Modelo autorregresivo de orden (p) aplicado al PIB de México

José Julián Mateo López

2022/09/18

Introducción

Una de las principales funciones de los modelos econométricos es poder explicar el comportamiento futuro de múltiples fenómenos económicos. Los modelos dinámicos son una herramienta muy útil para poder realizar pronósticos de múltiples variables macroeconómicas, principalmente series de tiempo. Podemos dividir estos modelos en modelos dinámicos univariados y multivariados. En este trabajo, se elaborará un modelo univariado, denominado modelos autorregresivo de orden (p), que en términos generales, representa un trabajo de baja escala comparado con los múltiples modelos econométricos avanzados de series temporales. Este tipo de modelos es una introducción a la metodología econométrica de series temporales, y surgen como la base teórica de modelos de medias móviles, modelos autorregresivos y de medias móviles, modelos autorregresivos integrados de medias móviles y vectores autorregresivos, mismos que se desarrollarán mas adelante. Este trabajo tiene la finalidad de servir como base para comprender el tratamiento estadístico matemático empleado en este tipo de modelos (dinámicos). Además, este conjunto de trabajos serán aplicados mediante el lenguaje de programación R, que al pertenecer a la programación requiere un poco mas de esfuerzo y dedicación por parte del lector.

Obtención de la información

```r
library(readxl)
datos <- read_excel("Modelo autorregresivo PIB.xlsx")
attach(datos)
head(datos, n = 8)
```

```
## # A tibble: 8 x 1
##         PIB
##       <dbl>
## 1 12725022.
## 2 12994569.
## 3 13008792.
## 4 13003302.
## 5 12796119.
## 6 12967530.
## 7 12857409.
## 8 12901429.
```

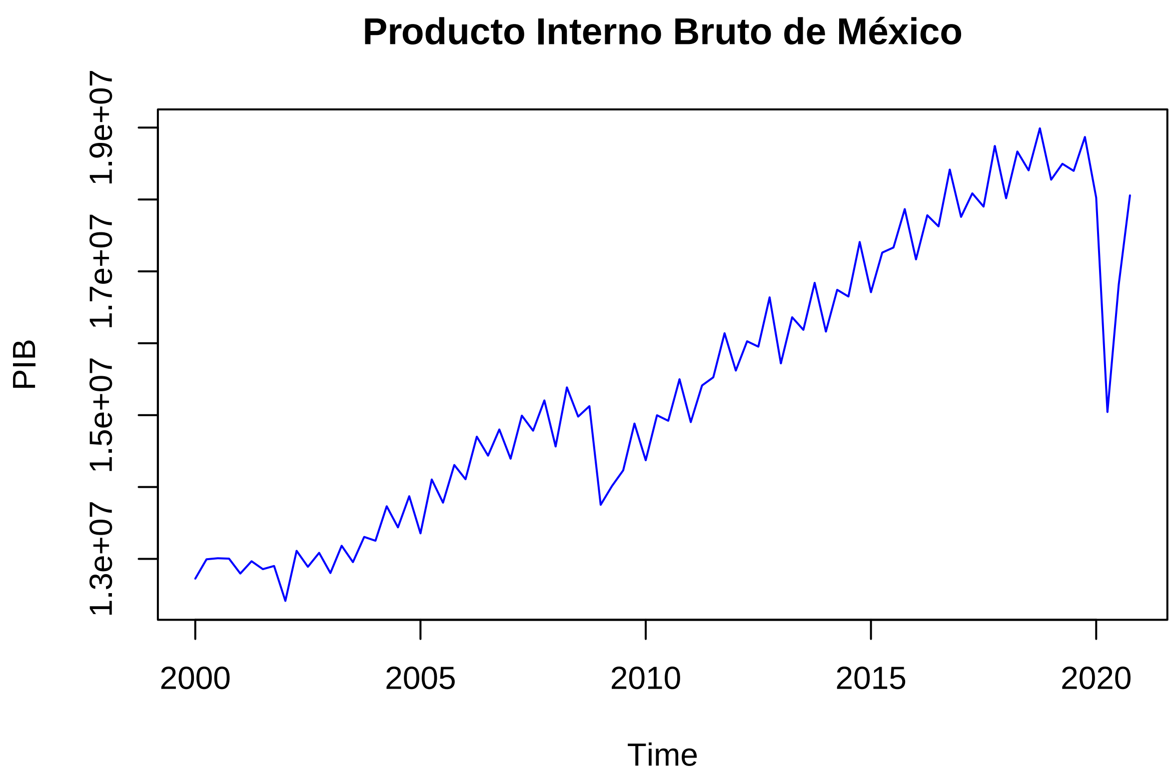
Una vez importada la base de datos, procedemos a activar las librerías que utilizaremos en esta sección. Las cuales son **tseries** y **dynlm**.

```
library(tseries)
library(dynlm)
PIB_ts <- ts(PIB, start = 2000, frequency = 4)
head(PIB_ts, n = 8)
```

```
## [1] 12725022 12994569 13008792 13003302 12796119 12967530 12857409 12901429
```

