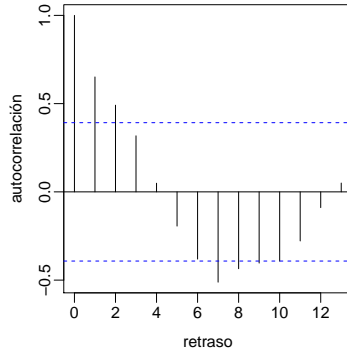
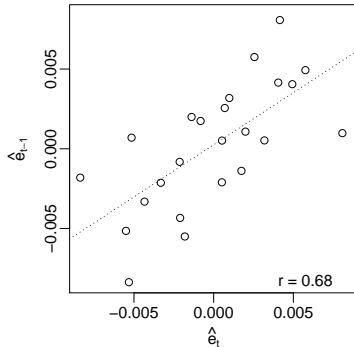


8 Verificación de Supuestos



Contenido

1

Introducción

- Supuestos

2

Normalidad

- Prueba Jarque-Bera
- Robustez de la prueba t

3

Heteroscedasticidad

- Detección
- Mínimos Cuadrados Generalizados y Ponderados

4

Autocorrelación

- Fuentes y consecuencias de la autocorrelación
- Detección
- Estadístico Durbin-Watson
- Corrección de Cochran-Orcutt

5

Colinealidad

- Fuentes de la colinealidad
- Efectos de la colinealidad.
- Diagnóstico de la colinealidad
- Corrección

Modelo de regresión lineal y supuestos

Considere el modelo de *regresión lineal múltiple*:

$$y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} + \epsilon_i$$

donde,

- y_i : respuesta al nivel i -ésimo nivel del regresor ($i = 1, \dots, n$)
- x_{ij} : j -ésimo regresor a su i -ésimo nivel ($j = 1, \dots, p$)
- β_0 : ordenada al origen
- β_j : coeficiente (tasa de cambio) del j -ésimo regresor

O bien, matricialmente

$$y = X\beta + \epsilon$$

Supuestos:

$$\epsilon \sim N_n(0, \sigma^2 I_n)$$

$$X_{n \times q} \text{ fija, } \text{rango}(X) = p + 1 = q (\leq n)$$

En ocasiones, no se satisfacen uno o más de los supuestos por lo que las conclusiones alcanzadas son cuestionables.

Una manera de tratar de verificar si los supuestos se cumplen o no es mediante el *análisis de residuales*.

Para algunos tipos de problemas (supuestos no satisfechos) hay pruebas estadísticas formales. Veremos algunas de ellas.

Algunos de los problemas más comunes de los modelos de regresión lineal son:

- Errores con distribución distinta a la normal.
- Heteroscedasticidad (varianza no constante).
- Autocorrelación (correlación de errores).
- Colinealidad (regresores linealmente dependientes).

Problema de No Normalidad

Los residuales son “errores observados” *si el modelo es correcto*. Luego, intentamos verificar los supuestos analizando los residuales.

- *Gráfica cuantil-cuantil normal* (gráfica de probabilidad normal).
- *Prueba Jarque-Bera*. Compara de manera conjunta el coeficiente de asimetría y curtosis contra los correspondientes parámetros de la distribución normal.
- Otras pruebas de *bondad de ajuste*, como son: *Lilliefors*, *Shapiro-Wilks*, *Anderson-Darling*, *Cramér-Von Mises*. Sin embargo, todas las pruebas suponen que las observaciones son *independientes*. (R: paquetes `moments` y `norstest`.)

Consecuencias

Jarque and Bera (1987) comentan que “*violaciones al supuesto de normalidad puede llevar al uso de estimadores subóptimos, aseveraciones inferenciales inválidas y conclusiones inexactas*”.

Prueba Jarque-Bera

La prueba de *Jarque-Bera* se basa en la *prueba de score*. Compara de manera conjunta el coeficiente de asimetría y curtosis contra los correspondientes parámetros de la distribución normal, resumido en el siguiente estadístico de prueba

$$JB = n \left(\frac{s^2}{6} + \frac{(k-3)^2}{24} \right)$$

donde,

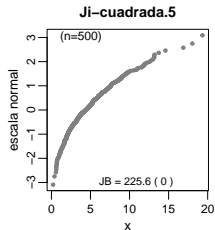
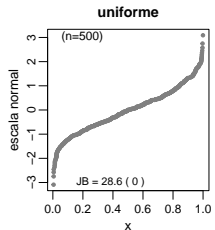
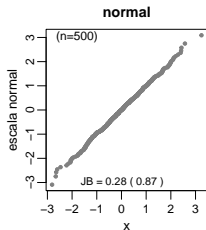
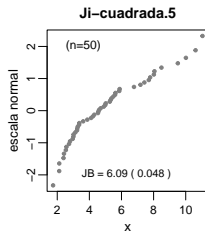
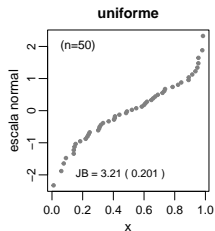
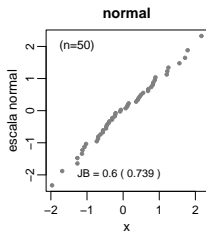
$$s = \frac{m_3}{(m_2)^{3/2}} = \frac{\frac{1}{n} \sum (x_i - \bar{x})^3}{\left[\frac{1}{n} \sum (x_i - \bar{x})^2 \right]^{3/2}} \quad \text{y} \quad k = \frac{m_4}{(m_2)^2} = \frac{\frac{1}{n} \sum (x_i - \bar{x})^4}{\left[\frac{1}{n} \sum (x_i - \bar{x})^2 \right]^2}$$

con $m_r = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^r$, el r -ésimo momento central muestral.

Los autores muestran que bajo el supuesto de normalidad,

$$JB \xrightarrow{d} \chi_2^2$$

ya la prueba es asintóticamente eficiente, que no lo es para muestras pequeñas ($n \leq 100$).

Muestras simuladas y estadístico Jarque-Bera JB (valor- p)

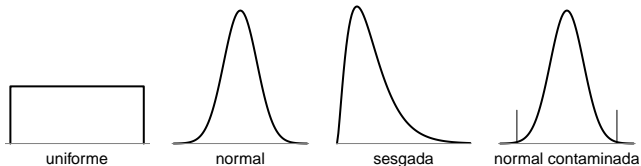
Corrección

En ocasiones la normalidad se logra mediante alguna transformación, por ejemplo, del tipo Box-Cox. A saber, Box and Cox (1964) mencionan al inicio de su artículo que “las contribuciones de la transformación a la normalidad, homoscedasticidad y aditividad son por separado”.

Las transformaciones Box-Cox favorecen tanto la estabilidad de la varianza como la *normalidad* de los datos.

Nota: La prueba [Jarque-Bera](#) es popular en la práctica de la econometría. Sin embargo, en muchos casos, el análisis gráfico es suficiente debido a la robustez de los estadísticos t y F a desviaciones moderadas de la normalidad.

Sobre la robustez de la prueba t



correlación serial ϕ	estadístico de prueba	distribución			
		uniforme	normal	sesgada	normal contaminada
0	t	0.053	0.051	0.045	0.052
	W	0.056	0.054	0.050	0.550
-0.4	t	0.008	0.007	0.007	0.018
	W	0.010	0.008	0.008	0.017
+0.4	t	0.185	0.186	0.185	0.118
	W	0.180	0.185	0.188	0.133

donde t y W son los estadísticos de las pruebas t y de Wilcoxon, respectivamente.
 La correlación serial supone que $X_k = \phi X_{k-1} + a_k$, con $\{a_k\}$ ruido blanco gaussiano.

Problema de Heteroscedasticidad

Detección

El problema de la *heteroscedasticidad* se da cuando la varianza no es constante.

Una manera de verificar el supuesto $\text{var}(\epsilon_i) = \sigma^2, i = 1, \dots, n$ es mediante las gráficas de residuales.

En el contexto del *diseño de experimentos* y los *modelos de análisis de varianza* existen pruebas formales, pero todas ellas hacen uso de *réplicas legítimas* que pocas veces se tienen en los modelos de regresión.

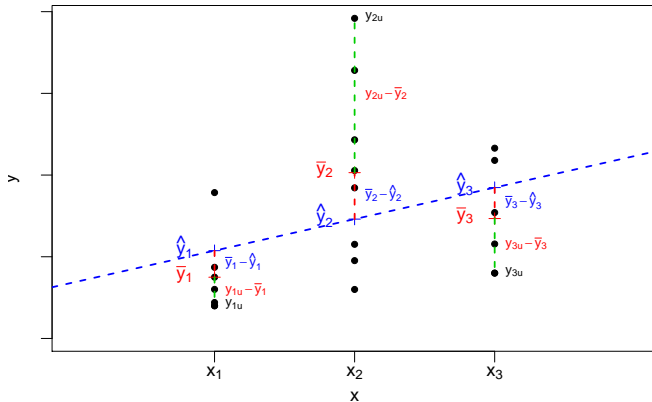
Entre las pruebas formales para probar la igualdad de varianzas están:

- Prueba de Bartlett, prueba de Levene, prueba de Box.
- Prueba de Breuch, Pagan y Koenker, prueba de White.

