T5. Ampliación del modelo lineal básico

Ana J. López y Rigoberto Pérez

Dpto Economía Aplicada. Universidad de Oviedo

Curso 2010-2011

Índice

Ampliación del modelo lineal básico

Competencias

Los contenidos de este tema suponen una ampliación de los desarrollados en el tema 3 y proporcionan una visión más realista de las variables económicas cuyo comportamiento se pretende modelizar. El objetivo es que los estudiantes lleguen a ser capaces de:

- Identificar los principales problemas asociados al incumplimiento de las hipótesis habituales y tratar de resolverlos
- Interpretar adecuadamente los resultados de los contrastes de homocedasticidad, no autocorrelación y normalidad
- Detectar la presencia de multicolinealidad y sus posibles consecuencias
- Contrastar la existencia de cambios estructurales y estudiar posibles soluciones

Ampliación del modelo lineal básico: $\mathbf{y} = \mathbf{X}\beta + \mathbf{u}$

Problemas de Especificación	Omisión de variables relevantes, Inclusión de variables irrelevantes, Forma funcional errónea
Violación de hipótesis sobre $\mathbf{u}$	Esperanza no nula, Matriz de Var-Cov no escalar (heterocedasticidad, autocorrelación, no normalidad)
Matriz de regresores $\mathbf{X}$	Regresores estocásticos, Multicolinealidad
Parámetros β	Cambio estructural

Errores de especificación

Sobre las variables

- Omisión de variables relevantes
- Inclusión de variables irrelevantes

Sobre la forma funcional

- Contrastes sobre la linealidad
  - Mínimos cuadrados no lineales

Errores de especificación

Omisión de variables relevantes

- Modelo correcto: $\mathbf{y} = \mathbf{X_1}\boldsymbol{\beta_1} + \mathbf{X_2}\boldsymbol{\beta_2} + \mathbf{u}$
- Modelo propuesto: $\mathbf{y} = \mathbf{X_1}\boldsymbol{\beta_1} + \mathbf{v}$

Los estimadores MCO son sesgados

$$\hat{\beta}_1 = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{y}$$

$$E\left(\hat{\beta}_1\right) = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{E}(\mathbf{y}) = (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'(\mathbf{X}_1\boldsymbol{\beta_1} + \mathbf{X}_2\boldsymbol{\beta_2}) =$$
$$= \beta_1 + (\mathbf{X}_1'\mathbf{X}_1)^{-1}\mathbf{X}_1'\mathbf{X}_2\beta_2$$

Los estimadores serían insesgados si las variables excluídas no están correlacionadas con las incluídas

Errores de especificación

Omisión de variables relevantes

- Modelo correcto: $\mathbf{y} = \mathbf{X_1}\beta_1 + \mathbf{X_2}\beta_2 + \mathbf{u}$
- Modelo propuesto: $\mathbf{y} = \mathbf{X_1}\beta_1 + \mathbf{v}$

- Los estimadores MCO son inconsistentes
- Los residuos son superiores a los del modelo correcto (esto afecta a la capacidad explicativa del modelo y también a la matriz var-cov).

$$S^2(\hat{\beta}) = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}(\mathbf{X}'\mathbf{X})^{-1}$$

- Los tests de significación t student son excesivamente exigentes

Errores de especificación

Inclusión de variables irrelevantes

- Las consecuencias de incluir en el modelo variables irrelevantes son **menos graves que las de omitir variables relevantes**
- Los Estimadores MCO son **insesgados y consistentes**
- Los Estimadores MCO **dejan de ser eficientes**, y por lo tanto las estimaciones pierden precisión
- Disminuyen los grados de libertad como consecuencia del aumento de k
- Al incorporar nuevas variables podría aparecer multicolinealidad

Contrastes de especificación

Omitir variables

H_0: Las r variables seleccionadas pueden ser omitidas (sus coeficientes son nulos)

Añadir variables

H_0: Las r variables seleccionadas no deben ser añadidas (sus coeficientes son nulos)

$$\left(\frac{\hat{\mathbf{u}}'_R\hat{\mathbf{u}}_R - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right)\frac{n-k}{r} \approx F^r_{n-k}$$

$$LR = -2\left(\ln L_R - \ln L\right) = n\ln\left(\frac{\hat{\mathbf{u}}'_R\hat{\mathbf{u}}_R}{\hat{\mathbf{u}}'\hat{\mathbf{u}}}\right) \rightarrow \chi^2_r$$

Test F de restricciones lineales y test LR de razón de verosimilitudes

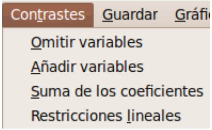
Para p bajos se rechaza la nulidad de los coeficientes y se concluye que las variables deben estar en el modelo

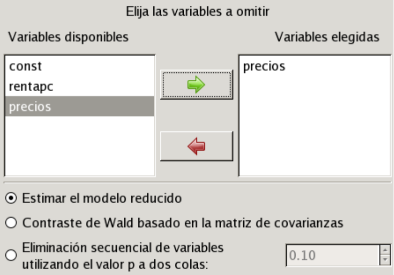
Test para omitir variables

Modelo 1: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35170	0.348359	−3.8802	0.0010
rentapc	−9.35994e–05	5.13489e–05	−1.8228	0.0841
precios	0.0498576	0.00860015	5.7973	0.0000

R^2	0.854650	R^2 corregido	0.839351
Log-verosimilitud	−9.230929	Criterio de Akaike	24.46186
Criterio de Schwarz	27.73499	Hannan--Quinn	25.23291





Test de variables omitidas

Modelo 2: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

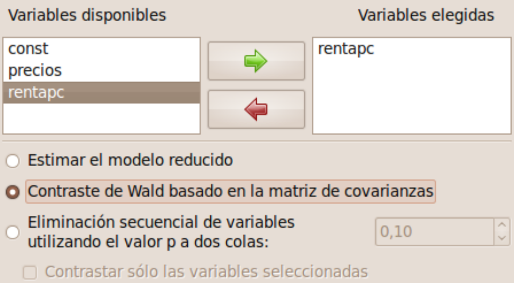
	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−0.00368661	0.420705	−0.0088	0.9931
rentapc	0.000179742	3.29843e−05	5.4493	0.0000

Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	8.254703	D.T. de la regresión	0.642445
R^2	0.597545	R^2 corregido	0.577422
$F(1,20)$	29.69498	Valor p (de F)	0.000025
Log-verosimilitud	−20.43379	Criterio de Akaike	44.86759
Criterio de Schwarz	47.04967	Hannan--Quinn	45.38162
$\hat{\rho}$	0.831284	Durbin--Watson	0.655456

Comparación entre el modelo 1 y el modelo 2:

Hipótesis nula: el parámetro de regresión es cero para precios
Estadístico de contraste: F(1, 19) = 33.6087, con valor p = 1.38425e-05
De los 3 estadísticos de selección de modelos, 0 han mejorado.

