# ECTRÍA-T1. MLG

## 1. INTRODUCCIÓN – EL MODELO LINEAL GENERAL.

La Econometría se puede definir como la aplicación de métodos matemáticos y estadísticos al análisis de datos económicos con el fin de dar contenido empírico a las teorías económicas y verificarlas o refutarlas.

Se sirve de un modelo, el modelo econométrico, que trata de ser una representación simplificada de la realidad económ, y que para ser operativo ha de estar expresado en forma matemática.

El modelo econométrico ha de ser capaz de:

- especificar las relación existente entre las var. económicas
- utilizar información muestral de estas variables que cuantifique la dependencia,
- evaluar la validez de los modelos de la Tª Económica,
- efectuar un seguimiento coyuntural y de previsión.

Para conseguir estos objetivos se siguen estos pasos:

1_ Formular un modelo simplificado que reproduzca los patrones de comportamiento de las var. económicas.
2_ Estimar los parámetros del modelo.
3_ Contrastar las hipótesis económicas que sean relevantes.
4_ Realizar predicciones.

El análisis de regresión (*) es una de las técnicas más utilizadas en Econometría. Se trata de descubrir la relación existente entre las var. que queremos estudiar.

$Y \equiv$ var. dependiente ó var. endógena

y un cjto de variables $X_1 \dots X_K \equiv$ var. indep, exógenas ó explicativas ó regresores.

EL Modelo Lineal General es el modelo de regresión más sencillo. Se basa en descomponer el comportamiento de Y en Señal + Ruido, donde la Señal recoge el comportamiento determinista explicado por X y el Ruido recoge el comport. aleatorio o estocástico (perturbación aleatoria). $Y_t = X_t + u_t$

~~Cada observac~~

Para cada valor de Y esta relación se puede expresar mediante la ecuación

$$Y_{i(t)} = \beta_1 + \beta_2 X_{2i(t)} \dots + \beta_K X_{Ki(t)} + u_{i(t)} \qquad (i/t)$$

para $i = 1 \dots N$, en datos de corte transversal

para $t = 1 \dots T$, en datos de serie temporal (a partir de ahora, es k)

En muchas ocasiones, el modelo incorpora una cte,

$\to \beta_1$, se ~~modeliza~~ interpreta acompañando a una 1ª var. explicativa que siempre vale 1, $X_{1i/t} = 1 \ \forall i, \forall t$.

$\to \beta_j$ = parámetros del modelo. Son los coef. de las var. explicativas que interpretan (o recogen) el impacto de $X_i$ sobre $Y_i$.

$j = 2 \dots K$

$\to u_{i(t)} \equiv$ término de error del modelo. Es una variable aleatoria.

(*) El término regresión fue utilizado por 1ª vez por Galton, al estudiar la talla de las familias y observar que los hijos "regresaban" a la talla de sus padres.

## 2. ESPECIFICACIÓN del MODELO

El modelo se puede escribir en notación matricial:

$$\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{pmatrix}_{T\times 1} = \begin{pmatrix} 1 & X_{21} & \cdots & X_{K1} \\ 1 & X_{22} & \cdots & X_{K2} \\ \vdots & \vdots & & \vdots \\ 1 & X_{2T} & \cdots & X_{KT} \end{pmatrix}_{T\times K} \cdot \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_K \end{pmatrix}_{K\times 1} + \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_T \end{pmatrix}_{T\times 1}$$

En forma reducida

$$\underset{T\times 1}{Y} = \underset{T\times K}{X}\,\underset{K\times 1}{\beta} + \underset{T\times 1}{u}$$

Se considera que el $\boxed{\text{MLG}}$ satisface los siguientes supuestos:

1. Respecto al modelo (3)

a) Es estocástico. Por ser Y función de u.
~~No lineal~~ si
Debido a que el modelo es solamente una aproximación, está sujeto a errores de medida, no incluye todos los factores que influyen sobre Y.

b) Es lineal. Es lo más sencillo. Se considera la relac. lineal como una aproximación,
Tb. se pueden hacer transformaciones (ln) para hacerlo lineal. → modelos no lineales

c) Relación causa-efecto. *En Economía es muy frecuente ⇄
$X_1 \ldots X_K \equiv$ causas
$Y \equiv$ efecto.

2. Respecto al término de error (perturbaciones):

a) Es aleatorio. (misma explicación que 1a).

b) $E[u_t] = 0, \forall t \Rightarrow E[u] = 0_T$
Si fuera $\neq 0$, habría que incorporarla ~~a~~ de la parte determinist.
Una posible causa es la no incorporación de una var. explic. relevante.

c) $\left.\begin{array}{l} V[u_t] = \sigma_u^2, \forall t \\ E[u_t u_s] = 0, \forall t \neq s \end{array}\right\} V[u] = \sigma_u^2 I_T$

Hipótesis de homoscedasticidad y de incorrelación.

Observación: Si $\sigma_u^2 \neq$ cte, $\forall t \Rightarrow$ Heteroscedasticidad
Si $E[u_t u_s] \neq 0 \Rightarrow$ Autocorrelación

d) $u \longrightarrow N(0_T, \sigma_u^2 I_T)$. Hptsis de normalidad.

