

Modelos no lineales

Román Salmerón Gómez

Grado en Economía

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Contenidos

Introducción

Especificaciones no
lineales

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Introducción

- Contenidos
- Introducción
- Especificaciones no lineales
- Modelos intrínsecamente lineales
- Modelos intrínsecamente no lineales

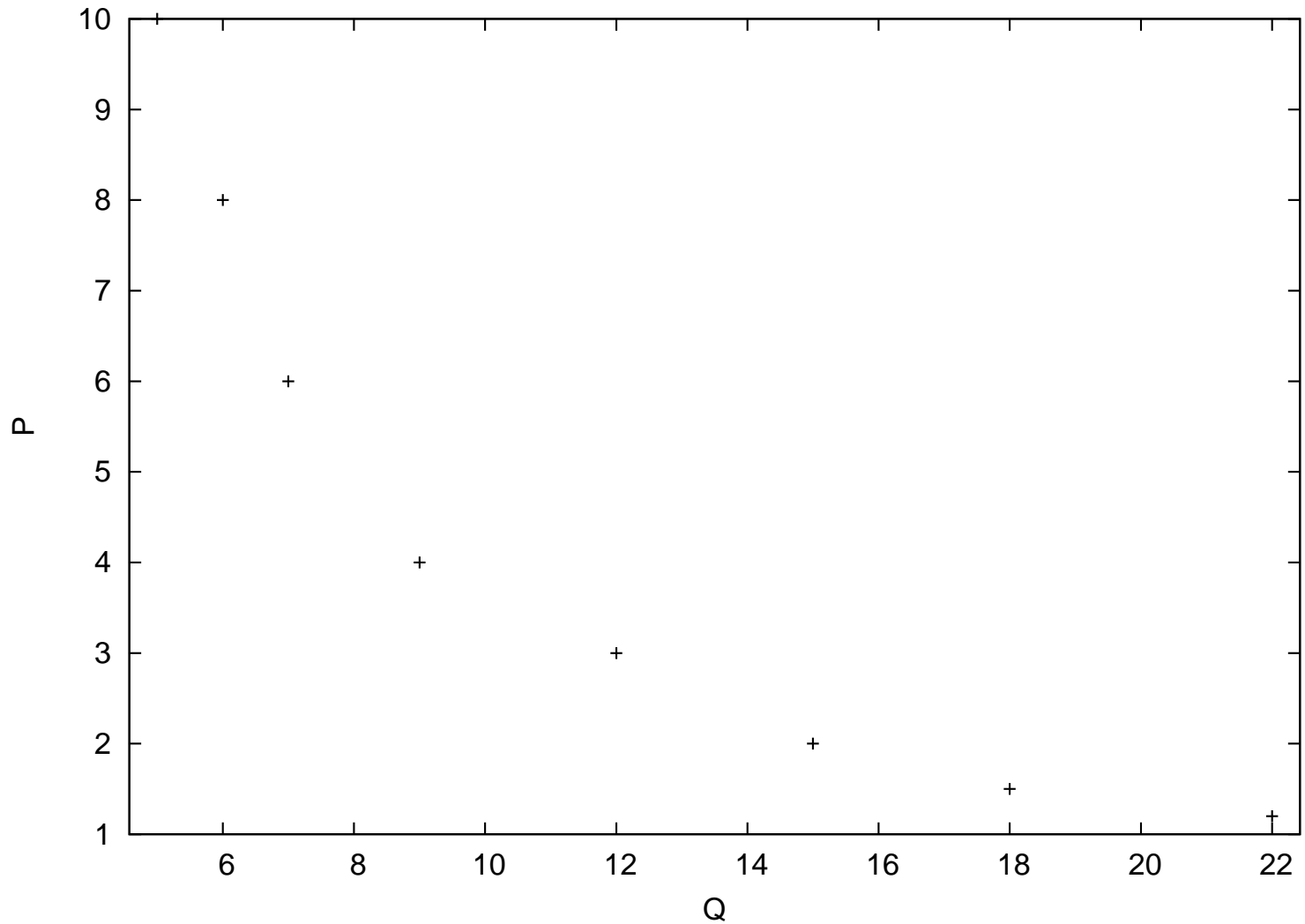


Figura 1: Relación no lineal

Hasta ahora, el efecto sobre Y de un cambio unitario en X no dependía del valor de X . ¿Qué ocurre si el efecto sobre Y de un cambio en X depende del valor de una (o más) de las variables independientes?. En este caso, la función de regresión es no lineal. Una función no lineal es una función con una pendiente que no es constante. Hay distintos tipos de especificaciones no lineales:

1. Modelos intrínsecamente lineales (se pueden estimar por los métodos conocidos hasta ahora haciendo un simple cambio de variable y/o transformación):

$$y_i = \beta_1 + \beta_2 x_i^2 + \epsilon_i, \quad y_i = \beta_1 + \beta_2 \frac{1}{x_i} + \epsilon_i, \quad y_i = \beta_1 x_i^{\beta_2} e^{\epsilon_i}.$$

2. Modelos intrínsecamente no lineales (no se pueden estimar por los métodos conocidos hasta ahora debido a la no linealidad existente en los parámetros):

$$y_i = \beta_1 + x_i^{\beta_2} + \epsilon_i, \quad y_i = \beta_1 x_i^{\beta_2} + \epsilon_i, \quad y_i = \frac{\beta_1}{1 + e^{\beta_2 + \beta_3 x_i}} + \epsilon_i.$$

Supongamos la siguiente información acerca de precios, P , y cantidades, Q , demandadas:

Q	P	$\ln Q$	$\ln P$
5	10	1.609	2.302
6	8	1.791	2.079
7	6	1.945	1.791
9	4	2.197	1.386
12	3	2.484	1.098
15	2	2.708	0.693
18	1.5	2.890	0.405
22	1.2	3.091	0.182

En la Figura 1 se observaba que la relación entre P y Q no es lineal, sin embargo, haciendo una transformación sencilla (aplicando logaritmos) se puede obtener dicho tipo de relación (Figura 2).

