

Trabajo Práctico 2

Modelo Lineal General y algunas extensiones

Contenidos: Modelo Lineal con K variables, notación matricial, test “F”.

Parte I ¹

1. Notación matricial: modelo con 2 variables

Considerar nuevamente el modelo de ahorro S e ingresos I de las familias propuesto en ejercicio 1 del Trabajo Práctico 1:

$$S_i = \alpha + \beta I_i + \mu_i \quad i=1, \dots, n$$

Por comodidad, los datos se reproducen a continuación:

Hogar N°	Ahorro S	Ingreso I
1	75	125
2	85	132
3	54	114
4	62	112
5	95	125
6	26	75
7	125	201
8	85	166
9	80	145
10	75	142
11	40	75
12	56	94

Se pide:

- Definir las matrices que intervienen en el cómputo de los estimadores MC de este modelo. Es decir: ¿Cuáles son las matrices? ¿Qué dimensiones tienen? ¿Cuáles son los elementos de estas matrices?
- Utilizando matrices, estime nuevamente los parámetros del modelo de ahorro e ingresos familiares.
- Obtenga el R^2 a partir de las correspondientes matrices y/o vectores.
- Explique qué significa en el modelo de dos variables el supuesto de que la matriz de variables explicativas tiene rango columna completo.

¹ Puede utilizar calculadora o planilla de cálculo para hacer sumas o productos de datos, pero las operaciones con matrices deben hacerse *a mano*. Para los ejercicios de la Parte I no puede utilizarse STATA.

2. Notación matricial: modelo con 3 variables

Las empresas pueden llevar adelante distintas estrategias con el objetivo de aumentar las ventas: promociones en base al precio, publicidad, etc. Cuando estas estrategias se adoptan en forma simultánea la pregunta relevante es cuál es el efecto de cada una de ellas, por separado, sobre las ventas. Este ejercicio presenta una aplicación particular de esta cuestión.

La base de datos *hamburguesas.xls* contiene información de ingresos semanales por ventas (en \$1000) – variable *V* –, gastos en publicidad (en \$1000) –variable *A*– y precios (en \$) –variable *P*– de hamburguesas a lo largo de 52 semanas de una importante cadena de hamburguesas.

Suponga un modelo lineal de los ingresos semanales por ventas en función de los gastos en publicidad y el precio de las hamburguesas.

Se pide:

- Definir las matrices que intervienen en el cómputo de los estimadores mínimo cuadráticos de ese modelo. Es decir: ¿cuáles son las matrices? ¿qué dimensiones tienen? ¿cuáles son los elementos de estas matrices?
- Utilizando matrices, computar los estimadores mínimos cuadráticos del modelo.
- Interpretar el significado económico de los coeficientes estimados.

3. Complete

A continuación se presenta una salida de regresión de la variable *mpg* en las variables *weight* y *price*. Lamentablemente se ha omitido deliberadamente cierta información (denotada con **XXX** en la tabla). Complete la información faltante y describa los pasos seguidos para obtenerla.

Source	SS	df	MS	Number of obs = XXX		
Model	XXXXXXXXXX	2	XXXXXXXXXX	F(2, 68) =	159.83	
Residual	1805.16823	XXX	XXXXXXXXXX	Prob > F =	0.0000	
				R-squared =	0.8671	
				Adj R-squared =	0.8617	
Total	XXXXXXXXXX	XXX	XXXXXXXXXX	Root MSE =	6.0696	

mpg	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
weight	2.984299	0.1669361	XXXXX	0.000	2.64883	3.31977
price	-6.641929	XXXXXXXXXX	-2.08	0.043	-13.0549	-0.2290
_cons	XXXXXXXXXX	6.482719	16.16	0.000	91.758	117.813

4. Ejercicio teórico con matrices

Suponga que $Y = Xb + u$, y que se cumplen todos los supuestos clásicos.

Considere el estimador $b = HY$, en donde H es una matriz no estocástica $K \times N$ (K =número de variables, N =observaciones)

Se pide:

- a) ¿Qué requisitos debería imponerle a H para que el estimador sea insesgado?
- b) ¿Cómo debería ser H para que el estimador sea el mejor estimador lineal insesgado?
- c) Obtenga la varianza de este estimador.

5. Demostración suma de residuos al cuadrado igual a cero

En el modelo lineal con k variables $Y = Xb + u$, en donde Y es un vector de n observaciones, X una matriz de K variables explicativas y n observaciones y u un vector de n términos residuales.

Muestre que si el número de observaciones es igual al número de variables explicativas (incluyendo la constante), entonces la suma de los residuos al cuadrado es igual a cero.

6. Ejercicio estimador b^*

Supongamos que $Y = Xb + u$ es la representación matricial del modelo lineal discutida en clase, en donde X es una matriz de n observaciones de las K variables explicativas. Considere el siguiente estimador de b , llamado b^* :

$$b^* = (Z'X)^{-1}Z'Y$$

en donde Z es cualquier matriz ($n \times K$) pero distinta de X

- a) Muestre que el estimador es insesgado.
- b) Justifique porque este estimador es ineficiente comparado con el estimador mínimo cuadrático.