Estas cuatro hipótesis se pueden resumir diciendo que $u$ es un ruido blanco.

3. Respecto a la parte determinista

a) $X_i$ son deterministas. Para muestras ≠s, los valores de $X_i$ serían los mismos

Observación: Si se trabaja con datos temporales, aparecen valores retardados de la var. endógena, $Y_{t-k}$
No tiene sentido llamarles ~~valores~~ var. exógenas, se llaman var. predeterminadas.

b) $X_i$ no son linealmente dependientes.
Excluye que una var. se pueda escribir como c.l. de las demás. Se admite cierta correlación, pero no dep. total.
Se evita información redundante $\Rightarrow rg(X) = k < T$
(> observac que parám)

c) los coef. $\beta_i$ son constantes en el tiempo.
Presupone estabilidad en el tiempo, que no tiene por qué ser cierto, pero sí + sencillo.

Si las var. explicativas no son todas indep $\Rightarrow$ Multicolinealidad
Si sólo se consideran retardos $\Rightarrow$ modelos univ. de series temp
Si hay mezcla de var. exógenas y retardos $\Rightarrow$ modelos dinámicos
Si hay bidireccionalidad $\Rightarrow$ modelos de ec. simultáneas
Si la relac. es no lineal $\Rightarrow$ modelos no lineales

# ECTRÍA - T1

## 3- ESTIMADORES MÍNIMO CUADRÁTICO ORDINARIOS. PROPIEDADES

Una vez especificado el modelo $Y = X\beta + u$ hay que estimar los parámetros $\beta_1, \beta_2 \dots \beta_K$ (término independiente y pendientes de la regresión) → IMP. para la PREDICCIÓN y $\sigma_u^2$ → IMP. para $V(\hat{\beta})$ y CONTRASTACIÓN

Para cada valor observado de $Y$, tendremos su valor teórico obtenido con la parte determinista. Llamamos error o residuo a la diferencia entre ambos.

$$Y = X\beta + u \longrightarrow \hat{Y} = X\hat{\beta} \longrightarrow \hat{u} = Y - \hat{Y}$$

Modelo original con los valores verdaderos — Modelo estimado — Residuo

De entre todos los posibles valores de $\beta$, elegiremos aquéllos que minimicen el error total, expresado éste como la suma de los cuadrados de los residuos. Matricialmente,

MCO $\boxed{\min_{\beta} \underbrace{\hat{u}'\hat{u}}_{N^o}} = \min_{\beta} (Y-\hat{Y})'(Y-\hat{Y}) = Y'Y - 2\hat{\beta}'X'Y + \hat{\beta}'X'X\hat{\beta}$

$= \min_{\beta} SR(\hat{\beta}) \equiv$ suma residual.

Es un problema de optimización clásica, en el que la C.N. nos da $\hat{\beta} = (X'X)^{-1}X'Y$ y la C.S. se cumple siempre, al ser $(X'X)$ una matriz definida positiva.

## PROPIEDADES

$\hat{\beta}$ es un vector aleatorio, pues depende de $Y$

Bajo los supuestos anteriores podemos afirmar que es un estimador lineal y eficiente de $\beta$. Lo vemos:

2. $E[u] = 0_T \Rightarrow E[\hat{\beta}] = \beta$

$E[\hat{\beta}] = E[(X'X)^{-1}X'Y] = \quad\quad = \beta$

1. Es lineal de $\beta$

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

3. Si $Var(u) = \sigma_u^2 I_T$, entonces $Var(\hat{\beta}) = \sigma_u^2 (X'X)^{-1}$

$Var(\hat{\beta})$ (matriz) $= E[(\hat{\beta} - E(\hat{\beta}))(\hat{\beta} - E(\hat{\beta}))'] = V$

$= E[(\beta + (X'X)^{-1}X'u - \beta)(\beta + (X'X)^{-1}X'u - \beta)'] =$

$= (X'X)^{-1}X'E(uu')X(X'X)^{-1} = \sigma_u^2 (X'X)^{-1}$.

4. Tma. Gauss-Markov: El estimador MCO es el estimador lineal insesgado óptimo, en el sentido de que tiene varianza mínima.

Al ser $Var(\hat{\beta}) = \sigma_u^2 (X'X)^{-1}$ y $\sigma_u^2$ desconocido, hay que estimarla para:
— precisión de $\hat{\beta}$
— fase de contrastación

(iii) $\hat{\sigma}_u^2 = \frac{\hat{u}'\hat{u}}{T-K} = \frac{SR}{T-K}$ $\quad E[\hat{\sigma}_u^2] = E\left[\frac{\hat{u}'\hat{u}}{T-K}\right] = \frac{1}{T-K}E[\hat{u}'\hat{u}] = \frac{T-K}{T-K}\sigma_u^2$.

Esta estimación no proviene de ningún criterio de mín., sino del interés de obtener un estimador insesgado a partir de los errores mínimo cuadráticos.