Por lo general, el análisis de series temporales empieza por la visualización de los datos. Por lo que, procedemos a llevar a cabo el código que nos permita ver el comportamiento de la serie a través de un gráfico de línea.

```
plot(PIB_ts, main = "Producto Interno Bruto de México",
     ylab = "PIB", col = "Blue")
```



Modelo autorregresivo de orden 1

```
modelo1 <- dynlm(PIB_ts ~ L(PIB_ts))
summary(modelo1)
```

```
##
## Time series regression with "ts" data:
## Start = 2000(2), End = 2020(4)
```

```
##
## Call:
## dynlm(formula = PIB_ts ~ L(PIB_ts))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2896859  -437310    12605   481085  1684449
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) 9.441e+05  5.692e+05   1.659    0.101
## L(PIB_ts)   9.429e-01  3.663e-02  25.741   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 639500 on 81 degrees of freedom
## Multiple R-squared:  0.8911, Adjusted R-squared:  0.8897
## F-statistic: 662.6 on 1 and 81 DF,  p-value: < 2.2e-16
```

Ecuación 2 del modelo autorregresivo de orden (P)

En la próxima ecuación se puede observar los resultados de la regresión, donde se ha transformado la variable del PIB en términos logarítmicos y funcionan como variables explicativas los retardos 1,2 y 8, que en términos generales, son estadisticamente significativos. Sin embargo, el intercepto del modelo presenta una probabilidad de 0.36, por consiguiente, no cumple con el numero estadístico de referencia 2.

```r
logPIB_ts <- log(PIB_ts)
modelo2 <- dynlm(logPIB_ts ~ L(logPIB_ts,c(1,2,8)))
summary(modelo2)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(1), End = 2020(4)
##
## Call:
## dynlm(formula = logPIB_ts ~ L(logPIB_ts, c(1, 2, 8)))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.212097 -0.009887  0.003815  0.018924  0.051430
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)                0.51968    0.57439   0.905  0.36862
## L(logPIB_ts, c(1, 2, 8))1  0.42115    0.10082   4.177 8.17e-05 ***
## L(logPIB_ts, c(1, 2, 8))2  0.25037    0.11961   2.093  0.03985 *
## L(logPIB_ts, c(1, 2, 8))8  0.29794    0.09199   3.239  0.00182 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03479 on 72 degrees of freedom
## Multiple R-squared:  0.9157, Adjusted R-squared:  0.9122
## F-statistic: 260.7 on 3 and 72 DF,  p-value: < 2.2e-16
```

Ecuación 3 del modelo autorresivo de orden (P): Agregando tendencia al modelo

Ahora procederemos a agregar la tendencia al modelo para ver los resultados estadísticos principales y poder llevar a cabo la inferencia estadística correspondientes. Para poder agregar la tendencia al modelo, previamente necesitamos activar la librería **astsa**, por lo que, ya estando en condiciones podemos proceder a la estimación del modelo, tal como se muestra en el siguiente comando.

```
library(astsa)
modelo3 <- dynlm(logPIB_ts ~ L(logPIB_ts,c(1,2,8)) + trend(logPIB_ts))
summary(modelo3)
```

```
##
## Time series regression with "ts" data:
## Start = 2002(1), End = 2020(4)
##
## Call:
## dynlm(formula = logPIB_ts ~ L(logPIB_ts, c(1, 2, 8)) + trend(logPIB_ts))
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.21197 -0.01107  0.00393  0.02030  0.04791
##
## Coefficients:
##                           Estimate Std. Error t value Pr(>|t|)
## (Intercept)               1.903948   3.333369   0.571  0.56968
## L(logPIB_ts, c(1, 2, 8))1 0.395350   0.118428   3.338  0.00134 **
## L(logPIB_ts, c(1, 2, 8))2 0.242364   0.121786   1.990  0.05043 .
## L(logPIB_ts, c(1, 2, 8))8 0.246782   0.152578   1.617  0.11022
## trend(logPIB_ts)          0.001816   0.004306   0.422  0.67455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03499 on 71 degrees of freedom
## Multiple R-squared:  0.9159, Adjusted R-squared:  0.9112
## F-statistic: 193.4 on 4 and 71 DF,  p-value: < 2.2e-16
```