En ocasiones se conoce la estructura de la matriz de covarianza $\text{cov}(\epsilon) = \Sigma$ de los errores. De ser así, la podemos incorporar explícitamente en los cálculos.

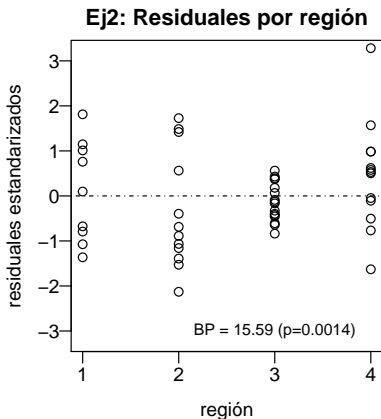
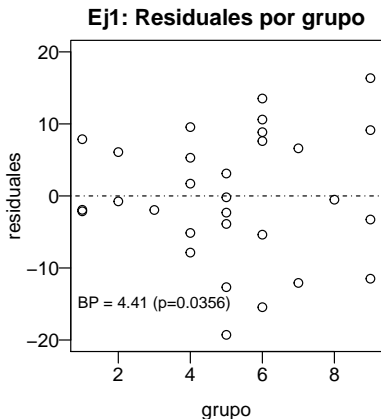
Problema de Heteroscedasticidad

¿Varianza constante?



Problema de Heteroscedasticidad

¿Varianza constante?



Prueba de Breusch y Pagan¹

Considere el modelo RLM

$$y_t = \beta_0 + \beta_1 x_{1t} + \cdots + \beta_p x_{pt} + \epsilon_t \quad (1)$$

Sea $\sigma_t^2 = \text{var}(\epsilon_t)$ y $\sigma_t^2 = h(z_t; \alpha)$, donde h es dos veces diferenciable, $z = (z_1, \dots, z_k)$ variables auxiliares posiblemente relacionadas con σ y α un vector de parámetros. Es común utilizar el modelo

$$\sigma_t^2 = \alpha_0 + \alpha_1 x_{1t} + \cdots + \alpha_p x_{pt} \quad (2)$$

Bajo la hipótesis

$$H_0 : \alpha_1 = \cdots = \alpha_p = 0 \quad (3)$$

$\sigma_t^2 = \alpha_0$ y las varianzas son constantes.

Sean $\hat{\epsilon}_t = y_t - x_t^T \hat{\beta}$, los correspondientes residuales del modelo (1). Los residuales $\hat{\epsilon}_t$ son la mejor representación disponible del error ϵ_t , por lo que $\hat{\epsilon}_t^2$ es una estimación de la varianza σ_t^2 .

Ajuste ahora el modelo

$$\hat{\epsilon}_t^2 = \alpha_0 + \alpha_1 x_{1t} + \cdots + \alpha_p x_{pt} + \xi_t \quad (4)$$

y determine R_e^2 el coeficiente de determinación múltiple del modelo (4). Entonces el estadístico de prueba es

$$BP = nR_e^2 \sim \chi_p^2 \quad (5)$$

donde n es el número de observaciones y p el número de regresores en (4).

¹Breusch and Pagan (1979); Koenker (1981).

Prueba de Breusch y Pagan, Koenker

- La prueba supone conocimiento previo de qué podría causar la heteroscedasticidad. Esto se refleja con las variables auxiliares z .
- La prueba de Breusch y Pagan es muy sensible a desviaciones de la normalidad.
- La *studentización* del estadístico de prueba original de Breusch y Pagan es el presentado en (5) fue propuesto por Koenker (1981) y depende menos del supuesto de normalidad.

Prueba de White²

- La Prueba de White está relacionada con la de Breusch y Pagan pero no supone conocimiento previo sobre la heteroscedasticidad.
- Parecida a la Prueba de Breusch y Pagan pero con un conjunto de variables z 's en particular y no depende de la normalidad de los errores.
- La función h representada por el modelo (2) en este caso queda como

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^p \alpha_i x_{it} + \sum_{i=1}^p \alpha_{ii} x_{it}^2 + \sum_{ij} \alpha_{ij} x_{it} x_{jt} \quad (6)$$

- Ajuste el modelo RLM (1) y obtenga los residuales $\hat{\epsilon}_t$.
- Ajuste el modelo (6) para las varianzas ($\hat{\epsilon}_t^2$) y aplique la prueba de Breusch y Pagan.
- Rechace la hipótesis de varianza constante si

$$nR_e^2 > \chi^2(1 - \gamma; k)$$

donde n es el número de observaciones, R_e^2 es el coeficiente de determinación del modelo(6), k el número de coeficientes en el modelo (sin considerar α_0) y γ el nivel de significancia de la prueba.

²White (1980).

Ejemplo: Televisión por cable (cont.)

```
> mod1 <- lm(renta ~ ninos + adultos + tvtot + I(valor/1000), data=dat)
> summary(mod1)
```

Coefficients:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	9.80566	10.32633	0.950 0.348838
ninos	-4.91432	2.73477	-1.797 0.080973
adultos	2.64006	2.44211	1.081 0.287065
tvtot	0.45053	0.11445	3.936 0.000375
I(valor/1000)	0.12989	0.03141	4.135 0.000211

Residual standard error: 11.99 on 35 degrees of freedom

Multiple R-squared: 0.5916, Adjusted R-squared: 0.545

F-statistic: 12.68 on 4 and 35 DF, p-value: 1.772e-06

```
> library(lmtest)
```

```
> bptest(mod1)
```

Prueba Breusch-Pagan:

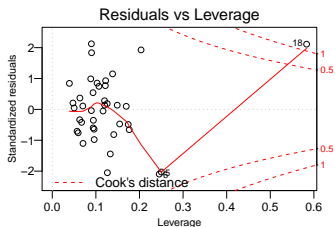
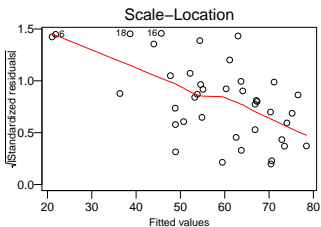
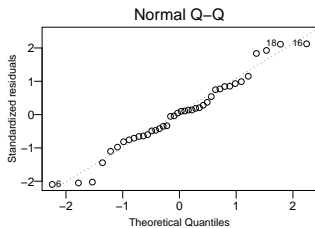
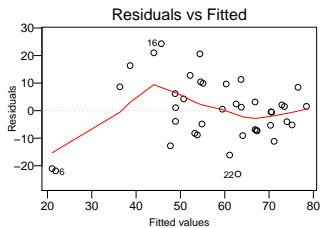
studentized Breusch-Pagan test

data: mod1

BP = 17.138, df = 4, p-value = 0.001817

Ejemplo: Televisión por cable (cont.)

Análisis de residuales mod1



Ejemplo: Televisión por cable (cont.)

```
> influence.measures(mod1)
```

```
Influence measures of
```

```
lm(formula = renta ~ ninos + adultos + tvtot + I(valor/1000), data = dat) :
```

dfb.1_	dfb.nins	dfb.adlt	dfb.tvtt	dfb.I.1.1	dffit	cov.r	cook.d	hat	inf
1	0.023748	-0.016212	0.01315	0.00733	-0.036422	0.0442	1.390	4.03e-04	0.1696
2	0.072893	0.021779	0.13039	0.09404	-0.236953	0.4633	1.106	4.25e-02	0.1383
3	0.224457	-0.034260	-0.14998	-0.00328	-0.146391	0.2851	1.208	1.65e-02	0.1218
4	-0.204652	0.080783	0.06338	-0.06754	0.197737	-0.3045	1.042	1.84e-02	0.0705
5	-0.633730	-0.489161	0.04981	0.93488	0.389236	-1.2277	0.826	2.74e-01	0.2501
6	-0.623916	-0.524295	0.05925	0.96471	0.362579	-1.2594	0.783	2.86e-01	0.2453
7	0.537617	-0.260280	-0.00841	-0.10745	-0.462822	0.5988	0.766	6.67e-02	0.0901
8	-0.293777	0.408067	-0.09501	-0.25050	0.380844	-0.5748	0.981	6.40e-02	0.1331
9	-0.065237	0.079371	0.06607	0.08086	-0.023482	0.2872	1.160	1.66e-02	0.1028
10	-0.083941	0.043308	0.06257	-0.07494	0.068407	-0.1722	1.141	6.02e-03	0.0569
11	0.010861	-0.042421	0.00359	0.07081	-0.032117	0.1019	1.299	2.13e-03	0.1199
12	-0.074077	0.042761	0.01136	0.00345	0.056821	-0.0856	1.214	1.50e-03	0.0629
13	0.080206	-0.055064	-0.01908	0.01057	-0.060905	0.0945	1.210	1.83e-03	0.0633
14	-0.147637	0.122029	0.05671	-0.05331	0.103188	-0.1893	1.131	7.25e-03	0.0594
15	0.078212	-0.054347	0.14409	-0.20080	-0.094412	-0.3288	1.223	2.18e-02	0.1411
16	0.291277	0.093546	-0.51527	0.03049	0.021680	0.7047	0.636	8.91e-02	0.0897
17	-0.012831	-0.009399	0.02933	0.02267	-0.009774	0.0564	1.355	6.55e-04	0.1490
18	-0.649953	1.927511	1.89461	-2.20973	0.014625	2.6380	1.402	1.25e+00	0.5833
19	-0.035754	0.009955	-0.07106	0.05272	0.048486	-0.1113	1.208	2.54e-03	0.0673
20	0.193892	-0.170827	-0.28681	0.16585	-0.026155	0.3466	1.162	2.41e-02	0.1224
21	-0.025931	0.001089	0.04435	0.01669	-0.002580	0.0703	1.316	1.02e-03	0.1259
22	0.348784	-0.381146	0.29761	-0.28315	-0.426902	-0.8200	0.697	1.22e-01	0.1264
23	0.001644	-0.000951	0.00640	-0.00370	-0.002470	0.0104	1.217	2.24e-05	0.0502
24	0.088159	-0.019787	-0.12113	-0.00985	0.012624	0.1722	1.223	6.05e-03	0.0935
25	-0.284442	0.575202	-0.57910	0.06758	0.601950	1.0153	0.829	1.90e-01	0.2037

