ECONOMETRÍA
APLICADA UTILIZANDO R.

PAPIME PE302513 LIBRO ELECTRÓNICO Y COMPLEMENTOS DIDÁCTICOS EN MEDIOS COMPUTACIONALES, PARA EL FORTALECIMIENTO DE LA ENSEÑANZA DE LA ECONOMETRÍA

Capítulo 2

Enfoque matricial de la regresión lineal

Javier Galán Figueroa

Objetivo

Que el usuario conozca las rutinas básicas que son necesarias para estimar los parámetros de la regresión lineal a través del enfoque matricial utilizando la paquetería del software R, los cuales podrán ser utilizados en sus variantes como es el RStudio.

Introducción

Para estimar los parámetros de la regresión lineal a través del enfoque matricial, se utilizarán datos de la economía mexicana para el periodo enero de 2009 a diciembre de 2013, con frecuencia mensual y cuya fuente provienen de la página web del Banco de México .

Se establecerá el modelo en el cual se explique el comportamiento de la deuda externa en función de la reserva internacional y del índice bursátil.

1. El modelo matricial

A continuación se estimará la siguiente relación:

$$y = f(X2, X3)(1)$$
$$y = X\beta + u(2)$$
$$y_t = \beta_1 + \beta_2 X2_t + \beta_t X3_t + u_t(3)$$

Donde la variable dependiente, y, esta en función de las variables explicativas, $X2, X3$.

$$\underset{(nx1)}{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_3 \end{bmatrix}} = \underset{(nxk)}{\begin{bmatrix} 1 & X_{21} X_{31} & \ldots\ldots & X_{k1} \\ 1 & X_{22} X_{32} & \ldots\ldots & X_{k2} \\ \vdots & \vdots \quad \vdots & \ldots\ldots & \vdots \\ 1 & X_{2n} X_{3n} & \ldots\ldots & X_{kn} \end{bmatrix}} \underset{(kx1)}{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \end{bmatrix}} + \underset{(nx1)}{\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}} (4)$$

La ecuación (2) es la representación matricial de la regresión lineal, donde y es un vector columna de orden $(nx1)$, X es una matriz de orden (nxk), β es un vector columna de orden $(kx1)$, por último u es un vector columna de orden $(nx1)$.

A través de los Mínimos Cuadrados Ordinarios (MCO) se procederá a estimar los parámetros de la ecuación (2). Para ello se considera que el vector β de la ecuación (2) es estimable a partir de la siguiente expresión:

$$\beta = (X'X)^{-1}X'y \quad (5)$$

Utilizar el enfoque matricial permite estimar con relativa sencillez las funciones de regresión simple (una sola variable explicativa) o múltiple (con k variables explicativas).

De la ecuación (3) la variable dependiente, y**,** es el nivel de deuda pública del gobierno mexicano (miles de millones de pesos) que es explicada por el nivel de reservas internacionales, $X2$, (miles de millones de dólares) y por el índice bursátil de la Bolsa Mexicana de Valores, $X3$ (miles de unidades).

Para encontrar el modelo en el cual explique el comportamiento de la deuda externa en función de la reserva internacional y del índice bursátil se utilizará los datos que se encuentran en el archivo CAP2_MCO con extensión CSV.

Para ejecutarlo en R es necesario direccionar la carpeta en la que se encuentra el archivo

```
deuda<-read.csv("cap2_mco.csv",header=T)
attach(deuda)
```

Si se desea visualizar los datos a través de una lista, basta con escribir:

```
deuda
```

2. Análisis exploratorio de datos

Si se desea obtener de manera individual: media aritmética, mediana, desviación estándar y varianza de la variable (y), se utilizan los siguientes comandos:

```
mean(y)
median(y)
sd(y)
var(y)
```

Para hacer lo manera conjunta usamos el comando:

```
summary(y)
```

Posteriormente se puede obtener el histograma y la gráfica de caja en un sólo gráfico:

```
split.screen(c(1,2))
hist(y)
screen(2)
boxplot(y)
```

Repetimos el mismo código para las variables X2 y X3 y observamos los resultados con el comando:

```
summary(X2,X3)
```

A continuación obtenemos la matriz de correlación entre las variables (y, X2, X3).

```
cor(deuda)
```

Para obtener los diagramas de dispersión para indicar a nivel gráfico como influye la reserva internacional (X2) y el índice bursátil (X3) al nivel de endeudamiento del gobierno mexicano (y) utilizamos los siguientes comandos:

```
scatter1<-plot(y~X2)
fit<-lm(y~X2)
abline(fit)
scatter1<-plot(y~X3)
fit2<-lm(y~X3)
abline(fit2)
```

3. Estimación por Mínimos Cuadrado Ordinarios

Con el análisis previo se procederá a estimar los parámetros de la ecuación (3) a través de los Mínimos Cuadrados Ordinarios (mco). Para ello se considera que el vector de la ecuación (2) es estimable a partir de la siguiente expresión:

$$\beta = (X'X)^{-1}X'y$$

Como primer paso se debe especificar en el programa R la matriz X así como el vector y para ello se sigue el siguiente algoritmo:

1) Para transformar un conjunto de variables a matriz se utiliza el código “cbind()”
2) Una vez que se ha dado de alta las matrices en r se procede a realizar las operaciones correspondientes para encontrar los componentes del vector β.

Para crear la matriz X y que esta se encuentre conformada de acuerdo a la ecuación (4) se utiliza el siguiente código:

```
X<-cbind(1,X2,X3)
```

Para el caso para transformar la variable deuda pública (y) a vector se utiliza el mismo código:

```
y1<-cbind(y)
```

Para estimar el vector (X'X) de la ecuación (5), primero se obtiene el parámetro para ello se sigue los siguientes pasos:1) transpuesta de X; 2) Producto de la transpuesta de X por X, cabe mencionar, en el programa R el producto de matrices se lleva a cabo mediante el código "%*%".

```
trX<-(t(X))
X_X<-trX %*% X
X_X
```

A continuación se obtiene el determinante de la matriz (X'X), para determinar si ésta tiene inversa o no. Para obtener la inversa (X'X)-1 se debe primero activar la librería "library(MASS)", después utilizar el código "ginv()".

```
det(X_X)
library(MASS)
invX_X<-(ginv(X_X))
invX_X
```

Una vez que se tiene la inversa (X’X)-1 , se procede a obtener el producto X’y:

```
Xy<-trX %*% y1
Xy
```

Por último, se procede a calcular al vector beta a través del siguiente código:

```
beta<-invX_X %*% Xy
beta
```

Un método de comprobación para tener la certeza que este vector, el cual fue obtenido paso a paso mediante algebra lineal, es correcto, se utiliza el código para estimar de manera directa la regresión lineal “lm(y~x)”.

```
modelo<-lm(y~x2+x3)
summary(modelo)
```

Se aprecia que el vector beta encontrado coincide con los coeficientes estimados por el código “lm(y~x)”. Por tanto la ecuación estimada se define como sigue:

$$y = 1.381549 + 0.022279X2 - 0.003898X\ 3$$