Bajo el supuesto de normalidad, $\cancel{\hat{\sigma}_u^2 \to \chi^2_{T-K}}$ $\quad (T-K)\frac{\hat{\sigma}_u^2}{\sigma_u^2} \to \chi^2_{T-K}$

$\frac{SR}{\sigma_u^2} \to \chi^2_{T-K}$

## 4_ CONTRASTES de NORMALIDAD

Las propiedades (→ distrib. de $\hat{\beta}$ y $\hat{\sigma}_u^2$) anteriores se basan en el supuesto de que la perturbación aleatoria del modelo se comporta con arreglo a una normal, por lo que sería bueno contrastarlo.

El contraste de Bera y Jarque (1981) se basa en la simetría y curtosis de la normal, utilizando (→ los residuos como una estimación del término de error) el estadístico

$$BJ = \underset{\uparrow \text{ tamaño muestral}}{T}\left(\frac{(\text{asimetría})^2}{6} + \frac{(\text{curtosis} - 3)^2}{24}\right)$$

Para muestras grandes, $BJ \longrightarrow \chi^2_2$.

$H_0$: $u$ es Normal
$H_1$: $H_0$ no es cierta
} si $BJ > K \Rightarrow$ rechazo $H_0$

Si $H_0$ cierta, $u \longrightarrow N(0_T, \sigma_u^2 I_T)$

$$\hat{\beta} \longrightarrow N(\beta_K, \sigma_u^2 (X'X)^{-1}).$$

$$\frac{(T-K)\hat{\sigma}_u^2}{\sigma_u^2} \longrightarrow \chi^2_{T-K}$$

## 5. ESTIMADOR MÁXIMA VEROSIMILITUD

Se trata de encontrar las estimaciones de los parámetros del modelo $\beta_1, \beta_2 \dots \beta_K$ y $\sigma_u^2$, de forma que hagan la información muestral más verosímil, que maximicen la probabilidad de obtener la muestra dada.

Necesitamos una distrib. de probabilidad, seguimos con $u_i \longrightarrow N(0, \sigma_u^2)$ v.a.i.i.d.

La función de verosimilitud conjunta, al ser indep. se obtiene multiplicando las funciones de densidad marginales.

$\max \mathcal{L} \sim \max \ln \mathcal{L} \longrightarrow$ C.N. $\rightsquigarrow \hat{\beta}_{MV} = \hat{\beta}_{MCO}$

$$\mathcal{L}(Y, X, \beta, \sigma_u^2) = \frac{1}{(\sqrt{2\pi\sigma_u^2})^T} \cdot e^{-\frac{1}{2\sigma_u^2} \sum^T \underbrace{(Y-X\hat{\beta})'(Y-X\hat{\beta})}_{u^2}}$$

$$\ln \mathcal{L}(Y, X, \beta, \sigma_u^2) = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln \sigma_u^2 - \frac{1}{2\sigma_u^2}(Y-X\beta)'(Y-X\beta).$$

$Y'Y - 2\beta'X'Y + \beta'X'X\beta$

$$\frac{\partial \ln \mathcal{L}}{\partial \beta} = 0 \iff -\frac{1}{2\sigma_u^2}[-2X'(Y-X\beta)] = 0_K$$

utilizamos $\hat{\beta}_{MV}$

$$\frac{\partial \ln \mathcal{L}}{\partial \sigma_u^2} = 0 \iff -\frac{T}{2\sigma_u^2} + \frac{1}{2\sigma_u^4}\underbrace{(Y-X\hat{\beta})'(Y-X\hat{\beta})}_{\hat{u}'\hat{u}} = 0$$

cuyas soluciones son:

$$\hat{\beta}_{MV} = (X'X)^{-1}X'Y = \hat{\beta}_{MCO} \Rightarrow \text{tiene las mismas propiedades.}$$

$$\hat{\sigma}^2_{u\,MV} = \frac{\hat{u}'\hat{u}}{T} \neq \hat{\sigma}^2_{u\,MCO}.$$

$(\hat{\sigma}_u^2)_{MV}$ no es insesgado, pero su varianza es menor.

~~mínima, alcanza~~

menor a la CCR

pero sesgado

# 6. ERRORES de ESPECIFICACIÓN

☒ Las var. explicativas incluidas en el modelo **NO** son las correctas.

1 _ Omisión de var. relevantes.

Las estimaciones MCO de los coef. de las var. incluidas serán sesgadas.

$\hat{\beta}_{MCO}$ sesgadas.

$\hat{\sigma}^2_{u\,MCO} >> \sigma^2_u$.

2 _ Inclusión de var. irrelevantes

$\hat{\beta}$ de las var. irrelevantes tiene media 0. Se espera que no resulten significativas.

Var ($\hat{\beta}$ relevantes) aumenta $\Rightarrow$ se pierde precisión y se puede pensar que ~~son irre~~ no son importantes.

---

Si se puede comparar $\neq$s especificaciones, elegir lo que tenga una SR menor, que equivale a elegir lo que tenga mayor coef. determinación corregido.