-ebz

Ejemplo: Televisión por cable (cont.)

```
> head(dat, 25)
```

obs	colonia	manzana	adultos	ninios	teles	renta	tvttot	tipo	valor
1	2	20	3	2	2	50	68	B	79928
2	2	25	3	3	1	65	82	B	94415
3	2	20	1	2	1	45	40	A	120896
4	2	8	2	2	2	35	56	A	132867
5	2	25	1	2	0	0	0	N	141901
6	2	14	1	2	0	0	0	N	147997
7	2	22	2	1	1	65	30	A	156410
8	2	20	3	1	3	45	62	C	156841
9	2	25	3	3	2	70	82	A	157041
10	2	20	2	2	3	45	60	C	161222
11	2	8	3	2	1	70	84	A	162509
12	2	8	2	1	3	45	34	A	180124
13	2	14	2	1	1	55	38	C	180437
14	2	8	2	1	2	45	42	A	190314
15	2	14	2	3	1	55	86	A	192265
16	2	25	1	2	3	70	40	B	192816
17	2	14	4	2	4	75	84	C	193279
18	2	20	4	3	3	55	14	C	205656
19	1	2	3	1	3	50	31	C	216190
20	2	22	1	1	3	65	42	C	216321
21	2	22	4	2	2	75	76	C	216465
22	2	22	2	3	2	40	74	C	225694
23	1	4	3	1	3	60	35	C	237752
24	2	25	1	1	1	55	22	C	241531
25	2	8	1	3	2	75	54	C	249098

Ejemplo: Televisión por cable (cont.)

```
> mod2 <- lm(renta ~ ninos + adultos + tvtot + I(valor/1000), data=dat, subset=seq(40)[-c(5,6)])
> print(summary(mod2))

Coefficients:
Estimate Std. Error t value Pr(>|t|)
(Intercept) 28.21648    9.50180   2.970  0.00552
ninos        -0.98699    2.42591  -0.407  0.68674
adultos      2.26288    2.01084   1.125  0.26856
tvtot        0.14240    0.11805   1.206  0.23632
I(valor/1000) 0.09640    0.02698   3.574  0.00111

Residual standard error: 9.866 on 33 degrees of freedom
Multiple R-squared:  0.3804, Adjusted R-squared:  0.3053
F-statistic: 5.065 on 4 and 33 DF,  p-value: 0.002701

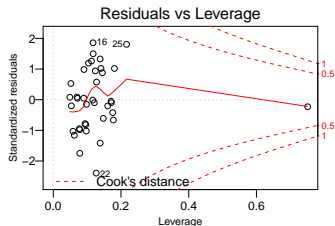
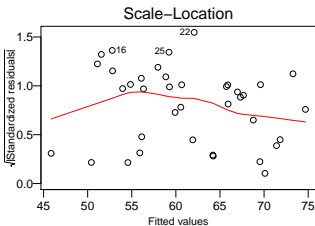
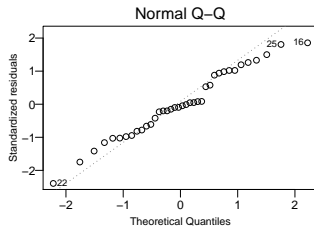
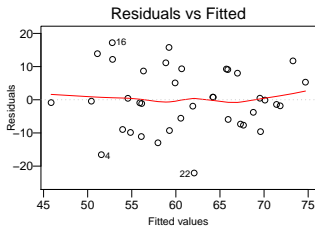
> print(bptest(mod2))

studentized Breusch-Pagan test

data:  mod2
BP = 7.6123, df = 4, p-value = 0.1069
```

Ejemplo: Televisión por cable (cont.)

Análisis de residuales mod2



Ejemplo: Televisión por cable (cont.)

```
> influence.measures(mod2)
Influence measures of
lm(formula = renta ~ ninos + adultos + tvtot + I(valor/1000), data = dat, subset = seq(40)[-c(5, 6)]) :

dfb.l_   dfb.nins dfb.adlt  dfb.tvtt  dfb.I..1  dffit cov.r   cook.d   hat inf
1 -0.010784  0.006404 -0.00610 -0.001675  0.016920 -0.02088  1.406 8.99e-05 0.1709
2  0.085559  0.031565  0.15075  0.074327 -0.270498  0.53988  1.028 5.69e-02 0.1386
3 -0.035636 -0.004593  0.01991  0.014206  0.024306 -0.04272  1.406 3.76e-04 0.1715
4 -0.375441  0.060129  0.11116  0.012909  0.359007 -0.52708  0.779 5.20e-02 0.0784
7  0.513812 -0.095662 -0.01860 -0.233472 -0.439748  0.55898  0.929 6.01e-02 0.1178
8 -0.207257  0.415175 -0.09836 -0.265557  0.326183 -0.57706  0.990 6.45e-02 0.1386
9 -0.102101  0.087976  0.09538  0.117597 -0.019371  0.40915  1.044 3.30e-02 0.1040
10 -0.128423  0.043915  0.09363 -0.063459  0.107928 -0.25637  1.052 1.31e-02 0.0583
11 -0.035269 -0.207181  0.02038  0.307987 -0.069317  0.41394  1.156 3.42e-02 0.1409
12 -0.242999  0.070432  0.03773  0.078418  0.190237 -0.27279  1.102 1.49e-02 0.0771
13  0.010903 -0.004829 -0.00258 -0.001497 -0.008543  0.01260  1.256 3.28e-05 0.0714
14 -0.233454  0.151581  0.09066 -0.027297  0.171728 -0.30049  1.010 1.79e-02 0.0622
15  0.076411 -0.017628  0.10543 -0.152278 -0.083121 -0.25189  1.293 1.29e-02 0.1478
16  0.381093  0.203976 -0.46494 -0.185337 -0.080149  0.70496  0.761 9.18e-02 0.1175
17 -0.178462 -0.142390  0.23481  0.260684 -0.014612  0.47950  1.213 4.59e-02 0.1807
18 -0.006902 -0.303369 -0.23982  0.342520  0.051347 -0.39196  4.649 3.16e-02 0.7511 *
19 -0.123107 -0.017516 -0.16396  0.163396  0.140520 -0.28452  1.092 1.62e-02 0.0780
20  0.193951 -0.144597 -0.29155  0.108601 -0.037300  0.35292  1.163 2.50e-02 0.1241
21 -0.169561 -0.044228  0.21752  0.142727  0.025824  0.36066  1.212 2.62e-02 0.1451
22  0.402256 -0.403052  0.35531 -0.308012 -0.509490 -0.98889  0.514 1.67e-01 0.1269 *
23 -0.011646 -0.000531 -0.02737  0.019927  0.013667 -0.04692  1.225 4.53e-04 0.0535
24 -0.022162 -0.003058  0.02296  0.011805  0.002711 -0.03597  1.325 2.67e-04 0.1211
25 -0.131969  0.591772 -0.55491 -0.093927  0.473884  0.98520  0.884 1.80e-01 0.2165
```

Mínimos Cuadrados Generalizados

Suponga el modelo $y = X\beta + \epsilon$ con

$$\mathbb{E}[\epsilon] = 0, \quad y \quad \text{var}(\epsilon) = \Sigma = \sigma^2 V \neq \sigma^2 I$$

En este caso, utilizar mínimos cuadrados ordinarios (MCO) no es lo apropiado pues las *condiciones de Gauss-Markov* no se cumplen. (Piense en la variación de observaciones o pesos de las observaciones.)

Puesto que $V = \frac{1}{\sigma^2} \text{var}(\epsilon)$, podemos suponer que la matriz V es definida positiva. Entonces, existe $R_{n \times n}$ no singular y simétrica tal que $V = R'R$.