Contenidos

Introducción

Especificaciones no
lineales

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

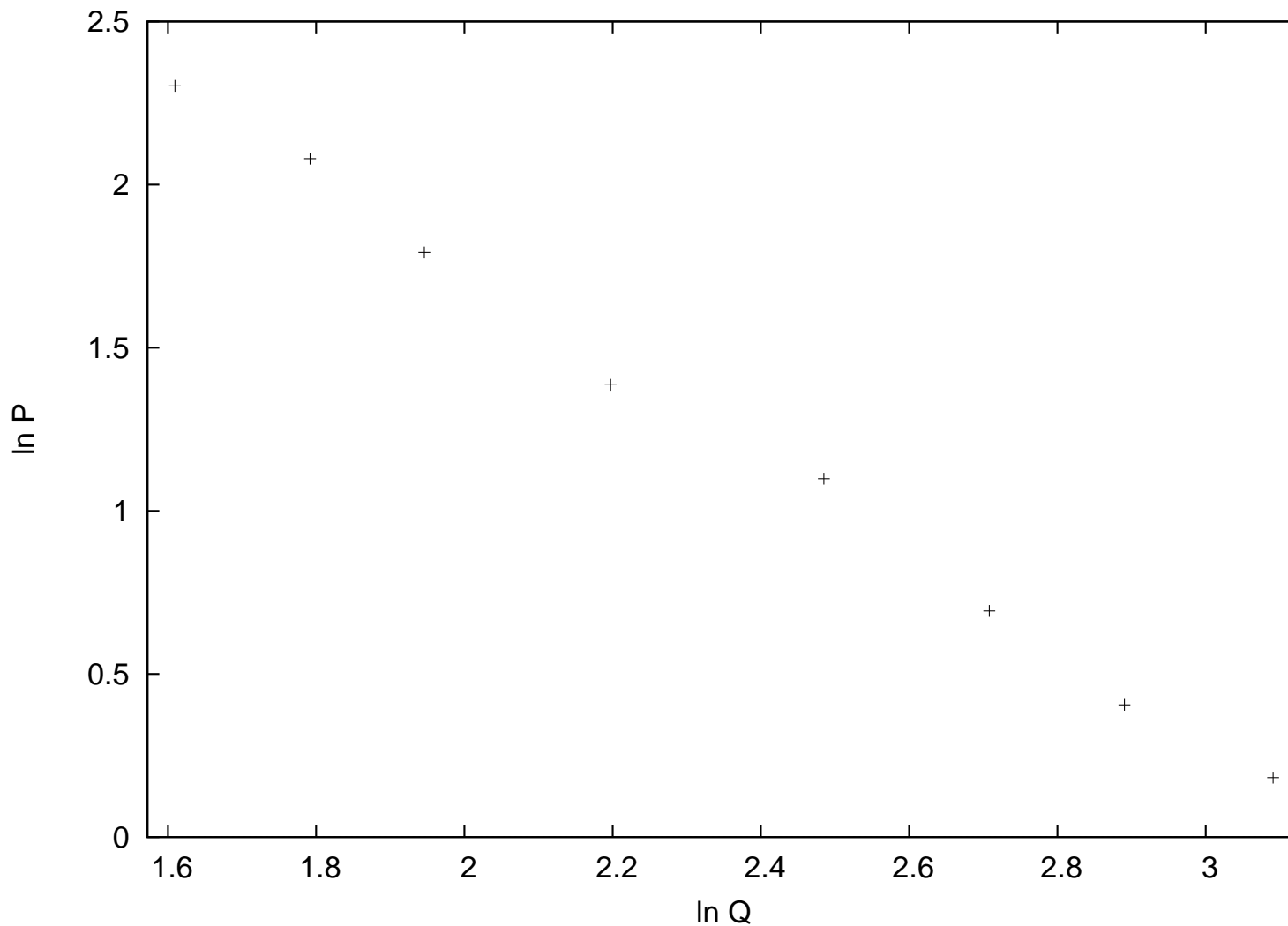


Figura 2: Relación lineal

Contenidos

Introducción

**Modelos
intrínsecamente lineales**

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

Modelos intrínsecamente lineales

La transformación de Box-Cox permiten corregir la no linealidad en la relación entre variables (también pueden ser usadas para solucionar problemas de normalidad y heterocedasticidad). A partir de la transformación de Box-Cox:

$$y^{(\lambda_1)} = \begin{cases} \frac{y^{\lambda_1} - 1}{\lambda_1} & \lambda_1 \neq 0 \\ \ln y & \lambda_1 = 0 \end{cases}, \quad x^{(\lambda_2)} = \begin{cases} \frac{x^{\lambda_2} - 1}{\lambda_2} & \lambda_2 \neq 0 \\ \ln x & \lambda_2 = 0 \end{cases}$$

modelos en principios no lineales podrán expresarse linealmente como sigue:

$$y^{(\lambda_1)} = \alpha + \beta x^{(\lambda_2)} + \epsilon.$$

Algunos casos particulares interesantes son los siguientes:

- Modelo lineal: $\lambda_1 = \lambda_2 = 1$.
- Modelo doblemente logarítmico: $\lambda_1 = \lambda_2 = 0$.
- Modelo semilogarítmico (log-lineal): $\lambda_1 = 0$ y $\lambda_2 = 1$.
- Modelo semilogarítmico (lineal-log): $\lambda_1 = 1$ y $\lambda_2 = 0$.
- Modelo hiperbólico: $\lambda_1 = 1$ y $\lambda_2 = -1$.

El modelo Box-Cox: casos particulares

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

- Modelo lineal: $\lambda_1 = \lambda_2 = 1$.

En este modelo, β representa el efecto marginal, es decir, una variación unitaria en la variable x provoca un cambio en la variable y igual a β :

$$\beta = \frac{\partial y}{\partial x} \rightarrow \Delta y = \beta \Delta x.$$

- Modelo doblemente logarítmico: $\lambda_1 = \lambda_2 = 0$.

En este caso, Box-Cox coincide con el modelo no lineal: $y = Ax^\beta e^\epsilon$. Aquí, β representa la elasticidad de y respecto de x (es decir, qué incremento porcentual de y se tendrá si se produce un incremento porcentual de x):

$$\frac{\partial y}{\partial x} = A\beta x^{\beta-1} = A\beta x^\beta x^{-1} = \beta y x^{-1} = \beta \frac{y}{x} \rightarrow \beta = \frac{\partial y}{\partial x} \frac{x}{y}.$$