Test de Wald variables omitidas




```
Hipótesis nula: el parámetro de regresión es cero para rentapc

Estadístico de contraste asintótico:
chi-cuadrado de Wald(1) = 3.32265, con valor p = 0.0683315
forma F: F(1, 19) = 3.32265, con valor p = 0.0841078
```

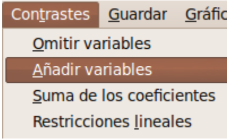
Estadístico de Wald(r=N.restricciones): $\boxed{W = n\dfrac{\hat{\mathbf{u}}_R'\hat{\mathbf{u}}_R - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} \approx \chi^2_r}$

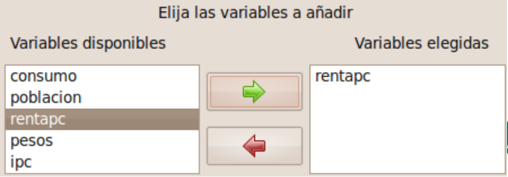
Test para añadir variables

Modelo 3: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−0.00368661	0.420705	−0.0088	0.9931
rentapc	0.000179742	3.29843e−05	5.4493	0.0000

Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	8.254703	D.T. de la regresión	0.642445
R^2	0.597545	R^2 corregido	0.577422
$F(1,20)$	29.69498	Valor p (de F)	0.000025
Log-verosimilitud	−20.43379	Criterio de Akaike	44.86759
Criterio de Schwarz	47.04967	Hannan--Quinn	45.38162
$\hat{\rho}$	0.831284	Durbin--Watson	0.655456





Test para añadir variables

Modelo 4: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35170	0.348359	−3.8802	0.0010
rentapc	−9.35994e−05	5.13489e−05	−1.8228	0.0841
precios	0.0498576	0.00860015	5.7973	0.0000

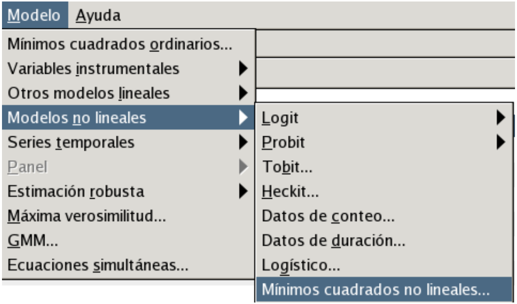
Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	2.981244	D.T. de la regresión	0.396116
R^2	0.854650	R^2 corregido	0.839351
$F(2,19)$	55.85969	Valor p (de F)	1.10e−08
Log-verosimilitud	−9.230929	Criterio de Akaike	24.46186
Criterio de Schwarz	27.73499	Hannan--Quinn	25.23291
$\hat{\rho}$	0.264285	Durbin--Watson	1.401145

Comparación entre el modelo 3 y el modelo 4:

Hipótesis nula: el parámetro de regresión es cero para precios
Estadístico de contraste: F(1, 19) = 33.6087, con valor p = 1.38425e-05
De los 3 estadísticos de selección de modelos, 3 han mejorado.

Forma funcional errónea

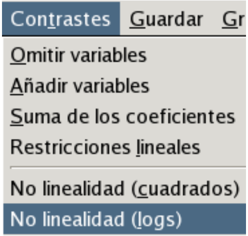
- Podemos tener **sospechas de no linealidad**, ligadas a la teoría económica o a la información muestral.
- La especificación errónea de modelos lineales tiene **consecuencias** similares a la omisión de variables relevantes.
- ¿Cómo tratar la no linealidad?



- En caso de no linealidad, la estimación de estos modelos se lleva a cabo por **Mínimos Cuadrados no lineales** (MCNL), que utilizan métodos de aproximación numérica (Gauss-Newton) y coinciden con los EMV.

Contrastes de linealidad

Se contrasta la hipótesis de linealidad, completando la especificación del modelo propuesto con los logaritmos o los cuadrados de las variables explicativas y contrastando la nulidad de sus coeficientes



Regresión auxiliar para el contraste de no linealidad (términos logarítmicos)\MCO, usando observaciones 1987-2008 (T=22)
Variable dependiente **uhat**

	Coeficiente	Desv.Típica	Estadístico t	Valor p	
const	30.8945	12.6802	2.436	0.0261	**
rentapc	-0.000227443	0.000140966	-1.613	0.1251	
precios	0.165767	0.0360676	4.596	0.0003	***
l_rentapc	5.21124	1.60434	3.248	0.0047	***
l_precios	-20.4815	4.39776	-4.657	0.0002	***

R-cuadrado = 0.561849, Estadístico de contraste: $TR^2 = 12{,}3607$, con valor $p = P(\chi^2_2 > 12{,}3607) = 0{,}00206974$

Contrastes de linealidad

Regresión auxiliar para el contraste de no linealidad (términos al cuadrado)\MCO, usando las observaciones 1987-2008 (T = 22)\Variable dependiente: uhat

	Coeficiente	Desv.Típica	Estadístico t	Valor p	
const	7.87416	1.56991	5.016	0.0001	$\star\star\star$
rentapc	0.000667027	0.000138571	4.814	0.0002	$\star\star\star$
precios	−0.231702	0.0386627	−5.993	1.45e-05	$\star\star\star$
sq_rentapc	−1.82857e-08	5.39076e-09	−3.392	0.0035	$\star\star\star$
sq_precios	0.000929836	0.000150836	6.165	1.04e-05	$\star\star\star$

R-cuadrado = 0.692012

Estadístico de contraste: $TR^2 = 15{,}2243$,

con valor $p = P(\chi^2_2 > 15{,}2243) = 0{,}000494415$

Bajo la hipótesis nula de linealidad los coeficientes de los términos cuadráticos son nulos. Para p bajos se rechaza la nulidad de estos coeficientes y se concluye que existe evidencia de no linealidad

Test RESET de Ramsey

Detecta errores de especificación en sentido amplio (variables omitidas, forma funcional incorrecta, ...)

Modelo estimado: $\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_k X_k$

$$\hat{Y} = \beta_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \delta_1 \hat{Y}^2 + \delta_2 \hat{Y}^3 + u$$

Hipótesis de especificación correcta: $\boxed{H_0 : \delta_1 = 0, \delta_2 = 0}$

Discrepancias:

$$\left(\frac{\hat{\mathbf{u}}'_R \hat{\mathbf{u}}_R - \hat{\mathbf{u}}'\hat{\mathbf{u}}}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} \right) \frac{n-k-2}{k+2} \approx F^{k+2}_{n-k-2} \ ; \ LR = n \ln \left(\frac{\hat{\mathbf{u}}'_R \hat{\mathbf{u}}_R}{\hat{\mathbf{u}}'\hat{\mathbf{u}}} \right) \rightarrow \chi^2_r$$

Si los tests no conducen al rechazo de la hipótesis, se asume como válida la especificación propuesta. En cambio, si p es bajo deberíamos rechazar el supuesto de especificación correcta

Test RESET de Ramsey

Regresión auxiliar para el contraste de especificación **RESET**\MCO, usando las observaciones 1987–2008 (T = 22)\Variable dependiente: consumopc\

	Coeficiente	Desv.Típica	Estadístico t	Valor p	
const	−0.116416	3.41005	−0.03414	0.9732	
rentapc	5.26458e−05	0.000161027	0.3269	0.7477	
precios	0.0293300	0.0781971	0.3751	0.7122	
$\hat{y}^2$	−0.532910	0.535877	−0.9945	0.3339	
$\hat{y}^3$	0.110908	0.0587397	1.888	0.0762	$\star$

Estadístico de contraste: F = 22.680726,

con valor p = P(F(2,17) > 22.6807) = 1.59e−05

Hipótesis básicas sobre **u**

$$E(\mathbf{u}) = \mathbf{0}$$
$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I}$$