Defina:

$$\begin{aligned} w &= R^{-1}y \\ Z &= R^{-1}X \\ \delta &= R^{-1}\epsilon \end{aligned}$$

Luego,

$$\begin{aligned} y &= X\beta + \epsilon \\ R^{-1}y &= R^{-1}X\beta + R^{-1}\epsilon \\ w &= Z\beta + \delta \end{aligned}$$

Mínimos Cuadrados Generalizados

Considere el modelo:

$$w = Z\beta + \delta$$

donde:

$$\begin{aligned} \mathbb{E}[\delta] &= R^{-1}\mathbb{E}[\epsilon] = 0 \\ \text{cov}(\delta) &= R^{-1}\text{cov}(\epsilon)R^{-1'} \\ &= \sigma^2 R^{-1}VR^{-1'} \\ &= \sigma^2 R^{-1}(R'R)R^{-1'} \\ &= \sigma^2 I_n \end{aligned}$$

Esto es, el modelo $w = Z\beta + \delta$ satisface las *condiciones de Gauss-Markov*, por lo que el estimador $\hat{\beta}_G$ de mínimos cuadrados es el de mínima varianza entre todos los estimadores lineales insesgados de β . A saber,

$$\begin{aligned} \hat{\beta} &= (Z'Z)^{-1}Z'w \\ &= (X'R^{-1'}R^{-1}X)^{-1}X'R^{-1'}R^{-1}y \\ &= (X'V^{-1}X)^{-1}X'V^{-1}y \\ &= \hat{\beta}_G \end{aligned}$$

$\hat{\beta}_G$ es el *estimador de mínimos cuadrados generalizados (MCG)*.

Mínimos Cuadrados Generalizados

$$\hat{\beta}_G = (X' V^{-1} X)^{-1} X' V^{-1} y$$

entonces,

$$\begin{aligned} \mathbb{E}[\hat{\beta}_G] &= \mathbb{E} \left[(X' V^{-1} X)^{-1} X' V^{-1} y \right] \\ &= \mathbb{E} \left[(X' V^{-1} X)^{-1} X' V^{-1} (X\beta + \epsilon) \right] \\ &= \beta + \mathbf{A} \mathbb{E}[\epsilon] \\ &= \beta \end{aligned}$$

$$\begin{aligned} \text{cov}(\hat{\beta}_G) &= \sigma^2 (Z' Z)^{-1} \\ &= \sigma^2 (X' R^{-1'} R^{-1} X)^{-1} \\ &= \sigma^2 (X' V^{-1} X)^{-1} \end{aligned}$$

Por *Gauss-Markov*, $\hat{\beta}_G$ es el mejor estimador lineal insesgado del vector de parámetros β .

Nota: Si $\text{cov}(\epsilon) = \Sigma$, y utilizamos MCO $\hat{\beta} = (X' X)^{-1} X' y$, el estimador $\hat{\beta}$ es aún insesgado pero no de varianza mínima, pues

$$\text{var}(\hat{\beta}) = \sigma^2 (X' X)^{-1} X' V X (X' X)^{-1}$$

Mínimos Cuadrados Ponderados

Cuando los errores no están correlacionados pero tienen distinta varianza, éstos corresponden a un caso particular de *mínimos cuadrados generalizados*. A saber, si

$$\Sigma = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_n^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1/w_1 & & \\ & \ddots & \\ & & 1/w_n \end{bmatrix} = \sigma^2 W^{-1}$$

donde

$$W = \text{diag}\{w_1, \dots, w_n\}$$

Si $\text{var}(\epsilon_j)$ es *grande*, entonces $w_j (> 0)$ es pequeño y a la correspondiente observación se le da *poco peso*.

Mínimos Cuadrados Ponderados

En el caso de errores no correlacionados, sea

$$R = W^{-1/2} = \text{diag}\{1/\sqrt{w_1}, \dots, 1/\sqrt{w_n}\}$$

$$w = R^{-1}y = \begin{bmatrix} \sqrt{w_1}y_1 \\ \vdots \\ \sqrt{w_n}y_n \end{bmatrix}; \quad Z = R^{-1}X = \begin{bmatrix} \sqrt{w_1} & \sqrt{w_1}x_{11} & \cdots & \sqrt{w_1}x_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{w_n} & \sqrt{w_n}x_{n1} & \cdots & \sqrt{w_n}x_{np} \end{bmatrix}$$

Entonces, el *estimador de mínimos cuadrados ponderados* (MCP)

$$\hat{\beta}_W = \hat{\beta}_G = (X'WX)^{-1}X'Wy$$

es insesgado y tal que

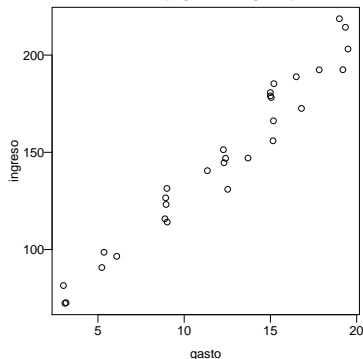
$$\text{cov}(\hat{\beta}_W) = \sigma^2(X'WX)^{-1}$$

Ejemplo: Venta alimentos vs. gastos en publicidad³

La siguiente tabla muestra los datos sobre los ingresos mensuales promedio (Y) por ventas de alimentos y los correspondientes gastos anuales en publicidad (x) para 30 restaurantes. La gerencia está interesada en establecer una relación entre estas variables.

obs	ingreso	gasto	grupo	obs	ingreso	gasto	grupo
1	81.464	3.000	1	16	144.630	12.310	5
2	72.661	3.150	1	17	147.041	13.700	5
3	72.344	3.085	1	18	179.021	15.000	6
4	90.743	5.225	2	19	166.200	15.175	6
5	98.588	5.350	2	20	180.732	14.995	6
6	96.507	6.090	3	21	178.187	15.050	6
7	126.574	8.925	4	22	185.304	15.200	6
8	114.133	9.015	4	23	155.931	15.150	6
9	115.814	8.885	4	24	172.579	16.800	7
10	123.181	8.950	4	25	188.851	16.500	7
11	131.434	9.000	4	26	192.424	17.830	8
12	140.564	11.345	5	27	203.112	19.500	9
13	151.352	12.275	5	28	192.482	19.200	9
14	146.926	12.400	5	29	218.715	19.000	9
15	130.963	12.525	5	30	214.317	19.350	9

Ventas (ingreso vs. gasto)



³Sec. 5.5.5: Montgomery, Peck, and Vining (2001).

Ejemplo: Venta alimentos (cont.)

Regresión lineal simple:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	49.4434	4.2889	11.53	3.81e-12
gasto	8.0484	0.3265	24.65	< 2e-16

Residual standard error: 8.999 on 28 degrees of freedom
Multiple R-squared: 0.9559, Adjusted R-squared: 0.9544
F-statistic: 607.5 on 1 and 28 DF, p-value: < 2.2e-16

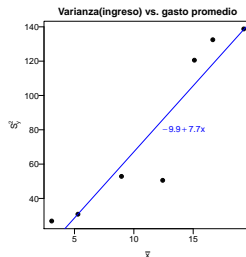
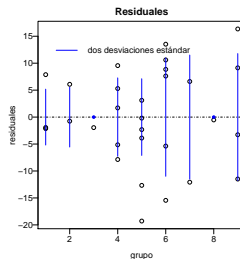
- Problema: Varianza creciente.

Regresión lineal S_Y^2 vs. \bar{x} :

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-9.903	16.764	-0.591	0.58040
xbar	7.703	1.309	5.885	0.00201

Residual standard error: 19.26 on 5 degrees of freedom
(2 observations deleted due to missingness)
Multiple R-squared: 0.8739, Adjusted R-squared: 0.8486
F-statistic: 34.64 on 1 and 5 DF, p-value: 0.002012



Ejemplo: Venta alimentos (cont.)

Mínimos cuadrados ponderados

```
w <- 1/predict(mod2,data.frame(x=dat$gasto))
lm(formula = ingreso ~ gasto, data = dat, weights = w)
```

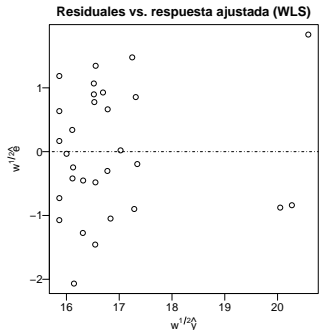
Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	51.0475	2.4095	21.19	<2e-16
gasto	7.9162	0.2503	31.62	<2e-16

Residual standard error: 0.9961 on 28 degrees of freedom

Multiple R-squared: 0.9728, Adjusted R-squared: 0.9718

F-statistic: 1000 on 1 and 28 DF, p-value: < 2.2e-16



Ejemplo: Gastos en educación en los E.E.U.U⁴

La Tabla 1 muestra los datos sobre gasto en educación para el año de 1975 en los E.E.U.U. Note que los estados han sido clasificados en alguna de las 4 regiones definidas: noreste, central, noroeste y sur.

variable	concepto
y	Gasto per cápita para educación.
x_1	Ingreso per cápita en 1973.
x_2	Menores de 18 años por millar en 1974.
x_3	Personas por millar viviendo en zona urbana en 1970.

	estado	y	x_1	x_2	x_3	región
1	ME	235	3944	325	508	1
.
9	PA	300	4894	300	715	1
10	OH	221	5012	324	753	2
.
21	KS	337	5057	304	661	2
22	DE	344	5540	328	722	3
.
37	TX	269	4336	335	797	3
38	MT	302	4418	335	534	4
.
50	HI	311	5309	333	831	4

⁴Sec. 7.4, Chatterjee, Hadi, and Price (2000).