El modelo Box-Cox: ejemplos gráficos

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

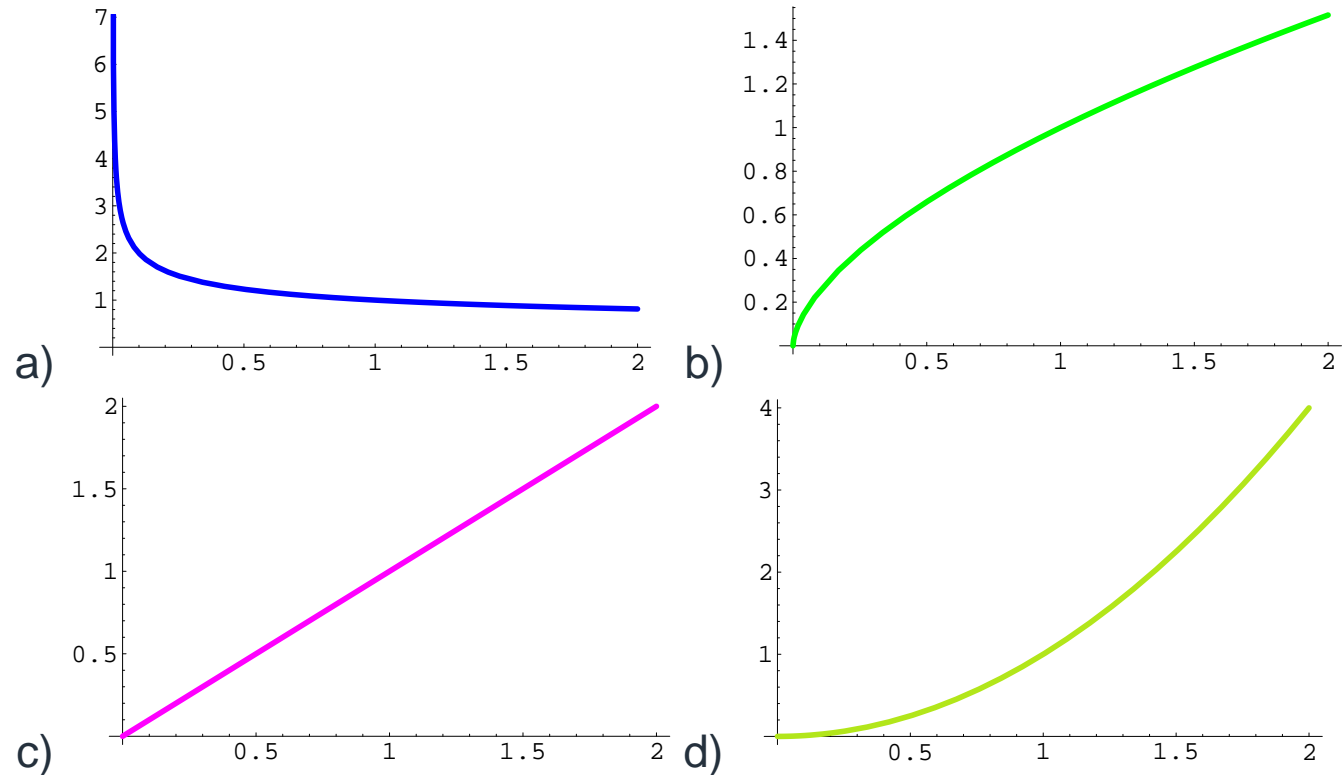


Figura 3: Modelo doblemente logarítmico con a) $\beta = -0.6$, b) $\beta = 0.6$, c) $\beta = 1$ y d) $\beta = 2$.

El modelo Box-Cox: ejemplos gráficos

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

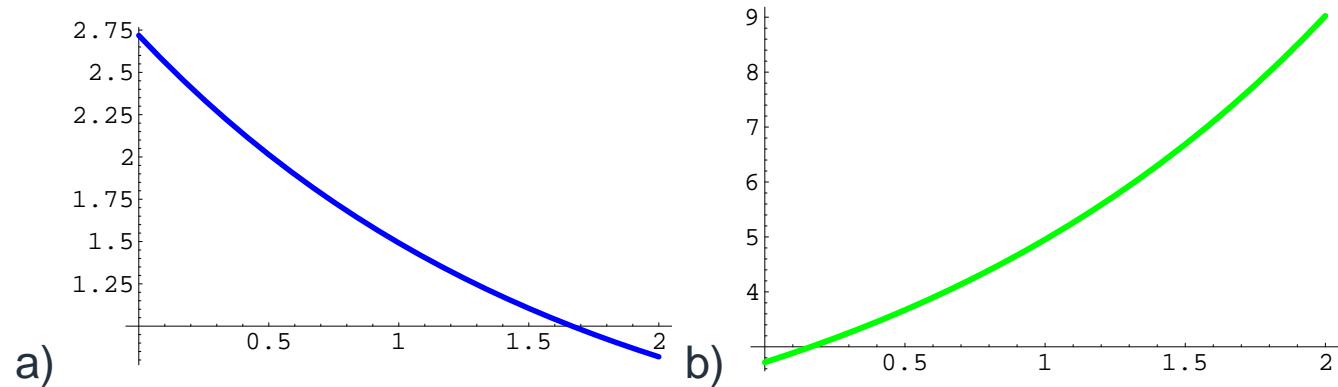


Figura 4: Modelo semilogarítmico en y con a) $\beta = -0.6$ y b) $\beta = 0.6$.

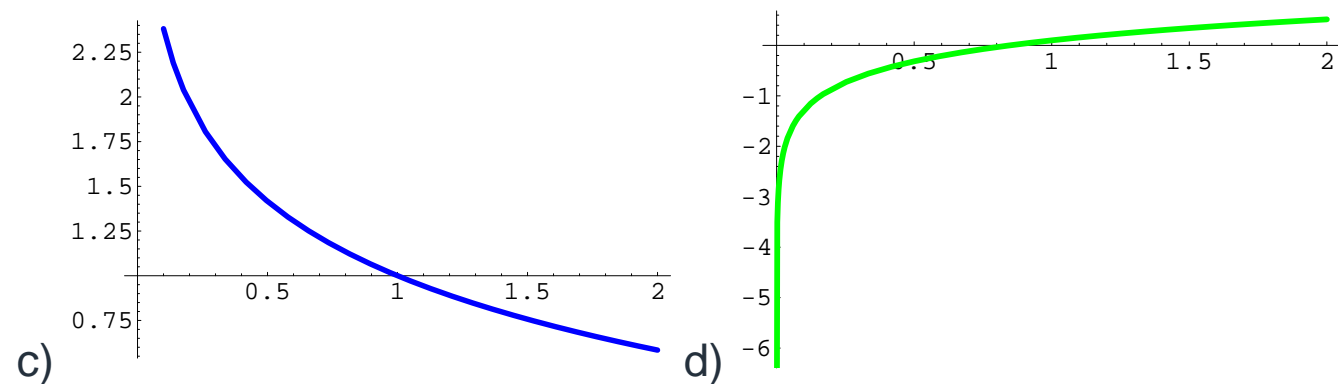


Figura 5: Modelo semilogarítmico en x con c) $\beta = -0.6$ y d) $\beta = 0.6$.

El modelo Box-Cox: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

Dados los siguientes datos ajustar un modelo doblemente logarítmico:

x	1	2	3	4	5
y	4	50	200	740	3000

El modelo doblemente logarítmico responde a la expresión $y = Ax^\beta e^\epsilon$, el cual se puede linealizar sin más que considerar logaritmos ($y^* = \ln y$ y $x^* = \ln x$):

$$\ln y = \ln A + \beta \ln x + \epsilon \ln e \Rightarrow y^* = A^* + \beta x^* + \epsilon.$$

Por tanto, se ha llegado a un modelo lineal que se puede estimar por el método de Mínimos Cuadrados Ordinarios:

$$\hat{A}^* = 1'2257 \Rightarrow \hat{A} = e^{1'2257} = 3'4066$$

$$\hat{\beta} = 3'9856, \quad R^2 = 0'865$$

El modelo estimado es $\hat{y} = 3.4066 \cdot x^{3.9856}$.

x^*	y^*
0	1'386
0'693	3'912
1'099	5'298
1'386	6'607
1'609	8'006

Ajustar un modelo del tipo $y = A\beta^x e^\epsilon$ y decidir cuál es el mejor.