$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \begin{pmatrix} E(u_1^2) & E(u_1u_2) & \cdots & E(u_1u_n) \\ E(u_2u_1) & E(u_2^2) & \cdots & E(u_2u_n) \\ \vdots & \vdots & \ddots & \vdots \\ E(u_nu_1) & E(u_nu_2) & \cdots & E(u_n^2) \end{pmatrix} = \begin{pmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{pmatrix}$$

$$\boxed{\begin{array}{lcr} Var(u_i) = E(u_i^2) = \sigma^2 & \Rightarrow & \text{Homocedasticidad} \\ Cov(u_i, u_j) = E(u_iu_j) = 0 & \Rightarrow & \text{No autocorrelación} \end{array}}$$

Hipótesis de normalidad

- **u** sigue una distribución normal multivariante: $\mathbf{u} \approx \mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$
- **y** sigue una distribución normal multivariante: $\mathbf{y} \approx \mathcal{N}(\mathbf{X}\beta, \sigma^2\mathbf{I})$

Perturbaciones de media no nula

La principal consecuencia de que las perturbaciones tengan esperanza no nula es la estimación sesgada del término independiente

$$\boxed{E(\mathbf{u}) \neq \mathbf{0}}$$

Esta hipótesis no puede ser contrastada porque la estimación MCO garantiza que los residuos tienen media nula

$$Y = \beta_1 + \beta_2 X_2 + \cdots + \beta_k X_k + u \Rightarrow E(Y) = \boxed{\beta_1 + E(u)} + \beta_2 X_2 + \cdots + \beta_k X_k$$

Problema relacionado con la especificación del modelo ¿Faltan variables explicativas?

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \boldsymbol{v}$$

$$\mathbf{y} = \mathbf{X}_1 \boldsymbol{\beta}_1 + \overbrace{\mathbf{X}_2 \boldsymbol{\beta}_2 + \mathbf{u}}$$

$$\boxed{E(\boldsymbol{v}) = \mathbf{X}_2 \boldsymbol{\beta}_2}$$

Matriz de varianzas-covarianzas no escalar

$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{I} \Rightarrow \begin{cases} E(u_i^2) = \sigma^2 \quad \forall i = 1, \ldots, n & \text{Homocedasticidad} \\ E(u_i u_j) = 0 \qquad \forall i \neq j & \text{No autocorrelación} \end{cases}$$

Matriz de var-cov no escalar: $Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2\boldsymbol{\Omega}$

Estimación por MCO

Estimador		$\hat{\beta}^{MCO} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$
Insesgados		$E\left(\hat{\beta}^{MCO}\right) = \beta$
Consistentes		$\lim_{n\to\infty} Var\left(\hat{\beta}^{MCO}\right) = 0$
No óptimos	La varianza no coincide con la mínima	$\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$

Estimador MCO con Var-Cov no escalar

- **Insesgado**

$$E\left(\hat{\beta}^{MCO}\right) = E\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}\right) = E\left((\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\beta + \mathbf{u})\right)$$
$$= \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'E(\mathbf{u}) = \beta$$

- **Estimadores MCO no óptimos**

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

Utilización de MCO con $E(\mathbf{uu}')$ no escalar

$$Var\left(\hat{\beta}^{MCO}\right) = \sigma^2 \left[(\mathbf{X}'\mathbf{X})^{-\mathbf{1}}\mathbf{X}' \boxed{\Omega} \mathbf{X}(\mathbf{X}'\mathbf{X})^{-\mathbf{1}})\right]$$

$$S^2 = \frac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{tr(\mathbf{M\Omega})}$$

$$M = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']$$

Simétrica: M=M'; Idempotente: MM=M; Traza: tr(M)=n-k

Doble sesgo

Modelización de MCO asumiendo $E(\mathbf{uu}')$ escalar

1. $Var\left(\hat{\beta}\right) = \sigma^2(X'X)^{-1}$, con lo que estamos identificando Ω, con la matriz identidad.
2. Estimar σ^2 como: $S^2 = \dfrac{\hat{\mathbf{u}}'\hat{\mathbf{u}}}{n-k}$

Estimación robusta de la matriz Var-Cov

La matriz de varianzas covarianzas puede estimarse de forma consistente o robusta mediante la expresión:

$$\hat{S}^2\left(\hat{\beta}^{MCO}\right) = S^2\left[(X'X)^{-1}X'\hat{\Omega}X(X'X)^{-1}\right] \quad ; \quad \hat{\Omega} = \begin{pmatrix} \hat{u}_1^2 & 0 & \cdots & 0 \\ 0 & \hat{u}_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{u}_n^2 \end{pmatrix}$$

Estima de forma consistente o robusta la matriz de varianzas- covarianzas, lo cual afecta a los contrastes de significación t de student

Estimación robusta de la matriz Var-Cov

Modelo 1: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35170	0.348359	−3.8802	0.0010
precios	0.0498576	0.00860015	5.7973	0.0000
rentapc	−9.35994e–05	5.13489e–05	−1.8228	0.0841

F(2, 19) 55.85969 Valor p (de F) 1.10e-08

Modelo 2: Desviaciones típicas HAC, con ancho de banda 2

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35170	0.464395	−2.9107	0.0090
precios	0.0498576	0.0116702	4.2722	0.0004
rentapc	−9.35994e–05	5.92050e–05	−1.5809	0.1304

F(2, 19) 31.17168 Valor p (de F) 1.00e-06

Mínimos Cuadrados Generalizados (MCG)

$$\boxed{\begin{gathered}\mathbf{y} = \mathbf{X}\beta + \mathbf{u} \\ Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \sigma^2\mathbf{\Omega}\end{gathered}}$$

Se transforma el modelo hasta tener matriz Var-Cov escalar

Se busca una matriz **P** cuadrada no singular tal que: $\mathbf{P}'\mathbf{P} = \mathbf{\Omega}^{-1}$

Modelo transformado: $\mathbf{P}\mathbf{y} = \mathbf{P}\mathbf{X}\beta + \mathbf{P}\mathbf{u}$

Características

$$E(\mathbf{P}\mathbf{u}) = \mathbf{P}E(\mathbf{u}) = \mathbf{0}$$

$$Cov(\mathbf{P}\mathbf{u}) = E(\mathbf{P}\mathbf{u}\mathbf{u}'\mathbf{P}') = \mathbf{P}E(\mathbf{u}\mathbf{u}')\mathbf{P}' = \mathbf{P}\sigma^2\mathbf{\Omega}\mathbf{P}' = \sigma^2\mathbf{P}\mathbf{P}^{-1}\mathbf{P}'^{-1}\mathbf{P}' = \sigma^2\mathbf{I}$$

Estimadores MCG

$$\hat{\beta}^{MCG} = [(PX)'(PX)]^{-1}(PX)'Py = \boxed{\left(X'\Omega^{-1}X\right)^{-1}X'\Omega^{-1}y}$$

Los estimadores de Mínimos Cuadrados Generalizados (MCG) coinciden con los EMV y son ELIO para el modelo transformado

Características de los estimadores MCG

Los estimadores MCG son insesgados

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

$$
\boxed{E\left(\hat{\beta}^{MCG}\right) = \beta}
$$

Los estimadores MCG son consistentes

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

Heterocedasticidad

$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \begin{pmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{pmatrix}$$

Causas

- Mala especificación
- Cambio estructural

Consecuencias

- Estimadores no óptimos
- Sesgo en las varianzas
- Test de significación no válidos

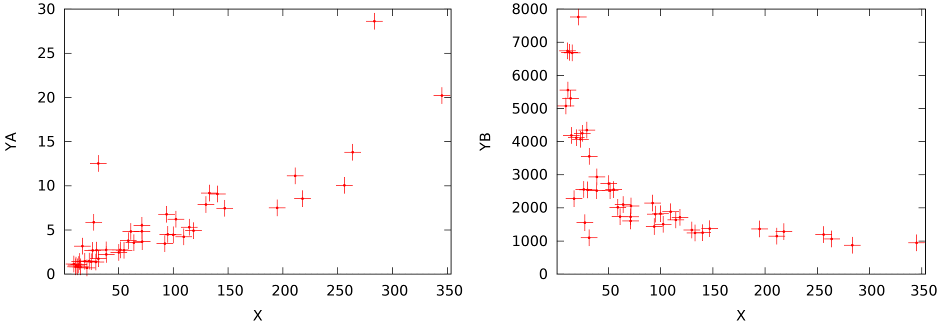
Detección

- Análisis gráfico
- Test de White, Breusch-Pagan, Goldfeld-Quandt, Glejser, ...