Parte II²

7. Estimación de una función de producción

En este ejemplo intentaremos estimar los parámetros de una función de producción lineal simple:

$$Y_t = \beta_0 + \beta_1 K_t + \beta_2 L + u_t$$

en donde Y es la producción K es el stock de capital, L es el stock de trabajadores, y u_t es un término aleatorio. En la planilla LINGEN2.XLS se presentan 23 observaciones anuales (1950 a 1972) de las siguientes variables: Y = Producto Bruto Interno a costo de factores en millones de pesos constantes de 1960 (Fuente: BCRA), K = stock de capital valuado en millones de pesos de 1960 (Fuente: Urbisaia y Brufman (1985)) y L = personal ocupado en miles (Fuente: BCRA).

- Muestre en un gráfico la evolución temporal del ingreso, la fuerza laboral y el stock de trabajo. Comente.
- Estime los parámetros del modelo utilizando mínimos cuadrados ordinarios.
- Evalúe las hipótesis de significatividad individual de cada uno de los coeficientes.
- Construya intervalos de confianza para cada coeficiente.
- Evalúe la significatividad conjunta de los coeficientes.
- Calcular R^2 y R^2 ajustado. Comente.

8. Valuación de inmuebles e impacto de política ambiental

El objetivo de este ejercicio consiste en construir un modelo econométrico para explicar los valores de las casas en los alrededores de la ciudad de Boston. Los datos son los utilizados por Harrison y Rubinfeld (1978, 'Hedonic prices and the demand for clean air', Journal of Environmental Economics and Management, 5, 81-102), disponibles en Carter Hill, R., Griffiths, W., y Judge, G., 2001, Undergraduate Econometrics, 2nd edition, Wiley, New York.

La base de datos (*houses.xls*) contiene información sobre 500 localidades en las afueras de Boston, con las siguientes variables:

VALUE = valor mediano de las casas ocupadas en la localidad, en miles de pesos.

CRIME = crímenes per cápita en la localidad.

NITOX = concentración de óxidos nítricos (partes por millón).

ROOMS = promedio de habitaciones por vivienda.

² Los ejercicios de la Parte II están pensados para resolver usando STATA.

AGE = proporción de viviendas ocupadas construidas antes de 1940.

DIST = distancia promedio a los cinco centros principales del área de Boston.

ACCESS = índice de accesibilidad a las autopistas radiales del área de Boston

TAX = tasa de impuesto a la vivienda por cada \$10.000.

PTRATIO = cociente alumnos por maestro.

Se solicita lo siguiente:

- a) Estime un modelo lineal para explicar el valor de las casas (VALUE) utilizando como variables explicativas el resto de las variables proporcionadas.
- b) Interprete cuidadosamente los coeficientes de las variables CRIME, PTRATIO y TAX
- c) Comente acerca del coeficiente R^2 .
- d) El óxido nítrico es un importante contaminante del aire, por lo que el gobierno planea tomar acciones para bajar sus niveles. El argumento de los funcionarios estatales es que esta es una cuestión relevante para los ciudadanos ya que de hecho se observa que los precios de las casas en donde hay menores niveles de contaminación son notoriamente más altos, de lo que se deduce que las familias están dispuestas a pagar precios mayores por vivir en lugares menos contaminados. Consecuentemente, en los lugares más contaminados, las familias deberían estar dispuestas a pagar un impuesto de carácter local que permita financiar el plan de reducción de la contaminación. No obstante, una empresa local, sostiene que la contaminación no es tenida en cuenta por las familias cuando deciden dónde mudarse. Use el modelo para refutar o validar esta aseveración.
- e) Suponga que usted es el dirigente de la localidad 32, y en una de las reuniones informativas se enteró que el plan de reducción de la contaminación se implementará, y que de acuerdo a las estimaciones de los técnicos, logrará reducir la contaminación en un 5% con respecto a los valores existentes. Los vecinos de su localidad le realizan la siguiente pregunta: ¿en cuánto deberíamos esperar que aumente el precio de nuestra vivienda luego de haberse implementado el plan de reducción de la contaminación? Use el modelo estimado anteriormente para responder esta cuestión.
- f) Suponga que usted es propietario de una de las casas de la zona. En términos del valor de reventa, ¿cuánto es la máxima cantidad de dinero que invertiría en agregar una habitación adicional a su casa? Obviamente, utilice el modelo estimado para justificar su respuesta.

Notas computacionales:

Para realizar un test t de que un coeficiente es 0

test x

Para realizar un test t de que un coeficiente es igual a otro

test x=y

Para realizar un test t de que un coeficiente es igual a un número, por ejemplo 5

test x=5

Para realizar un test F de que los coeficientes de las variables x_1 y x_2 son simultáneamente iguales a cero

test $x_1 x_2$