Capítulo 2

Modelo de Regresión Lineal con k-Variables

2.1. El modelo

En muchas aplicaciones es natural pensar que una variable economica de interés pueda depender de mas de una variable exogena. Por esta razón veremos ahora como generalizar el modelo estudiado en el capitulo anterior. Estudiaremos en esta seccion el modelo de regresión lineal con k-variables. A diferencia del modelo simple este modelo admite mas de una variable como regresor. El modelo de regresión lineal con k-variables puede expresarse a traves de la siguiente ecuación

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} + u_i \tag{2.1}$$

donde i, como antes, se refiere a la observación i en nuestra muestra y

1. $y_i -$ es la variable que queremos explicar y recibe el nombre de *variable dependiente* o *variable explicada*.

2. $x_{ij}, \quad j = 1, ..., k -$ son las k variables a través de las cuales queremos explicar y y reciben como antes el nombre de *variables independientes* o *explicativas*.

3. $u_i -$ es el término de error, es una variables aleatoria y representa factores non observables distintos a x_j que afectan a y.

4. $\beta_j, \;\; j = 0, ..., k -$ son los parámetros del modelo.

En el caso de k variables es muy útil escribir el modelo en forma matricial. Sea n el numero de observaciones en nuestra muestra aleatoria. Definimos

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

un vector $(n \times 1)$ de todas las observaciones de la variable dependiente

$$\mathbf{X} = \begin{pmatrix} \mathbf{x_1} \\ \mathbf{x_2} \\ \vdots \\ \mathbf{x_n} \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{pmatrix}$$

una matriz $(n \times k)$ con todas las observaciones en las filas y las variables independientes en las columnas

$$\mathbf{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}$$

un vector $(n \times 1)$ de errores y

$$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{pmatrix}$$

un vector $((k+1) \times 1)$ donde el primer elemento es el coeficiente del término constante y los demás son los coeficientes de las variables exógenas. Entonces podemos escribir el modelo como

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u} \tag{2.2}$$

2.2. Estimación

2.2.1. Estimador Mínimos Cuadrados Ordinarios (MCO)

Como en el modelo simple, el único estimador que veremos en el modelo a k-variables es el estimador MCO. Para poder derivar el estimador necesitamos algunos resultados y definiciones preliminares. Sea $\hat{\beta}$ un estimador de β. Definimos $\mathbf{\hat{u}} = \mathbf{y} - \mathbf{X}\hat{\beta}$ el vector de residuos de regresión y $\mathbf{\hat{y}} = \mathbf{X}\hat{\beta}$ el vector de valores ajustados. Como antes el estimador de MCO es el estimador que minimiza la suma de los residuos al cuadrado $\mathbf{\hat{u}}'\mathbf{\hat{u}} = \sum_{i=1}^{n} \hat{u}_i^2$. Obtenemos

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \tag{2.3}$$

Se puede demostrar que cada uno de los estimadores contenidos en el vector $\hat{\beta}$ se puede escribir como

$$\hat{\beta}_j = \frac{\sum_{i=1}^{n} y_i \hat{r}_{ij}}{\sum_{i=1}^{n} \hat{r}_{ij}^2}$$

donde $\hat{r}_j$ es el residuo de la regresión de x_j sobre las demás variables independientes. Este resultado es muy útil para poder entender la interpretación de cada uno de los elementos en el vector $\hat{\beta}$. Dado que $\hat{r}_j$ es la parte de x_j que no esta relacionada con las demás variables independientes, $\hat{\beta}_j$ mide el efecto de x_j sobre y una vez que descontamos los efectos de las demás variables exógenas. Por eso podemos interpretar $\hat{\beta}_j$ como el efecto parcial de x_j sobre y o sea el efecto de x_j cuando las demós variables se mantienen fijas.

Ejemplo 2.1 Extendemos el modelo utilizado en el Ejemplo 1.1 para explicar el salario horario incluyendo otro regresor: el tiempo trabajado en el actual puesto de trabajo. Utilizando el estimador que acabamos de ver obtenemos las siguientes estimaciones

$$\log(\widehat{salario_i}) = 0{,}216 + 0{,}097 educ + 0{,}010 exper.$$

La interpretación de $\hat{\beta}_1 = 0{,}097$ es que ahora un año mas de educación produce un incremento del salario de 9.7 %, mientras que un año mas de experiencia laboral aumenta el salario en un 1 %.

El ejemplo anterior evidencia una caracteristica muy importante del modelo a k-variables. En general, si añadimos una variable al modelo las estimaciones de los pará metros que ya eran incluidos cambian. Ejemplo en GRETL

Hay dos excepciones a este resultado. Se consideren dos modelos, el primero con una sola variable dependiente, $y = \beta_0 + \beta_1 x_{i1} + u_i$, y el segundo con dos variables independientes $y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + u_i$.

(1) Si $\hat{\beta}_2 = 0$ entonces el estimador de β_1 coinciden en los dos modelos.