El modelo Box-Cox: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

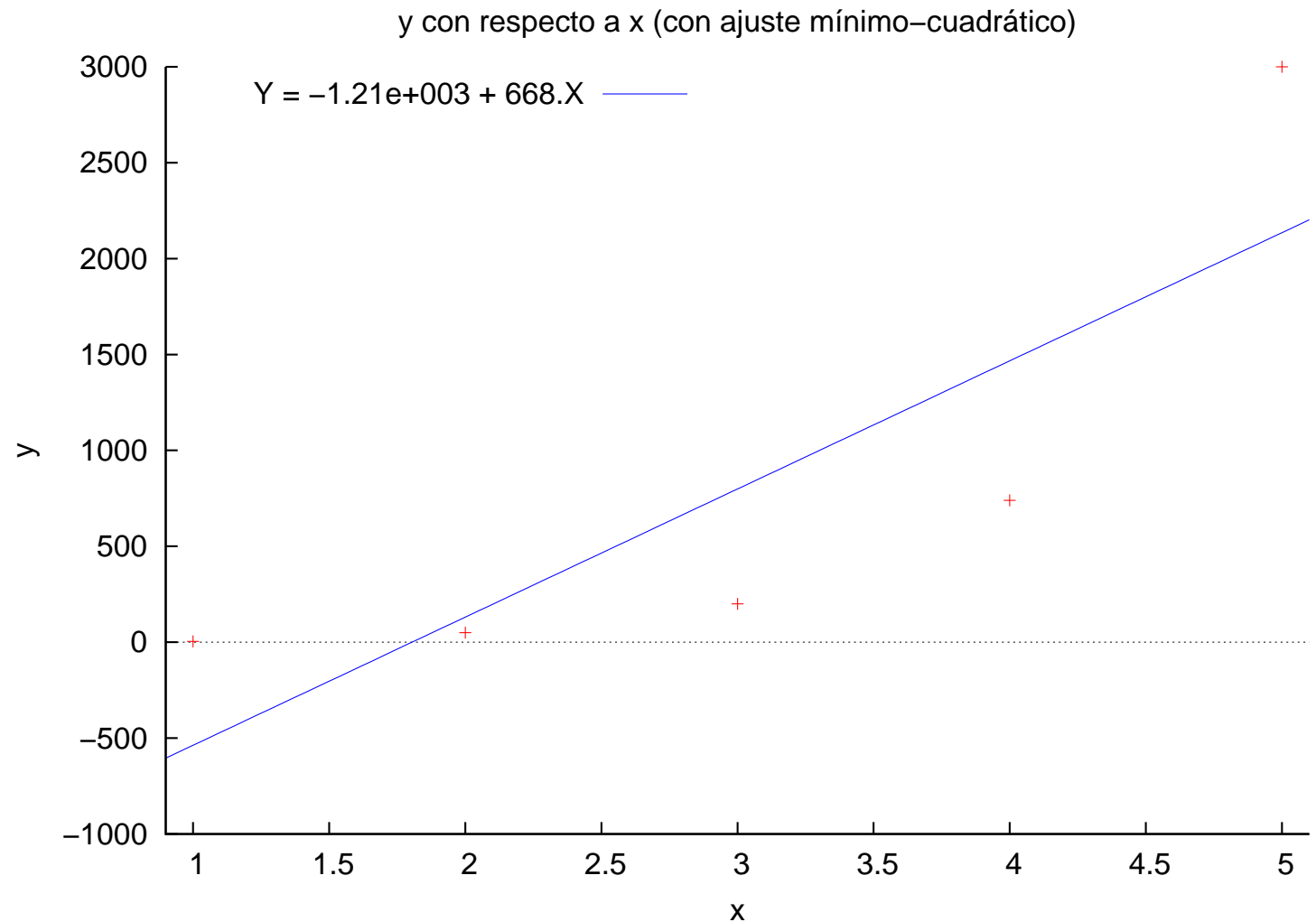


Figura 6: Representación gráfica de los datos originales.