Ejemplo: Gastos en educación en los E.E.U.U (cont.)

Ajuste MCO datos completos:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-5.566e+02	1.232e+02	-4.518	4.34e-05
x1	7.239e-02	1.160e-02	6.239	1.27e-07
x2	1.552e+00	3.147e-01	4.932	1.10e-05
x3	-4.269e-03	5.139e-02	-0.083	0.934

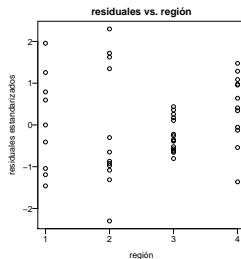
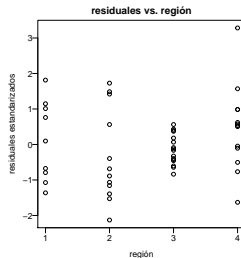
Residual standard error: 40.47 on 46 degrees of freedom
Multiple R-squared: 0.5913, Adjusted R-squared: 0.5647
F-statistic: 22.19 on 3 and 46 DF, p-value: 4.945e-09

Ajuste MCO sin Alaska:

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-277.57731	132.42286	-2.096	0.041724
x1	0.04829	0.01215	3.976	0.000252
x2	0.88693	0.33114	2.678	0.010291
x3	0.06679	0.04934	1.354	0.182591

Residual standard error: 35.81 on 45 degrees of freedom
Multiple R-squared: 0.4967, Adjusted R-squared: 0.4631
F-statistic: 14.8 on 3 and 45 DF, p-value: 7.653e-07



Ejemplo: Gastos en educación (cont.)

Mínimos cuadrados ponderados

Sin considerar Alaska, calculamos la varianza muestral de los residuales por región y usamos su recíproco como ponderación o peso para las observaciones.

R_i	1	2	3	4
s^2	1629.76	2640.89	185.82	810.21

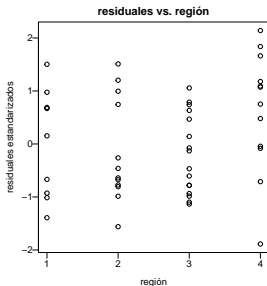
Ponderaciones: $w_i = 1/s_{R_i}^2$

Ajuste MCP datos sin Alaska:

Coefficients:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3.195e+02	7.122e+01	-4.486 4.99e-05
x1	6.372e-02	7.295e-03	8.735 3.01e-11
x2	8.733e-01	1.827e-01	4.781 1.90e-05
x3	2.576e-02	3.179e-02	0.810 0.422

Residual standard error: 1.017 on 45 degrees of freedom
Multiple R-squared: 0.7877, Adjusted R-squared: 0.7735
F-statistic: 55.65 on 3 and 45 DF, p-value: 3.494e-15



Problema de Autocorrelación

Uno de los supuestos de la regresión es la *no correlación* de los errores. Sin embargo, en series temporales o datos espaciales es común tener situaciones donde $\text{cov}(\epsilon_j, \epsilon_{i+k}) \neq 0$. En tal situación, se dice que los errores están *correlacionados serialmente*, o bien, *autocorrelacionados*.

Se define la *autocorrelación de orden k*, a

$$r_k = \text{corr}(\epsilon_j, \epsilon_{i+k}) = \frac{\text{cov}(\epsilon_j, \epsilon_{i+k})}{\text{var}(\epsilon_j)}, \quad k = 1, 2, \dots$$

para todo error ϵ_j , puesto que se supone que $\text{var}(\epsilon_j) = \sigma^2$.

Fuentes de la autocorrelación

- Ocurre cuando se tienen datos *temporales* como son las series económicas, o *espaciales* como lo son los datos geográficos y no se incluyen regresores importantes en el modelo, como aquellos que comparten relación entre observaciones. Por ejemplo, en series temporales quizás se deba incluir explícitamente la variable t *tiempo*; en estadística espacial alguna variable para indicar la cercanía.

Consecuencias de la autocorrelación

- 1 Los estimadores de Mínimos Cuadrados Ordinarios (MCO) siguen siendo insesgados pero no son de varianza mínima como lo garantizan las condiciones de *Gauss-Markov*. Esto es, son ahora *estimadores ineficientes*.
- 2 Si los residuales se correlacionan positivamente, s^2 *subestima* a σ^2 , lo que implica errores estándar pequeños. Esto conduce a indicar como “*significativos*” a términos que en realidad no lo son.

$$t_j = \frac{\hat{\beta}_j}{\text{ee}(\hat{\beta}_j)} = \frac{\hat{\beta}_j}{s\sqrt{C_{jj}}}$$

Error estándar pequeños sugieren *precisión* que en realidad no se tiene.

- 3 Estrictamente, una σ^2 *mal estimada* implica que las pruebas de hipótesis basadas en las distribuciones t y F no son válidas ni apropiadas.

Notas

- 1 Si el problema de la autocorrelación es debido a no haber incluido términos o regresores necesarios, incluirlos en el modelo elimina el problema.
- 2 Si no es posible el caso anterior, habrá que incluir la estructura del error en el modelo y estimar de manera apropiada (*estimación no lineal*).
- 3 La dependencia temporal es frecuente en series *económicas* y *financieras*, que muchas veces son estudiadas con modelos de *series de tiempo*. Por ejemplo, modelos *autorregresivos* y *de promedios móviles* – ARMA(p, q):

$$\Phi(B)y_t = \Theta(B)a_t$$

donde $\Phi(B)$ y $\Theta(B)$ se conocen como *polinomios de retraso* y los $a_t \sim N(0, \sigma_a^2)$ *i.i.d.* El modelo se puede escribir explícitamente como:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + a_t + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$$

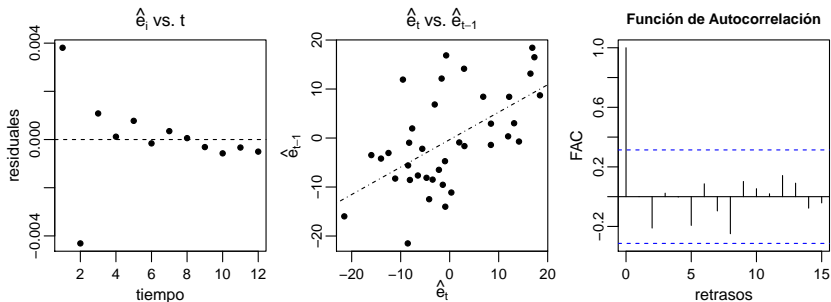
O más generalmente, los modelos ARIMA(p, d, q):

$$\Phi(B)Y_t = \Theta(B)a_t$$

donde $Y_t = (y_t - y_{t-1})^d = \nabla^d y_t$, es la *diferencia de orden d* .

¿Cómo identificar el problema de la autocorrelación?

- 1 Gráfica de residuales $\hat{\epsilon}_t$ con respecto al tiempo t .
- 2 Gráfica de dispersión de la autocorrelación de primer orden ($\hat{\epsilon}_{t-1}$ vs. $\hat{\epsilon}_t$).
- 3 Función de Autorrelación (ACF).



- 4 Prueba formal para la autocorrelación de primer orden en RLM, propuesta por Durbin and Watson (1950, 1951, 1971).

Problema de Autocorrelación

Suponga el modelo de primer orden para el error $\epsilon_t = \rho\epsilon_{t-1} + \mathbf{a}_t$, donde $\mathbf{a}_t \sim N(0, \sigma_a^2 I)$ es *ruído blanco*⁵ y $|\rho| < 1$. Luego, el modelo completo queda:

$$\begin{aligned} y_t &= \mathbf{x}'_t \beta + \epsilon_t \\ \epsilon_t &= \rho\epsilon_{t-1} + \mathbf{a}_t \end{aligned}$$

Ahora bien,

$$\epsilon_t = \rho\epsilon_{t-1} + \mathbf{a}_t = \rho(\rho\epsilon_{t-2} + \mathbf{a}_{t-1}) + \mathbf{a}_t = \dots = \sum_{u=0}^{\infty} \rho^u \mathbf{a}_{t-u}$$

De donde,

$$\begin{aligned} \mathbb{E}[\epsilon_t] &= 0 \\ \text{var}(\epsilon_t) &= \text{var}\left(\sum_{u=0}^{\infty} \rho^u \mathbf{a}_{t-u}\right) = \sigma_a^2 \frac{1}{1 - \rho^2} \\ \text{cov}(\epsilon_t, \epsilon_{t-k}) &= \rho^{|k|} \sigma_a^2 \frac{1}{1 - \rho^2} \end{aligned}$$

Entonces los ϵ_t tiene media cero pero están correlacionados a menos que $\rho = 0$.