De acuerdo a los resultados anteriores, tanto el intercepto de la regresión como el rezago 8 y el termino de tendencia no son estadísticamente significativos. Una de las recomendaciones en trabajos aplicados, es llevar a cabo diversas aplicaciones de forma que podamos encontrar la recta de mejor ajuste, que se ajuste a los datos. Continuando con la experimentación, procedemos a omitir el termino autorregresivo 8 de la ecuación.

```
modelo4 <- dynlm(logPIB_ts ~ L(logPIB_ts, c(1,2)) + trend(logPIB_ts))
summary(modelo4)
```

```
##
## Time series regression with "ts" data:
## Start = 2000(3), End = 2020(4)
##
## Call:
## dynlm(formula = logPIB_ts ~ L(logPIB_ts, c(1, 2)) + trend(logPIB_ts))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
```

```
## -0.208298 -0.012012  0.003019  0.020232  0.058598
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            6.188359   1.911504   3.237  0.00177 **
## L(logPIB_ts, c(1, 2))1 0.336614   0.108678   3.097  0.00271 **
## L(logPIB_ts, c(1, 2))2 0.284854   0.114247   2.493  0.01477 *
## trend(logPIB_ts)       0.007333   0.002410   3.042  0.00320 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03408 on 78 degrees of freedom
## Multiple R-squared:  0.9265, Adjusted R-squared:  0.9237
## F-statistic: 327.9 on 3 and 78 DF,  p-value: < 2.2e-16
```

Ahora si se puede apreciar la significancia estadística de todos los estimadores del modelo, tanto el intercepto como los dos primeros términos autorregresivos son estadísticamente significativos al segundo numero de referencia (5%). Por su parte, el termino de tendencia también cumple con lo indicado por la teoría estadística. Por lo que, el método de predicción seleccionado, al menos por el momento, toma la forma de la función 4.

Agregando estacionalidad al modelo

Como paso final en al elaboración de este método de predicción, procederemos a ver el comportamiento estadístico del modelo agregando estacionalidad al mismo, tal como se muestra en el código siguiente.

```
modelo5 <- dynlm(logPIB_ts ~ L(logPIB_ts, c(1,2)) + trend(logPIB_ts) + season(logPIB_ts))
summary(modelo5)
```

```
##
## Time series regression with "ts" data:
## Start = 2000(3), End = 2020(4)
##
## Call:
## dynlm(formula = logPIB_ts ~ L(logPIB_ts, c(1, 2)) + trend(logPIB_ts) +
##     season(logPIB_ts))
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.206989 -0.005038  0.003026  0.012459  0.034806
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)            5.095206   1.650520   3.087  0.00283 **
## L(logPIB_ts, c(1, 2))1 0.610232   0.115145   5.300 1.12e-06 ***
## L(logPIB_ts, c(1, 2))2 0.076177   0.120885   0.630  0.53051
## trend(logPIB_ts)       0.006058   0.002077   2.916  0.00467 **
## season(logPIB_ts)Q2    0.042563   0.011574   3.677  0.00044 ***
## season(logPIB_ts)Q3    0.026425   0.009187   2.876  0.00524 **
## season(logPIB_ts)Q4    0.057256   0.009725   5.888 1.03e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## Residual standard error: 0.02873 on 75 degrees of freedom
## Multiple R-squared:  0.9498, Adjusted R-squared:  0.9458
## F-statistic: 236.5 on 6 and 75 DF,  p-value: < 2.2e-16
```

Los resultados correspondientes a la ecuación 5 indican que los términos de estacionalidad son estadísticamente significativos. No obstante, debido al ingreso de la estacionalidad al modelo 5, el retardo 2 ha recogido un P-value de 0.53. Bien podemos dejar la ecuación tal cual, o si queremos, podemos eliminar el rezago 2 para obtener las predicciones del mismo.

PREDICCIONES DEL MODELO SELECCIONADO (M5)

```
library(urca)
```