

Interrogación I parte II

Series de tiempo

Profesor: Ricardo Olea
estudiantes: Maximiliano Norbu
Sebastián Cabezas
Fecha: 11/10/2022

Modelos

Mediante técnicas de regresión lineal o no lineal, ajuste el comportamiento de la serie de venta mensual

En este trabajo realizamos 11 modelos de regresión lineal, en los cuales fuimos añadiendo variables y analizando el aporte que tenían en la representación de un modelo correcto. Las variables dependientes para cada modelo fueron las siguientes:

1. IPC + drift: Con el IPC el modelo queda bastante mal ajustado, por lo que probamos con otra variable.
2. Apertura del dolar + drift: Con la apertura del dolar queda un poco mejor que con el IPC aunque tampoco creemos que sea la variable macroeconomica apropiada. Cambiamos la apertura del dolar con la variación.
3. Variación del dolar + drift: No ajusta para nada correcto. Seguimos probando con otras variables.
4. IMACEC + drift: La variable IMACEC al menos aporta a la caída final del modelo. Esta nos aporta información por lo que la dejamos y seguimos añadiendo más variables.
5. IMACEC + pase de movilidad + drift: El pase de movilidad ayuda un poco a modelar la tendencia.
6. IMACEC + pase de movilidad + mes(factor) + drift: Agregando los meses como factores nos acercamos mucho más a un ajuste apropiado. Seguimos añadiendo variables para ver si mejora el modelo.
7. IMACEC + pase de movilidad + mes(factor) + fecha estallido social + drift: mejora un poco al final del 2019.
8. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + drift(bs): Pareciera que el modelo mejora en general comparandolo con el modelo anterior. Para ver con cual seguimos analizamos medidas de error, en las cuales nos dio como resultados que el ultimo modelo tiene un mayor R^2 y un menor AIC. Analizando la medida de error que se comentó en el enunciado, el MAPE, obtenemos que este es menor en el ultimo modelo.
9. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + retiros (factor) + drift(bs): Se nota similar al anterior. Tambien analizamos las medidas de bondad de ajuste que son similares. intentamos otra cosa con los retiros del 10%:
10. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + retiro1 (factor) + retiro2 (factor) + retiro3 (factor) + drift(bs): Aparenta que el modelo ajusta mejor, para confirmar analizamos medidas de bondad de ajuste, ajustando mejor pero no conformes con esto, probamos con el primer mes de cada retiro, el segundo mes de cada retiro y el mes posterior a cada retiro ya que allí probablemente la gente tenia menos dinero:

11. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + mes 1 retiro (factor) + mes 2 retiro (factor) + mes post retiro (factor) + drift(bs): podemos ver que ajustan similar y que los indices de bondad de ajuste son similares. Probamos añadiendo la temperatura.
12. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + mes 1 retiro (factor) + mes 2 retiro (factor) + mes post retiro (factor) + temperatura + drift(bs): Tambien probamos el mismo modelo con los retiros:
13. IMACEC + pase de movilidad + mes(factor) + fecha estallido social (factor) + retiro1 (factor) + retiro2 (factor) + retiro3 (factor) + temperatura + drift(bs): Analizamos los indices de bondad de ajuste de estos ultimos modelos y notamos que el ultimo modelo y el 11 ajustan mejor, el ultimo en base al MAPE y AIC mientras que el 11 en base al R^2 . Realizando un summary notamos que la temperatura, el IMACEC y la parte cubica del spline no son significativos. iremos eliminando estas variables en ambos modelos y analizamos que tal.
14. pase de movilidad + mes(factor) + fecha estallido social (factor) + mes 1 retiro (factor) + mes 2 retiro (factor) + mes post retiro (factor) + drift(bs): Los indices de bondad de ajuste estn muy bien, ahora eliminamos el grado 3 del bs:
15. pase de movilidad + mes(factor) + fecha estallido social (factor) + mes 1 retiro (factor) + mes 2 retiro (factor) + mes post retiro (factor) + drift(bs, degree = 2): Enpeoran un poco los indices de bondad de ajuste, pero seguimos eliminando por parsimonia.

Modelos finales:

1. Pase de movilidad + mes (factor) + estallido (factor) + drift (bs, degree = 2) + Retiro2 (factor)
2. Pase de movilidad + mes (factor) + estallido (factor) + drift (bs, degree = 2) + retiro segundo mes (factor)

Se intentó realizar una transformación de Box-Cox a estos modelos finales, sin embargo, los intervalos contienen al 1 por lo que se decide no realizar la transformación.

Posterior a todo esto, realizamos una predicción simplemente para ver que tal se vería, el resultado se puede apreciar en imagen 1 al final del documento. Podemos notar que ambas predicciones se ven razonables. Analizaremos los residuos de ambos modelos y les realizaremos un teste de blancura. Las imgenes de estos residuos se presentan al final del documento, en la imagen 2. En esta, a simple vista, no se ve que se comporte como ruido blanco, realizamos graficos de ACF y PACF (imagen 3) y TS.diag (box-Ljung) (imagen 4), confirmamos lo comentado. Por los graficos de ACF y PACF creemos apostamos por un modelo MA(1) o MA(2)

Continuamos con un auto.arima() para ver que sugiere R, para los residuos en ambos sugiere un arima(1,0,0), graficamos para ver que tal estos modelos (imagen 5) y notamos que graficamente se ven bien. Analizamos los indices de bondad de ajuste y obtenemos que tienen un AIC muy similar aunque el MAPE del segundo es mucho más bajo. Ploteamos los residuos de los residuos para ver si se comportan como ruido blanco (imagen 6) y segun los graficos, parece que si. Graficamos ACF y PAFC (Imagen 7) y box-Ljung (Imagen 8) en ambos se ven muy bien, con algun punto cuestionable pero por lo general bien. Realizamos un autoarima para los residuos y en este caso vamos a realizar una diferenciación. En estos obtenemos