El modelo Box-Cox: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

El modelo Box-Cox

Modelos
intrínsecamente no
lineales

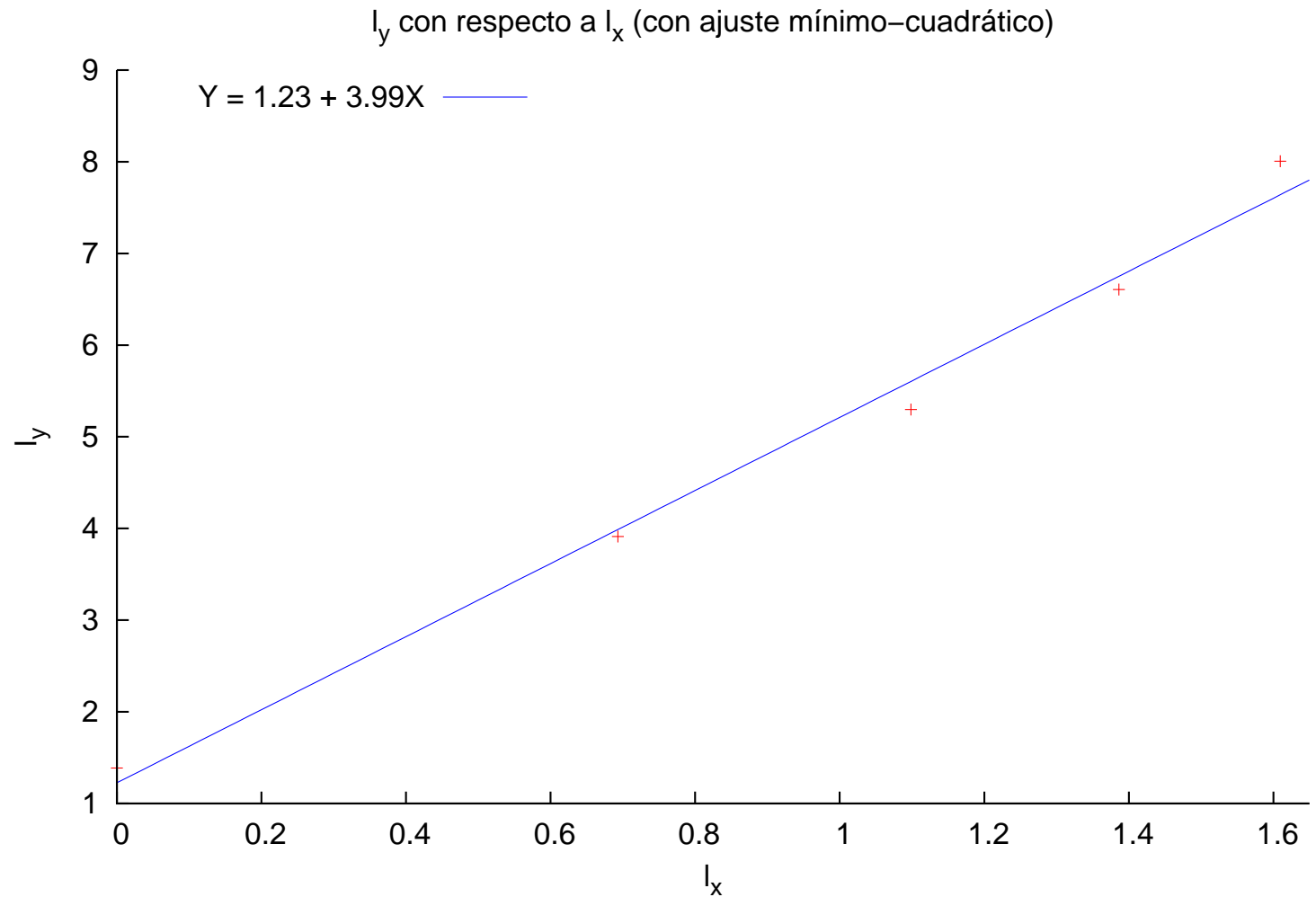


Figura 7: Representación gráfica de los datos transformados.

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Modelos intrínsecamente no lineales

Modelos intrínsecamente no lineales

Contenidos

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Aproximación lineal de Taylor

Mínimos cuadrados no lineales

Estimación por máxima verosimilitud

Algoritmos de búsqueda:
Newton–Raphson y Gauss–Newton

Contraste de restricciones sobre los parámetros

Dentro de este apartado, estudiaremos explícitamente las situaciones en las cuales no es posible transformar el modelo de manera que pueda estimarse con las técnicas de estimación correspondientes al modelo lineal general.

Aun así, en primer lugar, se aborda la aproximación lineal de Taylor que consiste en obtener una versión lineal aproximada del modelo. Es decir, se recupera la idea de aplicar las técnicas conocidas aunque el modelo no sea linealizabile de forma natural y sencilla.

En segundo lugar, asumiendo que la naturaleza de las relaciones son no lineales, se aborda el uso de los métodos de mínimos cuadrados y máxima verosimilitud en este tipo de modelos. Recordemos, que la idea de partida en ambos casos no exige en ningún momento la linealidad del modelo. Sin embargo, como se verá, la resolución analítica del mismo se complica bastante cuando el modelo no es lineal, siendo necesario recurrir a métodos distintos a los usados hasta el momento para obtener las estimaciones buscadas.

Aproximación lineal de Taylor

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Supongamos que pretendemos estimar una relación entre una magnitud y y un conjunto de variables explicativas recogidas en el vector \mathbf{X}_t . Representaremos tal relación mediante el modelo econométrico:

$$y_t = f(\mathbf{X}_t, \beta) + \epsilon_t, \quad t = 1, 2, \dots, T,$$

donde f representa una función no lineal cualquiera. Podemos tomar una aproximación lineal de la función en un entorno de un punto cualquiera, pudiendo elegir inicialmente el valor $\hat{\beta}_0$. El modelo quedaría:

$$y_t = f(\mathbf{X}_t, \hat{\beta}_0) + \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0) + \epsilon_t, \quad t = 1, 2, \dots, T.$$

Operando:

$$y_t - f(\mathbf{X}_t, \hat{\beta}_0) + \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \hat{\beta}_0 = \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \beta + \epsilon_t,$$

con $t = 1, 2, \dots, T$.

Si denotamos:

$$y_t^* = y_t - f(\mathbf{X}_t, \hat{\beta}_0) + \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \hat{\beta}_0,$$

el modelo puede reescribirse como

$$y_t^* = \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \beta + \epsilon_t, \quad t = 1, 2, \dots, T.$$

El modelo así transformado es un modelo lineal general donde la nueva variable dependiente es y^* y el vector de variables explicativas es $\left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0}$.

Aplicando Mínimos Cuadrados Ordinarios:

$$\hat{\beta} = \left[\left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_0} \right]^{-1} \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}_0} \mathbf{y}^*.$$

Esta estimación dará buenos resultados si los valores iniciales $\hat{\beta}_0$ son próximos a los verdaderos valores, lo cual no es conocido.

Aproximación lineal de Taylor: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Dado el modelo $y_t = \beta_1 e^{\beta_2 x_t} + \epsilon_t$ vamos a obtener una aproximación lineal del mismo.

Puesto que $f(\mathbf{X}_t, \beta) = \beta_1 e^{\beta_2 x_t}$ depende de β_1 y β_2 en este caso se verifica que:

$$\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} = \left(\frac{\partial \beta_1 e^{\beta_2 x_t}}{\partial \beta_1} \quad \frac{\partial \beta_1 e^{\beta_2 x_t}}{\partial \beta_2} \right) = (e^{\beta_2 x_t} \quad \beta_1 x_t e^{\beta_2 x_t}).$$

Por tanto, dado un valor inicial $\hat{\beta}$, se tiene:

$$\begin{aligned} y_t &= f(\mathbf{X}_t, \hat{\beta}) + \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)'_{\beta=\hat{\beta}} (\beta - \hat{\beta}) + \epsilon_t \\ &= \hat{\beta}_1 e^{\hat{\beta}_2 x_t} + \left(e^{\hat{\beta}_2 x_t} \quad \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t} \right) \cdot \begin{pmatrix} \beta_1 - \hat{\beta}_1 \\ \beta_2 - \hat{\beta}_2 \end{pmatrix} + \epsilon_t \\ &= \hat{\beta}_1 e^{\hat{\beta}_2 x_t} + e^{\hat{\beta}_2 x_t} \cdot (\beta_1 - \hat{\beta}_1) + \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t} \cdot (\beta_2 - \hat{\beta}_2) + \epsilon_t. \end{aligned}$$

Aproximación lineal de Taylor: ejemplo

Contenidos

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Aproximación lineal de Taylor

Mínimos cuadrados no lineales

Estimación por máxima verosimilitud

Algoritmos de búsqueda:

Newton–Raphson y Gauss–Newton

Contraste de restricciones sobre los parámetros

Operando de forma conveniente:

$$y_t - \hat{\beta}_1 e^{\hat{\beta}_2 x_t} + e^{\hat{\beta}_2 x_t} \cdot \hat{\beta}_1 + \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t} \cdot \hat{\beta}_2 = e^{\hat{\beta}_2 x_t} \cdot \beta_1 + \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t} \cdot \beta_2 + \epsilon_t.$$

Como se puede observar esta expresión es intrínsecamente lineal, es decir, haciendo un cambio de variable puede estimarse por MCO.