⁵Variables aleatorias independientes, idénticamente distribuidas

Estadístico d de Durbin-Watson

Puesto que el problema más común es la correlación positiva, se contrasta

$$H_0 : \rho = 0 \quad \text{vs.} \quad H_a : \rho > 0$$

Estadístico *Durbin-Watson*:

$$d = \frac{\sum_{t=2}^n (\hat{\epsilon}_t - \hat{\epsilon}_{t-1})^2}{\sum_{t=1}^n \hat{\epsilon}_t^2}$$

La distribución del estadístico d depende del tamaño de la muestra y de la matriz de regresión X . Asimismo, $d \in (0, 2)$ donde 0 indica autocorrelación.

- Por tanto, valores pequeños de d implican que $\hat{\epsilon}_t$ y $\hat{\epsilon}_{t-1}$ son parecidos, por lo que H_0 será rechazado.
- Para contrastar autocorrelación negativa ($H_0 : \rho = 0$ vs. $H_1 : \rho < 0$), se toma $d' = 4 - d$ y procedemos como antes.

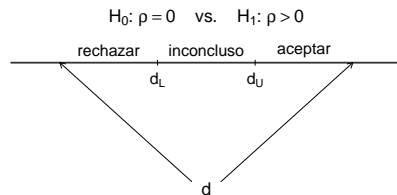
Estadístico d de Durbin-Watson

Durbin y Watson mostraron que existen cotas o límites independientes de los datos X tales que, con

$$d = \frac{\sum_{t=2}^n (\hat{e}_t - \hat{e}_{t-1})^2}{\sum_{t=1}^n \hat{e}_t^2}$$

si

- $d < d_L \implies$ Rechazar $H_0 : \rho = 0$
 $d > d_U \implies$ Aceptar $H_0 : \rho = 0$
 $d_L < d < d_U \implies$ Prueba inconclusa



Los límites d_L y d_U dependen del número de observaciones n , el número de regresores p y la significancia de la prueba α .

Tablas de Durbin y Watson⁶

J. DURBIN AND G. S. WATSON

173

The weight function was chosen to be the density of the Beta-distribution with the correct mean and variance. (An alternative weight function giving the right order of vanishing of (17) at $x = 0$ and $x = 1$ was also tried but it was found to be less satisfactory.) With this weight function, the coefficients a_1 and a_2 in (17) were zero. Terms as far as $G_4(x)$ were used.

Table 4. Significance points of d_L and d_U : 5 %

n	k' = 1		k' = 2		k' = 3		k' = 4		k' = 5	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83

⁶Durbin and Watson (1950,1951)

Tablas del estadístico de Durbin y Watson

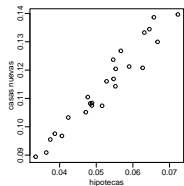
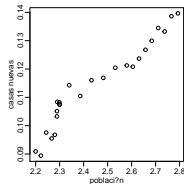
Tabla 10A.1 Cotas críticas $dL_{(\alpha;n,k)}$, $dU_{(\alpha;n,k)}$ del estadístico de Durbin-Watson.

n	α = 0.01																					
	k = 1		k = 2		k = 3		k = 4		k = 5		k = 6		k = 7		k = 8		k = 9		k = 10		k = 11	
	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU	dL	dU
6	0.390	1.142																				
7	0.435	1.036	0.294	1.676																		
8	0.488	1.004	0.346	1.489	0.229	2.102																
9	0.554	0.998	0.409	1.390	0.279	1.875	0.183	2.433														
10	0.604	1.000	0.466	1.332	0.340	1.734	0.230	2.194	0.150	2.690												
11	0.653	1.010	0.519	1.297	0.396	1.641	0.286	2.029	0.193	2.453	0.125	2.893										
12	0.698	1.023	0.570	1.274	0.448	1.574	0.339	1.912	0.244	2.280	0.164	2.664	0.105	3.053								
13	0.739	1.039	0.616	1.261	0.499	1.526	0.391	1.856	0.294	2.149	0.211	2.490	0.140	2.838	0.090	3.182						
14	0.776	1.054	0.660	1.254	0.547	1.491	0.441	1.757	0.343	2.049	0.257	2.354	0.183	2.667	0.122	2.982	0.078	3.287				
15	0.811	1.071	0.700	1.251	0.592	1.465	0.487	1.704	0.391	1.967	0.303	2.244	0.226	2.530	0.161	2.817	0.107	3.101	0.068	3.374		
16	0.844	1.087	0.738	1.253	0.632	1.445	0.532	1.663	0.437	1.900	0.349	2.153	0.269	2.415	0.200	2.681	0.142	2.944	0.094	3.201	0.060	3.446
17	0.874	1.103	0.773	1.256	0.672	1.433	0.574	1.631	0.480	1.846	0.393	2.078	0.313	2.319	0.241	2.566	0.179	2.811	0.127	3.053	0.084	3.286
18	0.902	1.118	0.805	1.260	0.708	1.423	0.613	1.604	0.522	1.803	0.435	2.016	0.355	2.237	0.283	2.466	0.217	2.697	0.160	2.954	0.114	3.146
19	0.928	1.132	0.835	1.265	0.742	1.416	0.650	1.584	0.561	1.767	0.476	1.963	0.396	2.169	0.322	2.382	0.255	2.597	0.195	2.813	0.144	3.023
20	0.953	1.147	0.863	1.270	0.773	1.410	0.685	1.567	0.598	1.737	0.515	1.919	0.436	2.108	0.362	2.307	0.293	2.510	0.232	2.714	0.178	2.915
21	0.976	1.161	0.890	1.278	0.803	1.408	0.718	1.554	0.634	1.712	0.552	1.881	0.474	2.060	0.400	2.244	0.332	2.434	0.268	2.625	0.212	2.816
22	0.997	1.174	0.915	1.284	0.832	1.408	0.748	1.543	0.667	1.691	0.587	1.848	0.510	2.015	0.437	2.189	0.369	2.367	0.305	2.547	0.246	2.729
23	1.017	1.186	0.938	1.290	0.857	1.406	0.778	1.535	0.698	1.673	0.620	1.821	0.545	1.978	0.473	2.140	0.404	2.308	0.340	2.479	0.281	2.650
24	1.037	1.199	0.960	1.298	0.882	1.407	0.804	1.528	0.728	1.658	0.652	1.797	0.578	1.944	0.507	2.098	0.439	2.255	0.375	2.417	0.315	2.580
25	1.055	1.211	0.981	1.305	0.906	1.409	0.831	1.523	0.756	1.646	0.682	1.777	0.610	1.915	0.540	2.060	0.473	2.209	0.409	2.362	0.348	2.518
26	1.072	1.222	1.001	1.312	0.928	1.410	0.855	1.519	0.783	1.635	0.711	1.759	0.641	1.890	0.571	2.026	0.505	2.168	0.441	2.313	0.381	2.461
27	1.088	1.233	1.019	1.318	0.950	1.413	0.879	1.516	0.808	1.626	0.738	1.744	0.669	1.867	0.603	1.997	0.536	2.131	0.473	2.269	0.413	2.410
28	1.104	1.244	1.037	1.325	0.969	1.415	0.900	1.513	0.832	1.618	0.764	1.739	0.696	1.847	0.631	1.971	0.566	2.097	0.504	2.239	0.444	2.363
29	1.119	1.254	1.054	1.332	0.988	1.418	0.922	1.512	0.855	1.611	0.788	1.717	0.723	1.829	0.658	1.947	0.595	2.068	0.534	2.193	0.474	2.321
30	1.133	1.264	1.070	1.339	1.006	1.422	0.942	1.510	0.877	1.606	0.812	1.707	0.748	1.814	0.684	1.925	0.622	2.041	0.562	2.160	0.503	2.282

★ Ejemplo: Construcción de casas habitación⁷

La Tabla muestra los datos sobre nuevas construcciones de casas habitación (N) en los E.E.U.U. La tabla presenta también la población (P) en millones de personas y un índice de hipotecas disponibles (H).

obs	N	P	H	obs	N	P	H
1	0.09090	2.200	0.03635	13	0.11048	2.386	0.04770
2	0.08942	2.222	0.03345	14	0.11604	2.433	0.05282
3	0.09755	2.244	0.03870	15	0.11688	2.482	0.05473
4	0.09550	2.267	0.03745	16	0.12044	2.532	0.05531
5	0.09678	2.280	0.04063	17	0.12125	2.580	0.05898
6	0.10327	2.289	0.04237	18	0.12080	2.605	0.06267
7	0.10513	2.289	0.04715	19	0.12368	2.631	0.05462
8	0.10840	2.290	0.04883	20	0.12679	2.658	0.05672
9	0.10822	2.299	0.04836	21	0.12996	2.684	0.06674
10	0.10741	2.300	0.05160	22	0.13445	2.711	0.06451
11	0.10751	2.300	0.04879	23	0.13325	2.738	0.06313
12	0.11429	2.340	0.05523	24	0.13863	2.766	0.06573
				25	0.13964	2.793	0.07229



⁷Sec. 7.8, Chatterjee, Hadi, and Price (2000).

Ejemplo: Construcción de casas habitación (cont.)

