según esta función, un ***ARMA(3, 12)*** con parámetros $\theta_i, i = 1,2,3,5,6,7,8,10,11$ fijos en cero, por lo que los errores estructurales estarían siendo modelados como $E_t = \phi_1 E_{t-1} + \phi_2 E_{t-2} + \phi_3 E_{t-3} + \phi_{12} E_{t-12} + a_t + \theta_4 a_{t-4} + \theta_9 a_{t-9}$, con $\{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0, \sigma^2)$. No obstante, si se tiene en cuenta a ϕ_7 y a θ_{10}, se tiene que el modelo de regresión estaría dado por:

$$Y_t = exp\left[\beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \beta_4 t^4 + \beta_5 t^5 + \beta_6 t^6 + \alpha_1 sin\left(\frac{1}{6}\pi t\right) + \gamma_1 cos\left(\frac{1}{6}\pi t\right) + \alpha_2 sin\left(\frac{1}{3}\pi t\right) + \gamma_2 cos\left(\frac{1}{3}\pi t\right) + \alpha_3 sin\left(\frac{1}{2}\pi t\right) + \gamma_3 cos\left(\frac{1}{2}\pi t\right) + \alpha_4 sin\left(\frac{2}{3}\pi t\right) + \gamma_4 cos\left(\frac{2}{3}\pi t\right) + \alpha_5 sin\left(\frac{5}{6}\pi t\right) + \gamma_5 cos\left(\frac{5}{6}\pi t\right)\right] + E_t, \text{ donde}$$

$$E_t = \phi_1 E_{t-1} + \phi_2 E_{t-2} + \phi_3 E_{t-3} + \phi_7 E_{t-7} + \phi_{12} E_{t-12} + a_t + \theta_4 a_{t-4} + \theta_9 a_{t-9} + \theta_{10} a_{t-10},$$

$$\text{con } \{a_t\}_{t\in\mathbb{Z}^+}\ un\ R.B. \sim N(0, \sigma^2).$$