Soluciones

- Especificaciones alternativas
- Estimación del modelo por MCP

Gráficos heterocedasticidad



Contraste de White

$$\boxed{\begin{array}{ll} H_0 : \sigma_i^2 = \sigma^2 & \forall i = 1, \ldots, n \\ H_1 : \sigma_i^2 \neq \sigma_j^2 & \text{para algún } i \neq j \end{array}}$$

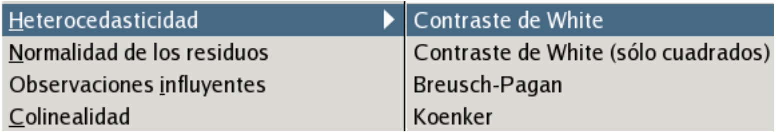
- Se estima el modelo propuesto, calculando sus residuos $\hat{\mathbf{u}}$
- Se efectúa una regresión de los residuos cuadráticos sobre las variables y sus cuadrados, incluyendo generalmente también productos cruzados:

$$\hat{u}^2 = \alpha_1 + \alpha_2 X_1 + \alpha_3 X_2 + \alpha_4 X_1^2 + \alpha_5 X_2^2 + \alpha_6 X_1 X_2 + v$$

- Para el coeficiente de determinación de esta regresión auxiliar (con m regresores) bajo H_0 se cumple: $\boxed{nR^2 \approx \chi^2_m}$

Si $p = P(\chi^2_m > nR^2)$ es bajo se rechaza la homocedasticidad y se concluye que las perturbaciones son heterocedásticas.

Contraste de White



Contraste de heterocedasticidad de White
MCO, usando las observaciones 1987–2008 (T = 22)
Variable dependiente: $\hat{u}^2$

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	3.92948	2.30350	1.706	0.1074	
precios	−0.0310550	0.0433067	−0.717	0.4837	
rentapc	−0.000359329	0.000177957	−2.019	0.0605	*
sq_precios	−0.00135966	0.000592252	−2.296	0.0355	**
X2_X3	2.11531e-05	9.13833e-06	2.315	0.0342	**
sq_rentapc	−5.96052e-08	2.88298e-08	−2.067	0.0553	*

R^2 = 0.501119

Estadístico de contraste: TR^2 = 11.024629,

con valor $p = P(\chi^2_5 > 11{,}024629) = 0{,}050894$

Contraste de Breusch-Pagan

$$\boxed{\begin{array}{ll} H_0 : \sigma_i^2 = \sigma^2 & \forall i = 1, \ldots, n \\ H_1 : \sigma_i^2 \neq \sigma_j^2 & \text{para algún } i \neq j \end{array}}$$

- Se estima el modelo propuesto, calculando sus residuos $\hat{\mathbf{u}}$
- Los residuos cuadraticos se normalizan con EMV de la varianza: $e_i^2 = \dfrac{\hat{u}_i^2}{\hat{\sigma}^2}$
- Se efectúa una regresión de estos residuos cuadráticos normalizados o escalados sobre las variables explicativas y se contrasta la nulidad de los coeficientes
- Este contraste se basa en la **Variación Explicada**, que bajo la hipótesis nula sigue una distribución **chi-cuadrado con k-1** grados de libertad

Se obtiene el nivel crítico, si es bajo se rechaza la homocedasticidad, concluyendo que las perturbaciones son heterocedásticas

Contraste de Breusch-Pagan

Contraste de heterocedasticidad de Breusch-Pagan
MCO, usando las observaciones 1987-2008 (T = 22)
Variable dependiente: $\hat{u}^2$ escalado

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	-3.16883	1.69508	-1.869	0.0771	*
precios	0.0350232	0.0418476	0.8369	0.4130	
rentapc	7.51488e-05	0.000249859	0.3008	0.7669	

Suma de cuadrados explicada = 29.6793

Estadístico de contraste: LM = 14.839664, $p = P(\chi^2_2 > 14{,}839664) = 0{,}000599$

Variable dependiente: $\hat{u}^2$ escalado (variante robusta de Koenker)

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	-0.564923	0.229703	-2.459	0.0237	**
precios	0.00474603	0.00567081	0.8369	0.4130	
rentapc	1.01835e-05	3.38587e-05	0.3008	0.7669	

Suma de cuadrados explicada = 0.545009

Estadístico de contraste: LM = 6.512092, $p = P(\chi^2_2 > 6{,}512092) = 0{,}038540$

Resumen de los test de Heterocedasticidad

```
Contraste de heterocedasticidad de White
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 11.0246
con valor p = P(Chi-cuadrado(5) > 11.0246) = 0.0508938
 Contraste de heterocedasticidad de White (cuadrados sólo)
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 7.34916
con valor p = P(Chi-cuadrado(4) > 7.34916) = 0.118548
 Contraste de heterocedasticidad de Breusch-Pagan
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 14.8397
con valor p = P(Chi-cuadrado(2) > 14.8397) = 0.00059925
 Contraste de heterocedasticidad de Breusch-Pagan (variante robusta)
Hipótesis nula: No hay heterocedasticidad
Estadístico de contraste: LM = 6.51209
con valor p = P(Chi-cuadrado(2) > 6.51209) = 0.0385405
```

Corrección de heterocedasticidad: Mínimos Cuadrados Ponderados (MCP)

La corrección por Mínimos Cuadrados Ponderados (MCP) consiste en buscar una matriz **P** de ponderaciones para el modelo inicial:

$$\mathbf{Py} = \mathbf{PX}\beta + \mathbf{Pu} \qquad \text{(Modelo transformado)}$$

Características

$$E(\mathbf{Pu}) = \mathbf{P}E(\mathbf{u}) = \mathbf{0}$$
$$Cov(\mathbf{Pu}) = E(\mathbf{Puu'P'}) = \sigma^2\mathbf{PP}^{-1}\mathbf{P}'^{-1}\mathbf{P}' = \sigma^2\mathbf{I}$$

Las ponderaciones W_i que dan lugar a la matriz **P** se aproximan analizando la estructura de la heterocedasticidad

$W_i = \frac{1}{X_i}$; $W_i = \frac{1}{\sqrt{X_i}}$; $W_i = \frac{1}{X_i^2}$ Varianza creciente con X

$W_i = X_i$; $W_i = \sqrt{X_i}$ Varianza decreciente con X

MCP: Ilustración $\boxed{\sigma_i^2 = \sigma^2 X_i}$

$$E\left[(W_i u_i)^2\right] = \sigma^2 \Rightarrow W_i^2 E\left[u_i^2\right] = \sigma^2 \Rightarrow$$

$$W_i^2 \sigma^2 X_i = \sigma^2 \Rightarrow \boxed{W_i = \frac{1}{\sqrt{X_i}}}$$

$$P = \begin{pmatrix} \frac{1}{\sqrt{X_1}} & & \\ & \frac{1}{\sqrt{X_2}} & \\ & & \frac{1}{\sqrt{X_n}} \end{pmatrix}$$