En efecto, llamando:

$$\begin{aligned} y_t^* &= y_t + \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t} \cdot \hat{\beta}_2, \\ x_{1t}^* &= e^{\hat{\beta}_2 x_t}, \\ x_{2t}^* &= \hat{\beta}_1 x_t e^{\hat{\beta}_2 x_t}, \end{aligned}$$

se obtendría la aproximación lineal $y_t^* = x_{1t}^* \beta_1 + x_{2t}^* \beta_2 + \epsilon_t$.

En tal caso, para el valor inicial $\hat{\beta}_1 = 1 = \hat{\beta}_2$ se obtendría el modelo linealizado:

$$y_t^* = \beta_1 x_{1t}^* + \beta_2 x_{2t}^* + \epsilon_t,$$

donde $y_t^* = y_t + x_t e^{x_t}$, $x_{1t}^* = e^{x_t}$ y $x_{2t}^* = x_t e^{x_t}$.

Aproximación lineal de Taylor: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Así, por ejemplo, a partir de los siguientes datos:

y	x	y^*	x_{1t}^*	x_{2t}^*
0.3	2.2	20.15503	9.025013	19.85503
1.1	3.8	170.96450	44.701184	169.8645
3.4	5	745.46580	148.413159	742.0658
10	5.1	846.51173	164.021907	836.51173

Se obtienen las siguientes estimaciones por MCO para el modelo original y linealizado por Taylor:

$$\hat{y} = -6.033 + 2'418x_t, \quad R^2 = 0'5523$$

$$\hat{y}^* = -0.08252x_{1t}^* + 1'02501x_{2t}^*, \quad R^2 = 1$$

Aproximación lineal de Taylor: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

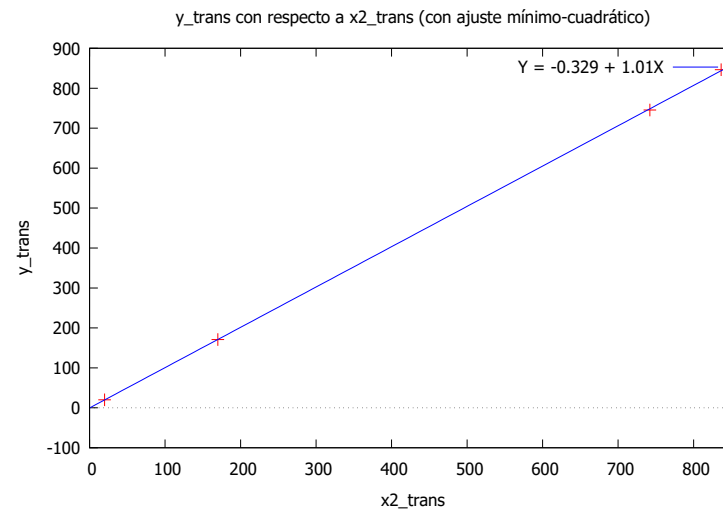
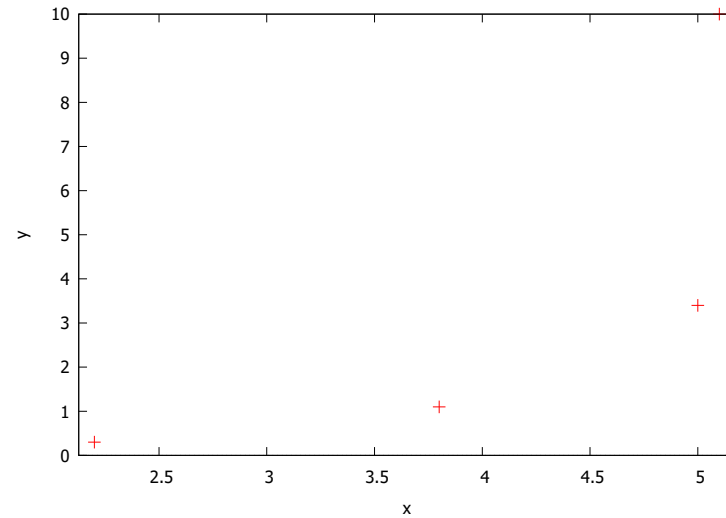


Figura 8: Representación gráfica de los originales y transformados.

Mínimos cuadrados no lineales

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Supongamos que pretendemos estimar un modelo cuya especificación genérica es:

$$y_t = f(\mathbf{X}_t, \beta) + \epsilon_t, \quad t = 1, 2, \dots, T,$$

donde \mathbf{X}_t es el vector de variables independientes, β es el vector de parámetros del modelo a estimar y f es una función no lineal de las componentes de los vectores \mathbf{X}_t y β , y cuya primera derivada vamos a suponer que es no lineal en β . El método de mínimos cuadrados no lineales, al igual que su homólogo lineal, trata de minimizar la suma de los residuos al cuadrado, es decir, minimizar la siguiente expresión:

$$SRC(\beta) = \sum_{t=1}^T e_t^2 = \sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta))^2.$$

Derivando la expresión anterior obtenemos las condiciones de primer y segundo orden necesarias y suficientes para la obtención del mínimo.

Así, derivando una primera vez se obtiene:

$$\frac{\partial SRC(\beta)}{\partial \beta} = -2 \sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta)) \cdot \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta}, \quad \forall t$$

e igualando a cero dicha derivada parcial se obtienen las ecuaciones normales del modelo:

$$\sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta)) \cdot \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} = 0, \quad \forall t.$$

Mínimos cuadrados no lineales: ejemplo

Contenidos

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Aproximación lineal de Taylor

Mínimos cuadrados no lineales

Estimación por máxima verosimilitud

Algoritmos de búsqueda:

Newton–Raphson y Gauss–Newton

Contraste de restricciones sobre los parámetros

Obtener el sistema de ecuaciones normales del modelo $y_t = \beta_1 + x_t^{\beta_2} + \epsilon_t$.