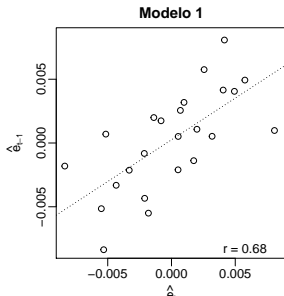
Modelo 1:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.060884	0.010416	-5.845 5.89e-06
P	0.071410	0.004234	16.867 1.91e-14

Residual standard error: 0.00408 on 23 degrees of freedom
 Multiple R-squared: 0.9252, Adjusted R-squared: 0.922
 F-statistic: 284.5 on 1 and 23 DF, p-value: 1.911e-14

lag	Autocorrelation	D-W Statistic	p-value
1	0.6511468	0.6208403	0

Alternative hypothesis: rho != 0



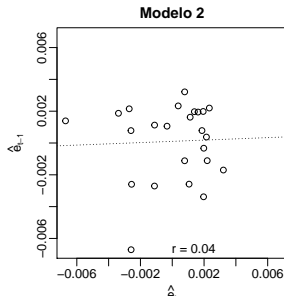
Modelo 2:

Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.010427	0.010291	-1.013 0.322
P	0.034656	0.006425	5.394 2.04e-05
H	0.760464	0.121588	6.254 2.70e-06

Residual standard error: 0.002503 on 22 degrees of freedom
 Multiple R-squared: 0.9731, Adjusted R-squared: 0.9706
 F-statistic: 397.6 on 2 and 22 DF, p-value: < 2.2e-16

lag	Autocorrelation	D-W Statistic	p-value
1	0.03957229	1.852409	0.462

Alternative hypothesis: rho != 0



Ejemplo: Pronóstico de un Índice Financiero⁸

133

Lagged Relationships in Economic Forecasting

By P. J. COEN, E. D. GOMME and M. G. KENDALL

Scientific Control Systems Ltd, London

[Read before the ROYAL STATISTICAL SOCIETY on Wednesday, January 15th, 1969,
the President, Mr F. A. COCKFIELD, in the Chair]

INTRODUCTION

1. FOR some years econometricians have been aware of the critical dependence of economic systems on the lag between cause and effect. It has been shown that the very stability of a system may depend on the length of time taken for a stimulus to pass through the transient to full response. However, not much is known about the actual lengths of these lags, at least in the U.K. economy, and surprisingly little effort has been devoted to determining them. Such writers as are concerned with dynamic systems usually write in lags of one or two time units (and occasionally up to four in quarterly data to describe seasonal effects) and hope that autoregression will take care of any important lag effects. We felt that work in this area being done, for example, in the U.S.A. (e.g. by the National Bureau of Economic Research in *Business Cycle Indicators*) might not be applicable in the U.K. and decided to approach the British problem *ab initio*.

⁸Coen, Gomme, and Kendall (1969).

Ejemplo: Pronóstico de un Índice Financiero (cont.)

TABLE 2
*Coefficients in regressions of F.T. ordinary share index
on the three variables shown*

Equation No.	Period of fit (year and quarter)	Constant	Coefficients			R ²
			(U.K. car production) _{t-6}	F.T. (commodity index) _{t-7}	(Reuter's commodity index) _{t-7}	
(1)	1948/1-1958/1	82.06	0.00063 (10.34)		-0.05 (-1.54)	0.75
(2)	1954/2-1960/4	750.96	0.00039 (2.99)	-7.07 (-5.65)		0.81
(3)	1954/2-1962/1	757.64	0.00041 (4.89)	-7.16 (-5.96)		0.85
(4)	1954/2-1964/1	716.87	0.00042 (5.83)	-6.74 (-7.66)		0.88
(5)	1954/2-1965/1	717.88	0.00044 (7.30)	-6.78 (-8.42)		0.90
(6)	1958/1-1965/4	897.51	0.00035 (7.879)	-8.675 (-8.49)		0.87
(7)	1954/2-1966/4	653.16	0.00047 (14.114)	-6.127 (-9.89)		0.90

Values of t are in brackets under the coefficients.

$$y_t = \alpha + \beta_1 x_{1,t-6} + \beta_2 x_{2,t-7} + a_t$$

Ejemplo: Pronóstico de un Índice Financiero⁹

1971]

229

Some Comments on a Paper of Coen, Gomme and Kendall

By GEORGE E. P. BOX¹ and PAUL NEWBOLD²

University of Wisconsin

[Received July 1970. Revised February 1971]

SUMMARY

The method of analysis used in a recent paper on economic forecasting is reviewed. Evidence is presented that what were believed to be highly significant relationships making possible the forecasting of the *Financial Times* share index arise because of the inflexibility of the assumed error structure.

1. INTRODUCTION

In a recent publication, by Coen, Gomme and Kendall (1969) which for convenience we refer to as the C.G.K. paper, the forecasting of the *Financial Times* ordinary share index using various other lagged series is discussed. It is sufficient for illustration to consider relation (7) of the C.G.K. paper which we write as

$$Y_t = \alpha + \beta_1 X_{1,t-6} + \beta_2 X_{2,t-1} + n_t \quad (1)$$

which is projected to obtain forecasts. In this expression the “output” Y_t is the *Financial Times* ordinary share index. The two “inputs” are $X_{1,t}$, United Kingdom car production, and $X_{2,t}$, the *Financial Times* commodity index, and n_t is an error term. Quarterly data were employed yielding time series containing 51 successive observations. Of course this is only one of a number of such relationships which the authors postulate. However, it is their methods which we are doubtful about and our reservations would apply equally to their other analyses.

⁹The authors built their model (1) by cross-correlating the detrended series which in

⁹Box and Newbold (1971).

Ejemplo: Pronóstico de un Índice Financiero (cont.)

$$y_t = \alpha + \beta_0 t + \beta_1 X_{1,t-6} + \beta_2 X_{2,t-7} + \eta_t$$

TABLE 1

Estimates and standard errors of coefficients in equation (3) for various noise structures

Assumed structure for noise in Y_t	White noise (C.G.K.)	Integrated moving average	Second-order autoregressive	First-order autoregressive	Random walk
	$n_t = a_t$	$n_t = a_t + (1 - \theta) \sum_{j=1}^{\infty} a_{t-j}$	$n_t = \phi_1 n_{t-1} + \phi_2 n_{t-2} + a_t$	$n_t = \phi_1 n_{t-1} + a_t$	$n_t = \sum_{j=0}^{\infty} a_{t-j}$
α	653 ± 57		306 ± 108	318 ± 106	
β_0		1.78 ± 2.7	2.31 ± 1.0	2.04 ± 1.1	1.74 ± 2.6
β_1	0.00047 ± 0.00004	0.00016 ± 0.00009	0.00017 ± 0.00009	0.00018 ± 0.00009	0.00017 ± 0.00008
β_2	-6.13 ± 0.62	-1.16 ± 1.18	-1.76 ± 1.22	-1.87 ± 1.18	-1.27 ± 1.17
θ		-0.06 ± 0.15			
ϕ_1			0.93 ± 0.16	0.82 ± 0.10	
ϕ_2			-0.14 ± 0.16		
σ_a^2	497	321	299	298	315
Durbin-Watson statistic	Significant at 1 per cent	Not significant	Not significant	Not significant	Not significant

De Coen et al. (1969) 1954/2–1966/4, ecuación (7)

Corrección de Cochran–Orcutt

Considere el modelo de regresión lineal simple

$$y_t = \beta_0 + \beta_1 x_t + \epsilon_t \quad \text{donde} \quad \epsilon_t = \rho \epsilon_{t-1} + a_t$$

y suponga que rechaza la *hipótesis nula* $H_0 : \rho = 0$ en favor de la *hipótesis alternativa* $H_1 : \rho > 0$.

Entonces, considere

$$\begin{aligned} Y_t &= y_t - \rho y_{t-1} \\ &= (\beta_0 + \beta_1 x_t + \epsilon_t) - \rho (\beta_0 + \beta_1 x_{t-1} + \epsilon_{t-1}) \\ &= \beta_0(1 - \rho) + \beta_1(x_t - \rho x_{t-1}) + (\epsilon_t - \rho \epsilon_{t-1}) \\ Y_t &= \alpha_0 + \alpha_1 X_t + a_t \end{aligned}$$

que sí satisface los supuestos usuales de un modelo de regresión lineal. El problema ahora es que Y_t y X_t dependen de ρ , que en general es desconocido.

Método iterativo:

$$\epsilon_t = \rho\epsilon_{t-1} + a_t \xrightarrow{\text{MC}} \hat{\rho} = \frac{\sum_{t=2}^n \hat{\epsilon}_t \hat{\epsilon}_{t-1}}{\sum_{t=1}^n \hat{\epsilon}_t^2}$$

Utilice $\hat{\rho}$ en la definición de Y_t y X_t , estime mediante *mínimos cuadrados ordinarios* (MCO)

$$Y_t = \alpha_0 + \alpha_1 X_t + a_t$$

y valide el modelo. Si los residuales aún están correlacionados itere una vez más. Note que el problema de estimación del modelo corregido es *no lineal*. A saber,

$$\min_{\beta_0, \beta_1, \rho} \sum [y_t - \rho y_{t-1} - \beta_0(1 - \rho) - \beta_1(x_t - \rho x_{t-1})]^2$$

Ejemplo: Pronóstico de un Índice Financiero (cont.)