Para obtener las ponderaciones Wi más adecuadas puede ser útil examinar la regresión auxiliar de los test de White o Breusch-Pagan, comprobando qué variables explicativas son más significativas

Corrección de la heterocedasticidad

Ponderaciones habituales

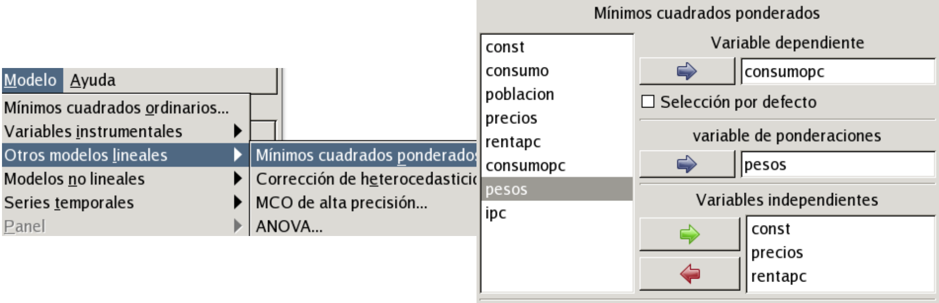
Dispersión directamente proporcional a X

Regresión auxiliar significativa		Estructura de heterocedasticidad		Ponderaciones
$\hat{u}_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 X_i$	$\Rightarrow$	$\sigma_i^2 = \sigma^2 X_i$	$\Rightarrow$	$W_i = \frac{1}{\sqrt{X_i}}$
$\hat{u}_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 X_i^2$	$\Rightarrow$	$\sigma_i^2 = \sigma^2 X_i^2$	$\Rightarrow$	$W_i = \frac{1}{X_i}$

Dispersión inversamente proporcional a X

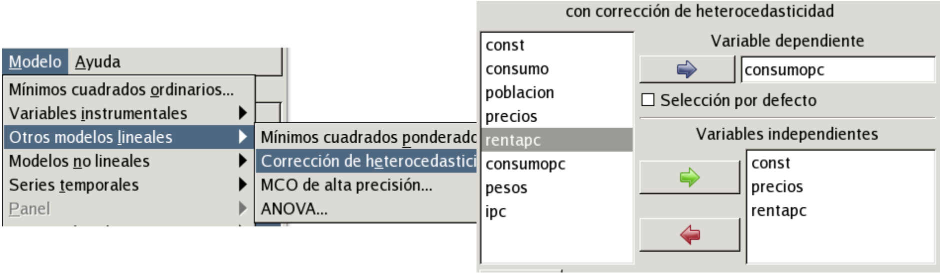
Regresión auxiliar significativa		Estructura de heterocedasticidad		Ponderaciones
$\hat{u}_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 \frac{1}{X_i}$	$\Rightarrow$	$\sigma_i^2 = \frac{\sigma^2}{X_i}$	$\Rightarrow$	$W_i = \sqrt{X_i}$
$\hat{u}_i^2 = \hat{\alpha}_1 + \hat{\alpha}_2 \frac{1}{X_i^2}$	$\Rightarrow$	$\sigma_i^2 = \frac{\sigma^2}{X_i^2}$	$\Rightarrow$	$W_i = X_i$

Estimación de MCP con Gretl



- Se introducen pesos en función de la estructura de heterocedasticidad existente
- Gretl estima el modelo ponderado con estos pesos

Corrección de heterocedasticidad automática en Gretl



- La corrección automática de Gretl se puede aplicar si no se conoce la estructura de la heterocedasticidad y asigna unas ponderaciones basadas en la regresión auxiliar de los logaritmos de los residuos cuadráticos

Corrección de heterocedasticidad

Modelo 2: con corrección de heterocedasticidad, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	0.409421	0.332413	1.2317	0.2331
precios	0.00935447	0.00763320	1.2255	0.2354
rentapc	5.82862e−05	3.13488e−05	1.8593	0.0785

Estadísticos basados en los datos ponderados:

Suma de cuad. residuos	35.14886	D.T. de la regresión	1.360125
R^2	0.872979	R^2 corregido	0.859608
$F(2, 19)$	65.29075	Valor p (de F)	3.07e−09
Log-verosimilitud	−36.37070	Criterio de Akaike	78.74139
Criterio de Schwarz	82.01452	Hannan--Quinn	79.51244
$\hat{\rho}$	0.665251	Durbin--Watson	0.549186

Estadísticos basados en los datos originales:

Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	9.043910	D.T. de la regresión	0.689924

- Los estadísticos comparables con el modelo inicialmente propuesto son los basados en datos originales (no en datos ponderados).
- La suma de cuadrados de los residuos habrá aumentado en esta estimación, ya que los estimadores MCO eran óptimos

Autocorrelación

$$Cov(\mathbf{u}) = E(\mathbf{u}\mathbf{u}') = \begin{pmatrix} \sigma^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{12} & \sigma^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \cdots & \sigma^2 \end{pmatrix}$$

Causas

- Mala especificación, inercia en las variables, cambios estructurales, ...

Consecuencias

- Ineficiencia de estimadores
- Estimación sesgada de las varianzas y test de significación no válidos

Detección

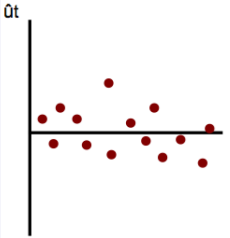
- Análisis gráfico de los residuos
- Test de Durbin y Watson
- Otros contrastes (LM, Q, ...)

Soluciones

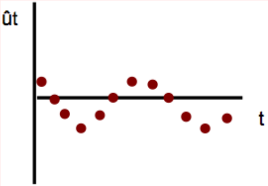
- Especificación de modelos autorregresivos AR

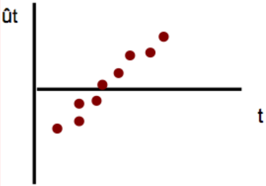
Análisis gráfico de los residuos

No autocorrelación



Autocorrelación





Contraste de Durbin-Watson

Esquema autorregresivo AR(1)

$$\mathbf{u}_t = \rho \mathbf{u}_{t-1} + \boldsymbol{\epsilon}_t \;\; ; \;\; |\rho| < 1 \;\; ; \;\; \boldsymbol{\epsilon}_t \approx \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

Hipótesis nula: no autocorrelación:

$$\boxed{\begin{array}{l} H_0 : \rho = 0 \\ H_1 : \rho \neq 0 \end{array}}$$

Información muestral: residuos $\hat{u}_t$

Discrepancia de Durbin-Watson:

$$\boxed{d_{DW} = \frac{\sum_{t=2}^{T} \left(\hat{u}_t - \hat{u}_{t-1}\right)^2}{\sum_{t=1}^{T} \hat{u}_t^2}}$$

Contraste de Durbin-Watson

$$
\begin{aligned}
d_{DW} &= \frac{\sum\limits_{t=2}^{T}(\hat{u}_t - \hat{u}_{t-1})^2}{\sum\limits_{t=1}^{T}\hat{u}_t^2} \\
&\approx \frac{2\sum\limits_{t=1}^{T}\hat{u}_t^2 - 2\sum_{i=2}^{T}\hat{u}_t^2\hat{u}_{t-1}^2}{\sum\limits_{t=1}^{T}\hat{u}_t^2} \\
&= 2\left(1 - \frac{Cov(\hat{u}_t, \hat{u}_{t-1})}{Var(\hat{u}_t)}\right) = 2(1-\hat{\rho})
\end{aligned}
$$