El objetivo es minimizar la suma de cuadrados de los residuos, esto es,

$$SRC(\beta) = \sum_{t=1}^T \left(y_t - \beta_1 - x_t^{\beta_2} \right)^2, \text{ donde } \beta' = (\beta_1 \ \beta_2).$$

Por tanto, habrá que derivar la expresión anterior (usando que si $f(x) = a^x$ entonces $f'(x) = a^x \ln a$) con respecto a cada uno de los elementos de β :

$$\frac{\partial SRC(\beta)}{\partial \beta_1} = -2 \cdot \sum_{t=1}^T \left(y_t - \beta_1 - x_t^{\beta_2} \right),$$

$$\frac{\partial SRC(\beta)}{\partial \beta_2} = -2 \cdot \sum_{t=1}^T \left(y_t - \beta_1 - x_t^{\beta_2} \right) \cdot x_t^{\beta_2} \cdot \ln x_t.$$

Igualando a cero estas derivadas se obtendrá el sistema de ecuaciones normales:

$$\sum_{t=1}^T \left(y_t - \beta_1 - x_t^{\beta_2} \right) = 0, \quad \sum_{t=1}^T \left(y_t - \beta_1 - x_t^{\beta_2} \right) \cdot x_t^{\beta_2} \cdot \ln x_t = 0.$$

Estimación por máxima verosimilitud

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Sabemos que la función de verosimilitud del modelo

$$y_t = f(\mathbf{X}_t, \beta) + \epsilon_t, \quad \epsilon_t \sim N(\mathbf{0}, \sigma^2 I_T),$$

viene dada por la expresión:

$$L(\beta, \sigma^2 | \mathbf{y}, \mathbf{X}) = \frac{1}{(2\pi\sigma^2)^{T/2}} \exp \left\{ -\frac{1}{(2\sigma^2)} (\mathbf{y} - f(\mathbf{X}_t, \beta))' (\mathbf{y} - f(\mathbf{X}_t, \beta)) \right\},$$

y aplicando logaritmos neperianos:

$$\ln L(\beta, \sigma^2 | \mathbf{y}, \mathbf{X}) = -\frac{T}{2} \ln(2\pi) - \frac{T}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} SCR(\beta).$$

Los posibles estimadores máximo verosímiles serán obtenidos tras derivar la expresión anterior con respecto a β y σ^2 e igualando el resultado a cero.

En el primer caso, se obtiene la derivada:

$$\frac{\partial \ln L(\beta, \sigma^2 | \mathbf{y}, \mathbf{X})}{\partial \beta} = \frac{\partial}{\partial \beta} \frac{-1}{2\sigma^2} SCR(\beta) = \frac{1}{\sigma^2} \sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta)) \cdot \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta},$$

Estimación por máxima verosimilitud

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

que igualada a cero queda:

$$\sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta)) \cdot \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} = 0, \quad \forall t.$$

Adviértase que los resultados obtenidos coinciden con el estimador por mínimos cuadrados no lineales, por tanto, al igual que antes, no es posible dar una solución analítica para las soluciones de este sistema de ecuaciones (y una vez más se hace necesario un método iterativo para obtener los valores de los parámetros).

Sin embargo, sí que es posible dar una expresión para la varianza de las perturbaciones. Derivando parcialmente con respecto a σ^2 e igualando a 0, obtenemos:

$$-\frac{T}{2} \frac{1}{\hat{\sigma}^2} + \frac{1}{2(\hat{\sigma}^2)^2} SCR(\beta) = 0.$$

$$\text{Despejando: } \hat{\sigma}^2 = \frac{SCR(\beta)}{T}.$$

En la presente sección veremos los procedimientos numéricos que serán utilizados para resolver las ecuaciones normales obtenidas que no pueden ser resueltas de forma directa mediante procedimientos algebraicos.

Dado el modelo:

$$y_t = f(\mathbf{X}_t, \beta) + \epsilon_t,$$

la suma de cuadrados de los residuos vendrá dada por la expresión:

$$SCR(\beta) = \sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta))^2,$$

siendo la condición necesaria de mínimo para esta función la dada por la ecuación normal:

$$\sum_{t=1}^T (y_t - f(\mathbf{X}_t, \beta)) \cdot \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} = 0.$$

Veremos a continuación dos métodos para resolver este tipo de ecuaciones.

Algoritmo de Newton-Raphson

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton-Raphson y
Gauss-Newton

Contraste de
restricciones sobre los
parámetros

Este algoritmo se basa en minimizar la suma de cuadrados de los residuos. En primer lugar se toma como aproximación a dicha suma el desarrollo del polinomio de Taylor de segundo orden en un entorno del valor inicial $\hat{\beta}_0$:

$$SCR(\beta) \simeq SCR(\hat{\beta}_0) + \left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0) + \frac{1}{2} (\beta - \hat{\beta}_0)' \left(\frac{\partial^2 SCR(\beta)}{\partial \beta \partial \beta'} \right)_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0),$$

Derivando respecto β :

$$\frac{\partial SCR(\beta)}{\partial \beta} \simeq \left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_0} + \left(\frac{\partial^2 SCR(\beta)}{\partial \beta \partial \beta'} \right)_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0).$$

Igualando a cero la primera derivada:

$$\left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_0} + \left(\frac{\partial^2 SCR(\beta)}{\partial \beta \partial \beta'} \right)_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0) = 0,$$

Algoritmo de Newton-Raphson

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

y despejando β :

$$\beta = \hat{\beta}_0 - \left(\frac{\partial^2 SCR(\beta)}{\partial \beta \partial \beta'} \right)_{\beta = \hat{\beta}_0}^{-1} \left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta = \hat{\beta}_0}.$$

Siempre que exista dicha inversa, a partir de la expresión anterior se puede plantear el siguiente procedimiento iterativo:

$$\hat{\beta}_{n+1} = \hat{\beta}_n - \left(\frac{\partial^2 SCR(\beta)}{\partial \beta \partial \beta'} \right)_{\beta = \hat{\beta}_n}^{-1} \left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta = \hat{\beta}_n}.$$

Este procedimiento se repite hasta que converja, es decir, hasta que exista h tal que $\hat{\beta}_{h+1} = \hat{\beta}_h$. En tal caso se tendría entonces que

$$\left(\frac{\partial SCR(\beta)}{\partial \beta} \right)_{\beta = \hat{\beta}_h} = 0,$$

de donde se deduce que $\hat{\beta}_h$ minimiza a $SCR(\beta)$.

Algoritmo de Newton-Raphson: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton-Raphson y
Gauss-Newton

Contraste de
restricciones sobre los
parámetros

Dado el modelo $y_t = x_t^\beta + \epsilon_t$, $t = 1, \dots, T$, obtener la estimación iterativa proporcionada por el algoritmo de Newton-Raphson.

En este caso $SCR(\beta) = \sum_{t=1}^T (y_t - x_t^\beta)^2$, por lo que:

$$\frac{\partial SCR(\beta)}{\partial \beta} = -2 \sum_{t=1}^T (y_t - x_t^\beta) x_t^\beta \ln x_t,$$

$$\begin{aligned} \frac{\partial^2 SCR(\beta)}{\partial \beta^2} &= \frac{\partial}{\partial \beta} \left(-2 \sum_{t=1}^T (y_t - x_t^\beta) x_t^\beta \ln x_t \right) \\ &= \frac{\partial}{\partial \beta} \left(-2 \sum_{t=1}^T y_t x_t^\beta \ln x_t + 2 \sum_{t=1}^T x_t^{2\beta} \ln x_t \right) \\ &= -2 \sum_{t=1}^T y_t x_t^\beta (\ln x_t)^2 + 4 \sum_{t=1}^T x_t^{2\beta} (\ln x_t)^2. \end{aligned}$$

Donde se ha usado que $\frac{\partial}{\partial \beta} x_t^\beta = x_t^\beta \ln x_t$.