(2) Un segundo caso en que $\hat{\beta}_1$ coincide en los dos modelos es cuando x_1 y x_2 no están correlacionadas. El estimador MCO de β_1 en el segundo modelo es $\hat{\beta}_i = \frac{\sum_{i=1}^{n} y_i \hat{r}_{i1}}{\sum_{i=1}^{n} \hat{r}_{i1}^2}$ donde $\hat{r}_{i1}$ es el residuo de la regresión de x_1 sobre x_2. Si los dos regresores no están correlacionados esto significa que la covarianza es cero y que el efecto de x_2 sobre x_1 es cero. Pero esto implica que $\hat{r}_{i1} = x_{i1} - \hat{\beta}_0 = x_{i1} - \bar{x}_1$. Resulta claro entonces que los estimadores de $\hat{\beta}_1$ coinciden en los dos modelos.

2.2.2. Error estandard

Para la varianza del error, el estimador insesgado que utilizaremos en este capitulo es parecido al anterior. La única diferencia es que ahora hay que normalizar por el numero de observaciones menos el numero total de parámetros. Así que en el modelo de k-variables el estimador de la varianza del error es

$$\hat{\sigma}^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k-1} \tag{2.4}$$

2.3. Supuestos y propiedades

2.3.1. Supuestos

1. Modelo de regresión.
2. Las variables salen de un muestreo aleatorio.
3. Zero esperanza condicional $E(u|x_1, ..., x_k) = 0$
4. No hay perfecta multicolinalidad.

Perfecta multicolinealidad significa que una o mas variables se pueden escribir como combinacíon lineal de las demas variables. Por ejemplo $x_1 = 2 + 0{,}5x_2 + 0{,}8x_3$. En este caso el modelo no se puede estimar.

2.3.2. Propiedades de los estimadores

Propiedades algebraicas

Además de **P1** y **P3** vale la siguiente propiedad

P1' $\mathbf{X}'\hat{\mathbf{u}} = 0$

$$\begin{aligned}
\mathbf{X}'\hat{\mathbf{u}} &= \mathbf{X}'(\mathbf{y} - \mathbf{X}\hat{\beta}) \\
&= \mathbf{X}'\mathbf{y} - \mathbf{X}'\mathbf{X}\hat{\beta} \\
&= \mathbf{X}'\mathbf{y} - (\mathbf{X}'\mathbf{X})(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= 0
\end{aligned}$$

Propiedades estadisticas

Vamos a estudiar ahora las propiedades estadísticas de los estimadores MCO. Podemos ahora volver escribir en forma matricial el supuesto **S2'** y **S3'**. El primero implica que $E(\mathbf{u}) = 0$, el segundo que $Var(\mathbf{u}) = \sigma^2 I$ donde I es una matriz identidad de dinemsión $n \times n$.

P2' *Los estimadores MCO son insesgados*, $E(\hat{\beta}) = \beta$.

$$\begin{aligned}
\hat{\beta} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \mathbf{u}) \\
&= (\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u} \\
&= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}
\end{aligned}$$

Tomando el valor esperado obtenemos

$$\begin{aligned}
E(\hat{\beta}) &= E[\beta) + E((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}] \\
&= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}) \\
&= \beta
\end{aligned}$$

donde la ultima igualdad se obtiene utilizando la propiedad *P1'*.

P3' $Var(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

Utilizando la definición de varianza de un vector de variables aleatorias

$$\begin{aligned}
Var(\hat{\beta}) &= E\left[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'\right] \\
&= E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right]
\end{aligned}$$

Desde la propiedad anterior sabemos que

$$\hat{\beta} - \beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}$$

Entonces

$$\begin{aligned}
Var(\hat{\beta}) &= E\left[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{u}\mathbf{u}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\right] \\
&= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}\mathbf{u}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \\
&= \sigma^2(\mathbf{X}'\mathbf{X})^{-1}
\end{aligned}$$

P4' *El estimador de la varianza del termino de error visto antes es insesgado* $E(\hat{\sigma}^2) = \sigma^2$.

P5' *Teorema de Gauss-Markov* $\hat{\beta}$ es el estimador lineal insesgado optimo (ELIO) de β.

2.4. Bondad del ajuste

Para poder estudiar la bondad del ajuste en el caso de k-variables necesitamos algunos resultados algï¿½bricos relativos a nuestro modelo. Primero, podemos observar que

$$\begin{aligned} STC &\equiv \sum_{i=1}^{n}(y_i - \bar{y})^2 = \mathbf{y}'\mathbf{y} - n\bar{y}^2 \\ SEC &\equiv \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2 = \hat{\mathbf{y}}'\hat{\mathbf{y}} - n\bar{y}^2 \\ SCR &\equiv \sum_{i=1}^{n}\hat{u}_i^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}} \end{aligned}$$

$$\begin{aligned} R^2 &= \frac{STC}{SEC} \\ &= 1 - \frac{SCR}{STC} \end{aligned} \tag{2.5}$$

Como interpretamos el R^2 en este caso? Como anteriormente R^2 se interpreta como *la fracción de la variación muestral de y_i explicada por todas las x_i conjuntamente.*

Existe una relación muy importante entre la varianza de estimador MCO y R^2. De hecho podemos escribir la varianza de cada uno del los $\hat{\beta}_j$ como

$$Var(\hat{\beta}_j) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2(1 - R_j^2)} \tag{2.6}$$

donde R_j^2 es el R-cuadrado que se obtiene de la regresión de x_j sobre todas las demás variables exógenas.

Se puede demostrar que el R^2 aumenta si incluyimos mas variables exogenas en el modelo.