Valores del DW

- Si existe correlación lineal directa, $\hat{\rho} = 1 \quad \Rightarrow \quad d_{DW} = 0$
- Si existe correlación lineal inversa, $\hat{\rho} = -1 \quad \Rightarrow \quad d_{DW} = 4$
- Si no hay autocorrelación $\hat{\rho} = 0 \quad \Rightarrow \quad d_{DW} = 2$

Contraste de Durbin-Watson en Gretl

Modelo 1: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35170	0.348359	−3.8802	0.0010
precios	0.0498576	0.00860015	5.7973	0.0000
rentapc	−9.35994e–05	5.13489e–05	−1.8228	0.0841

Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	2.981244	D.T. de la regresión	0.396116
R^2	0.854650	R^2 corregido	0.839351
$F(2, 19)$	55.85969	Valor p (de F)	1.10e–08
Log-verosimilitud	−9.230929	Criterio de Akaike	24.46186
Criterio de Schwarz	27.73499	Hannan--Quinn	25.23291
$\hat{\rho}$	0.264285	**Durbin--Watson**	**1.401145**

En la salida del modelo, opción de Contrastes

Estadístico de Durbin-Watson = 1.40115
Valor p = 0.0261283

Esquema autorregresivo AR(1)

Modelo inicial: $\mathbf{Y}_t = \beta_1 + \beta_2 \mathbf{X}_t + \mathbf{u}_t$

$$\mathbf{u}_t = \rho \mathbf{u}_{t-1} + \boldsymbol{\epsilon}_t \quad ; \quad |\rho| < 1 \quad ; \quad \boldsymbol{\epsilon}_t \approx \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$$

$$\mathbf{Y}_t = \beta_1 + \beta_2 \mathbf{X}_t + (\rho \mathbf{u}_{t-1} + \boldsymbol{\epsilon}_t)$$

$$\mathbf{Y}_{t-1} = \beta_1 + \beta_2 \mathbf{X}_{t-1} + \mathbf{u}_{t-1} \quad \Rightarrow \quad \rho \mathbf{Y}_{t-1} = \rho\beta_1 + \rho\beta_2 \mathbf{X}_{t-1} + \rho \mathbf{u}_{t-1}$$

$$\mathbf{Y}_t - \rho \mathbf{Y}_{t-1} = \beta_1(1-\rho) + \beta_2 (\mathbf{X}_t - \rho \mathbf{X}_{t-1}) + \boldsymbol{\epsilon}_t$$

Modelo transformado (perturbaciones incorreladas)

$$\boxed{\mathbf{Y}_t^* = \beta_1^* + \beta_2 \mathbf{X}_t^* + \boldsymbol{\epsilon}_t}$$

Procedimiento iterativo de Cochrane-Orcutt

Procedimiento de Cochrane-Orcutt

Estimación del modelo original por MCO

Contraste de la hipótesis $H_0 : \rho = 0$

NO RECHAZAR

Estimación válida

Punto de parada

RECHAZAR

Corrección de autocorrelación

- Sobre los residuos $\hat{\mathbf{u}}$ estimación de ρ y estimación del modelo transformado $\mathbf{Y}_t - \rho \mathbf{Y}_{t-1} = \beta_1(1-\rho) + \beta_2 (\mathbf{X}_t - \rho \mathbf{X}_{t-1}) + \epsilon_t$
- **REPETIR** (procedimiento iterativo)

Estimación Cochrane-Orcutt

Gretl: Modelos → Series temporales → Cochrane-Orcutt

Realizando el cálculo iterativo de ρ ...

ITERACIÓN	RHO	SCR
1	0.26428	2.79644
2	0.27012	2.79634
3	0.27036	2.79634

Modelo 2: Cochrane--Orcutt , usando las observaciones 1988--2008 (T = 21)

Variable dependiente: consumopc

$\hat{\rho}$ = 0.270356

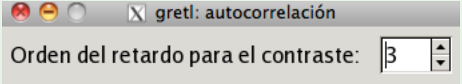
	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.39891	0.446858	−3.1305	0.0058
precios	0.0506968	0.0104446	4.8539	0.0001
rentapc	−9.50947e–05	6.85507e–05	−1.3872	0.1823

Estadísticos basados en los datos rho-diferenciados:

Media de la vble. dep.	2.201685	D.T. de la vble. dep.	0.996326
Suma de cuad. residuos	2.796343	D.T. de la regresión	0.394148
R^2	0.859396	R^2 corregido	0.843774
$F(2,18)$	36.45864	Valor p (de F)	4.67e–07
$\hat{\rho}$	0.156240	Durbin--Watson	1.618881

Otros contrastes de autocorrelación

- Durbin-Watson sólo contrasta la autocorrelación de primer orden y no puede ser aplicado en modelos que contienen como explicativas variables endógenas retardadas
- Para solucionar estos inconvenientes existen otros contrastes como los de Breusch-Godfrey (LM) o Ljung-Box (Q), en los que la hipótesis nula es ausencia de correlación hasta cierto orden de retardos



Contraste de Autocorrelación

Gretl: Salida de modelo, Contrastes → Autocorrelación

Contraste Breusch-Godfrey de autocorrelación hasta el orden 3
MCO, usando las observaciones 1987-2008 (T = 22)
Variable dependiente: $\hat{u}$

	Coeficiente	Desv. Típica	Estadístico t	Valor p	
const	0.591824	0.347077	1.705	0.1075	
precios	-0.0164693	0.00894872	-1.840	0.0843	*
rentapc	7.52747e-05	4.84594e-05	1.553	0.1399	
$\hat{u}_1$	0.209950	0.245743	0.854	0.4055	
$\hat{u}_2$	-0.910259	0.308967	-2.946	0.0095	***
$\hat{u}_3$	-0.105322	0.424803	-0.248	0.8073	

R-cuadrado = 0.490561
Estadístico de contraste: $LMF = 5{,}135702$, $p = P(F^3_{16} > 5{,}1357) = 0{,}0112$

Estadístico alternativo: $TR^2 = 10{,}792346$, $p = P(\chi^2_3 > 10{,}7923) = 0{,}0129$

Ljung-Box Q' = 8.71089, $p = P(\chi^2_3 > 8{,}71089) = 0{,}0334$

Rechazamos la hipótesis de que no hay autocorrelación hasta el orden 3

No Normalidad

Causas

- Presencia de observaciones atípicas (Outliers)
- Problemas en la especificación del modelo

Consecuencias

- Se desconoce el modelo probabilístico de los estimadores MCO y de $\hat{u}$
- Se desconocen las distribuciones necesarias para los procesos inferenciales

Detección

- Contrastes:
  - Chi-cuadrado
  - Otros: Jarque y Bera, Kolmogorov-Smirnov, Shapiro-Wilks, ...