Algoritmo de Newton-Raphson: Ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton-Raphson y
Gauss-Newton

Contraste de
restricciones sobre los
parámetros

Por tanto:

$$\beta_{n+1} = \beta_n + \frac{\sum_{t=1}^T y_t x_t^{\beta_n} \ln x_t - \sum_{t=1}^T x_t^{2\beta_n} \ln x_t}{\sum_{t=1}^T y_t x_t^{\beta_n} (\ln x_t)^2 - 2 \sum_{t=1}^T x_t^{2\beta_n} (\ln x_t)^2}.$$

Para el valor inicial $\beta_0 = 0$, la primera iteración corresponde a:

$$\begin{aligned} \beta_1 &= \frac{\sum_{t=1}^T y_t \ln x_t - \sum_{t=1}^T \ln x_t}{\sum_{t=1}^T y_t (\ln x_t)^2 - 2 \sum_{t=1}^T (\ln x_t)^2} \\ &= \frac{\sum_{t=1}^T (y_t - 1) \ln x_t}{\sum_{t=1}^T (y_t - 2) (\ln x_t)^2}. \end{aligned}$$

Algoritmo de Gauss-Newton

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Partiendo de la aproximación lineal del desarrollo en serie de Taylor para la función no lineal f en un entorno del punto $\hat{\beta}_0$ se obtiene:

$$f(\mathbf{X}_t, \beta) = f(\mathbf{X}_t, \hat{\beta}_0) + \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_0} (\beta - \hat{\beta}_0).$$

La expresión iterativa correspondiente al algoritmo de Gauss-Newton es:

$$\hat{\beta}_{n+1} = \hat{\beta}_n + \left(\sum_{t=1}^T \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right) \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)' \right)_{\beta=\hat{\beta}_n}^{-1} \sum_{t=1}^T \left(\frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} \right)_{\beta=\hat{\beta}_n} \hat{\epsilon}_t,$$

donde $\hat{\epsilon}_t = y_t - f(\mathbf{X}_t, \hat{\beta}_n)$ es el residuo obtenido en la estimación realizada. Esta expresión es similar al de algoritmo de Newton-Raphson, con la ventaja añadida que la matriz a invertir es simétrica y definida positiva.

Algoritmo de Gauss-Newton: ejemplo

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:
Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Dado el modelo no lineal $y_t = x_t^\beta + \epsilon_t$ se tiene que:

$$f(\mathbf{X}_t, \beta) = x_t^\beta \Rightarrow \frac{\partial f(\mathbf{X}_t, \beta)}{\partial \beta} = x_t^\beta \ln x_t.$$

En tal caso es claro que:

$$\hat{\beta}_{n+1} = \hat{\beta}_n + \frac{\sum_{t=1}^T x_t^{\beta_n} \ln x_t (y_t - x_t^{\beta_n})}{\sum_{t=1}^T x_t^{2\beta_n} (\ln x_t)^2}.$$

Para el valor inicial $\hat{\beta}_0 = 0$, la primera iteración del algoritmo de Gauss–Newton corresponde a:

$$\hat{\beta}_1 = \frac{\sum_{t=1}^T (y_t - 1) \ln x_t}{\sum_{t=1}^T (\ln x_t)^2}.$$

Criterios de finalización del algoritmo

Contenidos

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Aproximación lineal de Taylor

Mínimos cuadrados no lineales

Estimación por máxima verosimilitud

Algoritmos de búsqueda: Newton–Raphson y Gauss–Newton

Contraste de restricciones sobre los parámetros

Estos métodos iterativos no siempre son convergentes. Por ello, se requiere establecer a priori unos criterios que permitan finalizar el proceso iterativo cuando no se alcanza dicha convergencia. Estos criterios son:

1. Los valores de los parámetros se estabilizan.
2. El valor de la función objetivo se estabiliza.
3. El vector gradiente está próximo a cero.
4. Se alcanzó el número máximo de iteraciones.
5. Se alcanzó el límite máximo de tiempo de cálculo.

Contraste de restricciones sobre los parámetros

Contenidos

Introducción

Modelos
intrínsecamente lineales

Modelos
intrínsecamente no
lineales

Aproximación lineal de
Taylor

Mínimos cuadrados no
lineales

Estimación por máxima
verosimilitud

Algoritmos de
búsqueda:

Newton–Raphson y
Gauss–Newton

Contraste de
restricciones sobre los
parámetros

Supongamos que queremos plantear cualquier hipótesis que implique una combinación de los parámetros del modelo de regresión. El sistema de restricciones se puede expresar como:

$$R_{q \times k} \beta = \mathbf{r}_{q \times 1}, \quad q \leq k.$$

El modelo sobre el que se imponen las restricciones se denomina “modelo restringido” y el modelo sobre el que no se imponen las restricciones se llama “modelo sin restricciones”.

■ Contraste F:

$$F = \frac{(\mathbf{e}'_r \mathbf{e}_r - \mathbf{e}' \mathbf{e})/q}{\mathbf{e}' \mathbf{e}/(n - k)} \sim F_{q, n-k},$$

donde $\mathbf{e}'_r \mathbf{e}_r$ es la SCR del modelo restringido; $\mathbf{e}' \mathbf{e}$ es la SCR del modelo sin restringir y q es el número de restricciones.

■ Contraste de razón de verosimilitudes:

$$-2(\ln L_R - \ln L) = n \ln \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_R^2} \right) \sim \chi_q^2,$$

Contraste de restricciones sobre los parámetros

Contenidos

Introducción

Modelos intrínsecamente lineales

Modelos intrínsecamente no lineales

Aproximación lineal de Taylor

Mínimos cuadrados no lineales

Estimación por máxima verosimilitud

Algoritmos de búsqueda:

Newton–Raphson y Gauss–Newton

Contraste de restricciones sobre los parámetros

■ Contraste de Wald:

$$W = (R\beta - \mathbf{r})' \left[\left(\frac{\partial R(\hat{\beta})}{\partial \hat{\beta}} \right) \hat{\sigma}^2 (X'X)^{-1} \left(\frac{\partial R(\hat{\beta})}{\partial \hat{\beta}} \right)' \right]^{-1} (R\beta - \mathbf{r}) \sim \chi_q^2$$

Esta expresión no requiere estimar el modelo con y sin restricciones.