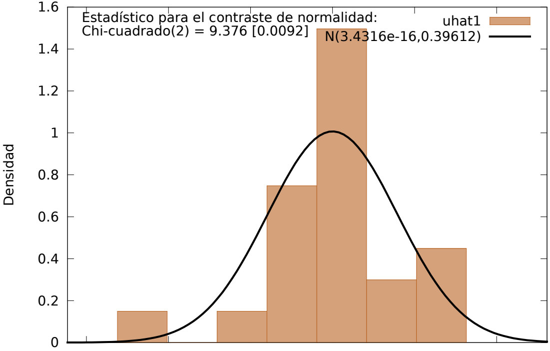
Soluciones

- Introducir variables dummy
- Cambiar la especificación

Contraste de normalidad en los residuos

Gretl: Salida de modelo, Contrastes → Normalidad de los residuos

- Para contrastar la normalidad se analizan los residuos del modelo estimado, clasificándolos en intervalos y comparando la frecuencia de cada intervalo con la que le correspondería bajo el supuesto de normalidad.
- Los resultados se resumen en una chi-cuadrado con 2 g.l.
- Para niveles críticos bajos se rechaza la normalidad



Otros contrastes de normalidad

Gretl: Variable → Contraste de normalidad

Guardar los residuos del modelo

Salida del modelo: Guardar → Residuos

Contraste de normalidad de $\hat{u}_1$:

Contraste de Doornik-Hansen = 9.3757, con valor $p = 0{,}00920645$
W de Shapiro-Wilk = 0.897126, con valor $p = 0{,}0260504$
Contraste de Lilliefors = 0.182989, con valor $p \cong 0{,}05$

Contraste de Jarque-Bera = 9.69534, con valor $p = 0{,}00784662$

Todos los test nos llevan a rechazar la hipótesis de normalidad de los residuos

Causas de la no normalidad de los residuos

- El incumplimiento de la hipótesis de normalidad puede ser debido a problemas en la especificación del modelo o a la presencia de observaciones atípicas (outliers).
- Cuando se identifica algún dato atípico o outlier sobre la serie de la variable dependiente, es aconsejable revisar la información para ver si existen errores en los datos.
- En caso de que el outlier sea debido a comportamientos anómalos como huelgas, accidentes, tormentas, ..., es aconsejable introducir en el modelo una variable dummy asociada a estos hechos.

Resumen: Problemas con $\mathbf{u}$ y soluciones

$E(\mathbf{u}) \neq 0$

- **Revisar especificación**

$Cov(\mathbf{u}) = \sigma^2 \mathbf{\Omega}$

- **Revisar especificación**
- Estimación consistente de la matriz de Var-Cov
- ¿Heterocedasticidad? Estimación MCP
- ¿Autocorrelación? AR(1)

$\mathbf{u}$ no normal

- **Revisar especificación**
- ¿Valores atípicos? Variables dummy

Regresores **X** estocásticos

La matriz **X** tiene carácter estocástico cuando:

- Se incluyen como explicativas variables endógenas retardadas
- Se especifican modelos multiecuacionales

Consecuencias

- Si la matriz **X** fuese independiente de las perturbaciones **u** entonces los estimadores MCO serían insesgados y consistentes
- En general la matriz **X** estará correlacionada contemporáneamente con **u**, y los estimadores MCO serán sesgados e inconsistentes, por lo que habría que realizar una estimación con **variables instrumentales**

Matriz **X** de rango no pleno

Tamaño muestral n insuficiente (Micronumerosidad)

- Situaciones extremas si n=0
- Problemas con pocos grados de libertad n-k

Correlación entre variables explicativas (Multicolinealidad)

- Situaciones extremas (Multicolinealidad perfecta):
  - Variable explicativa constante
  - Variables explicativas relacionadas linealmente
  - Excesivas variables cualitativas ("trampa" de v. ficticias)
- Problemas con multicolinealidad aproximada (alto grado de correlación entre las **X**)

Las propiedades de estimadores MCO no varían

Hay un aumento de varianzas

Principio de parsimonia

Multicolinealidad

Causas

- Existencia de relaciones lineales entre las variables explicativas del modelo

Consecuencias

- Multicolinealidad perfecta
  - La matriz $\mathbf{X}'\mathbf{X}$ es singular, no puede obtenerse el vector de estimadores $\hat{\beta}$
- Multicolinealidad aproximada
  - Aumenta la matriz Var-Cov
  - Modelo inestable
  - No afecta a la predicción

Detección

- Analizar:
  - Contrastes de significación (t y F)
  - Factores de inflación de la varianza (VIF)
  - Matriz de correlaciones de las variables $\mathbf{X}$

Soluciones

- El problema de la información muestral no tiene una solución concreta
- Recomendaciones: ampliar la información muestral, omitir variables del modelo, ...

Multicolinealidad

Consecuencias

- Con multicolinealidad perfecta no es posible calcular los estimadores MCO, ya que la matriz $\mathbf{X}'\mathbf{X}$ es singular:

$$\rho(\mathbf{X}) < k \Rightarrow \rho(\mathbf{X'X}) < k \Rightarrow |X'X| = 0$$

- Con multicolinealidad aproximada se mantienen las propiedades de los estimadores, pero aumenta la matriz var-cov, con lo que resulta difícil rechazar la nulidad de los coeficientes en los contrastes t.
- Con multicolinealidad el modelo resulta inestable, con lo que se observan alteraciones importantes en los coeficientes estimados y su significación al cambiar el tamaño muestral o la especificación del modelo.
- La multicolinealidad no afecta a las predicciones del modelo

Detección de la multicolinealidad con Gretl

Modelo 2: MCO, usando las observaciones 1987--2008 (T = 22)
Variable dependiente: consumopc

	Coeficiente	Desv. Típica	Estadístico t	Valor p
const	−1.35103	0.354700	−3.8089	0.0013
precios	0.0283484	0.0386238	0.7340	0.4724
rentapc	−9.30222e–05	5.22930e–05	−1.7789	0.0922
ipc	0.0281372	0.0492098	0.5718	0.5745

Media de la vble. dep.	2.163958	D.T. de la vble. dep.	0.988285
Suma de cuad. residuos	2.928062	D.T. de la regresión	0.403324
R^2	0.857243	R^2 corregido	0.833451
$F(3, 18)$	36.02957	Valor p (de F)	8.10e–08
Log-verosimilitud	−9.032929	Criterio de Akaike	26.06586
Criterio de Schwarz	30.43003	Hannan--Quinn	27.09392
$\hat{\rho}$	0.318925	Durbin--Watson	1.309339

Detección de la multicolinealidad con Gretl

Estabilidad del modelo

$$\widehat{\text{consumopc}} = \underset{(0{,}35470)}{-1{,}35103} + \underset{(0{,}038624)}{0{,}0283484}\,\text{precios} - \underset{(5{,}2293\text{e}{-}05)}{9{,}30222\text{e}{-}05}\,\text{rentapc} + \underset{(0{,}049210)}{0{,}0281372}\,\text{ipc}$$

$$T = 22 \quad \bar{R}^2 = 0{,}8335 \quad F(3,18) = 36{,}030 \quad \hat{\sigma} = 0{,}40332$$

(Desviaciones típicas entre paréntesis)

$$\widehat{\text{consumopc}} = \underset{(0{,}34836)}{-1{,}35170} + \underset{(0{,}0086001)}{0{,}0498576}\,\text{precios} - \underset{(5{,}1349\text{e}{-}05)}{9{,}35994\text{e}{-}05}\,\text{rentapc}$$

$$T = 22 \quad \bar{R}^2 = 0{,}8394 \quad F(2,19) = 55{,}860 \quad \hat{\sigma} = 0{,}39612$$

(Desviaciones típicas entre paréntesis)

$$\widehat{\text{consumopc}} = \underset{(0{,}35285)}{-1{,}14025} + \underset{(0{,}039701)}{0{,}0127562}\,\text{precios} + \underset{(0{,}051927)}{0{,}0298270}\,\text{ipc}$$

$$T = 22 \quad \bar{R}^2 = 0{,}8145 \quad F(2,19) = 47{,}097 \quad \hat{\sigma} = 0{,}42568$$

(Desviaciones típicas entre paréntesis)

Multicolinealidad en Gretl

Gretl: En la salida del modelo, Contrastes → Colinealidad

Factores de inflación de varianza (VIF)

Mínimo valor posible = 1.0
Valores mayores que 10.0 pueden indicar un problema de colinealidad

precios	124.024
rentapc	6.377
ipc	117.681

$VIF_j = \frac{1}{1-R_j^2}$, donde R_j^2 es el coeficiente de correlación múltiple entre la variable j y las demás variables independientes

En este caso admitimos que existe multicolinealidad

Cambio estructural

Causas

- Existencia de rupturas o cambios en las relaciones entre las variables del modelo

Consecuencias

- Pérdida de validez del modelo y riesgos de extrapolación

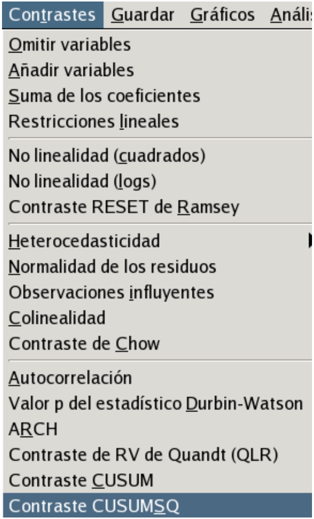
Detección

- Análisis gráfico
- Contrastes de CUSUM y CUSUMSQ, de Chow, RV de Quandt

Soluciones

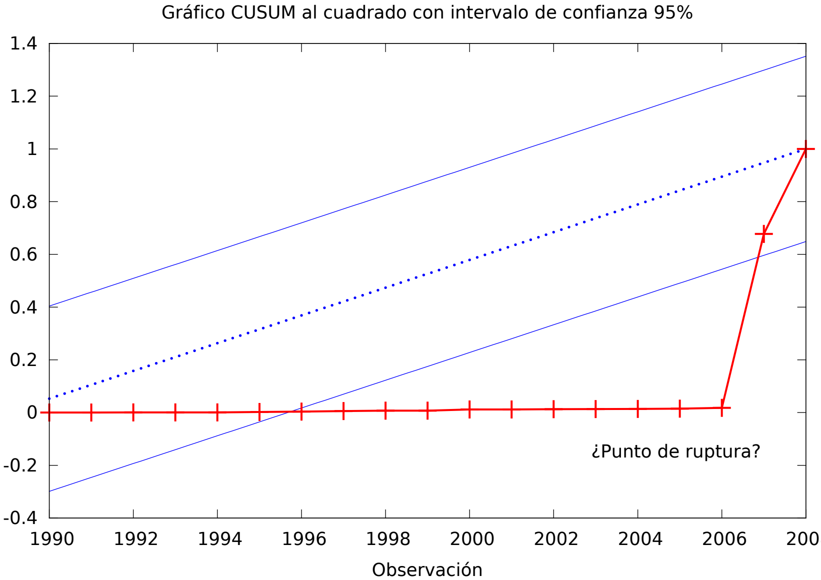
- Especificación de varios modelos diferentes
- Introducción de variables dummy

Test recursivos CUSUM y CUSUMSQ



- Se estima inicialmente el modelo para una muestra pequeña $1, \ldots, n_1$
- Se repite la estimación sucesivamente para las muestras $1, \ldots, n_1 + 1;\ 1, \ldots, n_1 + 2; \ldots$ y así hasta el periodo global $1, \ldots, n$.
- Se analizan las sumas acumuladas de residuos (CUSUM) o las sumas acumuladas de residuos cuadráticos (CUSUMSQ).
- **Cambios drásticos en las sumas acumulativas indican posibles puntos de ruptura del modelo**

Test recursivos CUSUM y CUSUMSQ



Contraste de Chow

Hipótesis: No hay cambio estructural

$$\boxed{\begin{array}{ll} H_0 : \beta_1 = \beta_2 = \beta & \sigma_1 = \sigma_2 = \sigma \\ H_1 : \beta_1 \neq \beta_2 \quad o & \qquad \sigma_1 \neq \sigma_2 \end{array}}$$

Regresiones propuestas

$$\left\{ \begin{array}{ll} \mathbf{y} = \mathbf{X}\beta + \mathbf{u} & con\ \mathbf{u} \approx N\left(\mathbf{0}, \sigma^2 \mathbf{I}\right) \\ \mathbf{y_1} = \mathbf{X_1}\beta_1 + \mathbf{u_1} & con\ \mathbf{u_1} \approx N\left(\mathbf{0}, \sigma_1^2 \mathbf{I}\right) \\ \mathbf{y_2} = \mathbf{X_2}\beta_2 + \mathbf{u_2} & con\ \mathbf{u_2} \approx N\left(\mathbf{0}, \sigma_2^2 \mathbf{I}\right) \end{array} \right.$$

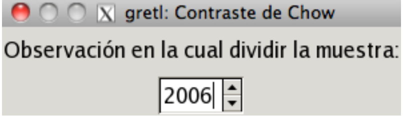
$$\frac{\frac{\hat{\mathbf{u}}`\hat{\mathbf{u}} - \left(\hat{\mathbf{u}}_1'\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2'\hat{\mathbf{u}}_2\right)}{k}}{\frac{\hat{\mathbf{u}}_1'\hat{\mathbf{u}}_1 + \hat{\mathbf{u}}_2'\hat{\mathbf{u}}_2}{n_1 + n_2 - 2k}} \approx F^k_{n_1 + n_2 - 2k}$$

Rechazamos para valores elevados de F (niveles críticos bajos)

Limitaciones: Exige conocer punto de ruptura y disponer de suficientes observaciones muestrales

Contraste de Chow en Gretl

Gretl: En la salida del modelo, Contrastes → Contraste de Chow



- Para contrastar la estabilidad estructural se realizan regresiones ampliadas incluyendo una variable dummy ligada al posible cambio estructural y funciones de las variables explicativas.
- El test contrasta la nulidad de los coeficientes de estas variables añadidas.

Contraste de Chow de cambio estructural en la observación 2006

$F^3_{16} = 354{,}652$ con valor $p = 0{,}0000$

Contraste RV de Quandt en Gretl

Gretl: En la salida del modelo, Contrastes → Contraste RV de Quandt (QRV)

Este test se basa en la razón de verosimilitudes y tiene la ventaja de que determina posibles puntos de ruptura

```
Contraste de razón de verosimilitudes de Quandt para cambio estructural
en un punto desconocido, con recorte del 15 por ciento:

El valor máximo de F^3_16 = 49,9244 corresponde a la observación 2005
Significativo al nivel del 1 por ciento (Valor crítico al 1% = 6.02)

Este estadístico no sigue la distribución F estándar; los valores
críticos provienen de Stock y Watson (2003).
```

Tratamiento del cambio estructural

Una vez detectada una ruptura en un modelo es posible:

- Estimar modelos diferentes antes y después del cambio (permite cambiar la especificación pero exige tamaños muestrales elevados)
- Introducir en el modelo variables cualitativas asociadas a los cambios estructurales

$$\boxed{W = \begin{cases} 0 & para\ t < t^* \\ 1 & para\ t \geq t^* \end{cases}}$$

Problemas y soluciones:

¿Multicolinealidad?

- Ampliar la muestra actual
- Omitir variables del modelo

¿Cambio estructural?

- Especificaciones de modelos diferentes
- Introducción de variables dummy