La Tabla muestra parte de los datos sobre índice financiero trimestral del *Financial Times* (y_t), la producción de automóviles (x_{1t}) y el índice de precios de artículos (x_{2t}). (En realidad las variables x_1 y x_2 han sido desfasadas 6 y 7 periodos respectivamente como lo indica el artículo original de Coen, Gomme y Kendall (1969).)

y = índice financiero de Financial Times al tiempo t .

x_1 = producción de automóviles al tiempo $t - 6$.

x_2 = índice de precios de artículos al tiempo $t - 7$.

obs	y	x_1	x_2
1	148.5	121874	96.21
2	165.4	126260	93.74
3	178.5	145248	91.37
4	187.3	160370	86.31
5	195.9	163648	84.98
6	205.2	178195	86.46
7	191.5	187197	90.04
8	183.0	195916	94.74
9	184.5	199253	92.43
10	181.1	227616	92.41
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮



Ejemplo: Pronóstico de un Índice Financiero (cont.)

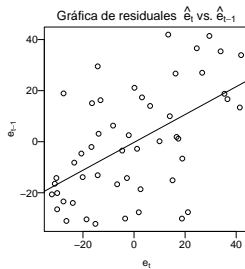
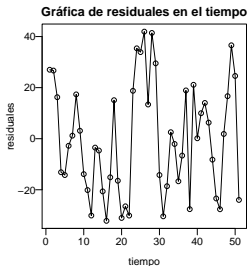
Modelo:

$$y_t = \beta_0 + \beta_1 X_{1,t-6} + \beta_2 X_{2,t-7} + \epsilon_t$$

Coefficients:

Estimate	Std. Error	t value	Pr(> t)		Durbin-Watson test:
(Intercept)	6.532e+02	5.653e+01	11.554	1.82e-15	Lag Autocorrelation D-W Statistic p-value
x1	4.747e-04	3.363e-05	14.114	< 2e-16	1 0.4828553 0.9796219 0
x2	-6.127e+00	6.199e-01	-9.885	3.70e-13	Alternative hypothesis: rho != 0

Residual standard error: 22.29 on 48 degrees of freedom
 Multiple R-Squared: 0.9018, Adjusted R-squared: 0.8977
 F-statistic: 220.5 on 2 and 48 DF, p-value: < 2.2e-16



Ejemplo: Pronóstico de un Índice Financiero (cont.)

Corrección de Cochran-Orcutt:

$$(y_t - \hat{\rho}y_{t-1}) = \beta_0(1 - \hat{\rho}) + \beta_1(x_{1,t-6} - \hat{\rho}x_{1,t-7}) + \beta_2(x_{2,t-7} - \hat{\rho}x_{2,t-8}) + (\epsilon_t - \rho\epsilon_{t-1})$$

$$Y_t = \alpha_0 + \alpha_1X_{1,t} + \alpha_2X_{2,t} + a_t$$

Coefficients:

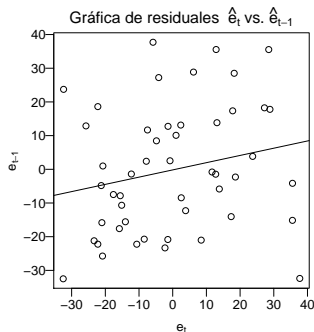
Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3.131e+02	4.218e+01	7.423 1.87e-09
X1	4.490e-04	5.004e-05	8.972 9.41e-12
X2	-5.494e+00	9.233e-01	-5.950 3.20e-07

Residual standard error: 19.07 on 47 degrees of freedom
 Multiple R-Squared: 0.7416, Adjusted R-squared: 0.7306
 F-statistic: 67.46 on 2 and 47 DF, p-value: 1.539e-14

Durbin-Watson test:

Lag	Autocorrelation	D-W Statistic	p-value
1	0.1407426	1.650757	0.14

Alternative hypothesis: rho != 0



Ejemplo: Pronóstico de un Índice Financiero (cont.)

Estimación mediante procedimiento no lineal del modelo con error autocorregresivo AR(2)

$$y_t = \alpha + \beta_0 t + \beta_1 x_{1,t-6} + \beta_2 x_{2,t-7} + \epsilon_t \quad \text{con} \quad \epsilon_t = \phi_1 \epsilon_{t-1} + \phi_2 \epsilon_{t-2} + a_t$$

```
library(nlme)
mod20 <- gls(y ~ t+x1+x2, correlation=corARMA(p=2,q=0), data=dat)
print(summary(mod20))
```

Generalized least squares fit by REML

Model: y ~ t + x1 + x2

Data: dat

AIC	BIC	logLik
451.109	464.06	-218.5545

Correlation Structure: ARMA(2,0)

Formula: ~1

Parameter estimate(s):

Phi1	Phi2
0.9938832	-0.1000777

Coefficients:

	Value	Std. Error	t-value	p-value
(Intercept)	268.44716	112.14436	2.393764	0.0207
t	2.28648	1.20227	1.901800	0.0633
x1	0.00016	0.00008	1.933058	0.0593
x2	-1.36845	1.17484	-1.164804	0.2500

Correlation:

(Intr)	t	x1	
t	-0.409		
x1	0.023	-0.448	
x2	-0.950	0.246	-0.117

Standardized residuals:

Min	Q1	Med	Q3	Max
-1.2333119	-0.2599283	0.1147286	0.6418843	1.7014727

Residual standard error: 40.10971

Degrees of freedom: 51 total; 47 residual

Colinealidad

Modelo:

$$y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip} + \epsilon_i$$

donde,

- y_i : respuesta al nivel i -ésimo nivel del regresor ($i = 1, \dots, n$)
- x_{ij} : j -ésimo regresor a su i -ésimo nivel ($j = 1, \dots, p$)
- β_0 : ordenada al origen
- β_j : coeficiente (tasa de cambio) del j -ésimo regresor

O bien, matricialmente

$$y = X\beta + \epsilon$$

Supuestos:

$$\epsilon \sim N(0, \sigma^2 I_n)$$

$$\text{rango}(X) = p + 1 = q (\leq n)$$

Luego, $(X'X)$ es una matriz invertible.

Si la matriz es “*cercanamente singular*” (columnas “*casi dependientes*”) se dice que se tiene el problema de *colinealidad*. Si $(X'X)$ es singular se dice que se tiene *colinealidad perfecta*.

- El problema de colinealidad implica *sensibilidad numérica* (números de condición muy altos).

Fuentes principales de la colinealidad.

- El método empleado en la recolección de datos. E. g., muestras en regiones limitadas que hacen que los casos sean parecidos. Por ejemplo, conjunto habitacional de un sindicato.
- Restricciones en el modelo o población. E. g., las casas habitación grandes están asociadas a altos ingresos. O bien, cuando se incluye un *índice* y la mayoría de sus componentes.
- Especificación del modelo. E. g., polinomios con x cercanos a 0, entonces x y x^2 se parecen mucho. Polinomios de grado alto tienden a ser colineales.
- Sobre-especificación del modelo. E. g., incluir un número índice y *todos* sus componentes. O bien, incluir k variables indicadoras para representar variables categóricas con k clases.

Efectos de la colinealidad.

Potencialmente, las consecuencias de la presencia de colinealidad son muchas. E. g., si se considera el modelo

$$y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

entonces las ecuaciones normales pueden escribirse como

$$\begin{bmatrix} 1 & r_{12} \\ r_{21} & 1 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} r_{y1} \\ r_{y2} \end{bmatrix}$$

con $r_{12} = \text{corr}(x_1, x_2)$, $r_{yi} = \text{corr}(y, x_i)$. Entonces,

$$C = (X'X)^{-1} = \frac{1}{1 - r_{12}^2} \begin{bmatrix} 1 & -r_{12} \\ -r_{12} & 1 \end{bmatrix}$$

Por lo que

$$\hat{\beta}_i = \frac{r_{yi} - r_{12}r_{yj}}{1 - r_{12}^2}, \quad i = 1, 2 \neq j = 1, 2$$

Por lo que si x_1 y x_2 están correlacionados, $r_{12}^2 \nearrow 1 \implies \hat{\beta}_i \nearrow \infty$.

Efectos de la colinealidad.

En el caso general de p regresores, Marquardt (1970) mostró que

$$C_{jj} = \frac{1}{1 - R_j^2}, \quad j = 1, \dots, p$$

donde R_j^2 es el *coeficiente de determinación múltiple* al explicar el j -ésimo regresor X_j en función de las otras $p - 1$ variables independientes. Una R_j^2 cercana a 1 implica un elemento C_{jj} muy grande, por lo que

$$\text{var}(\hat{\beta}_j) = \sigma^2 C_{jj} \nearrow \infty$$

Lo que hace muy incierta la estimación de β_j .

También, se puede mostrar que

$$\mathbb{E} [\|\hat{\beta} - \beta\|^2] = \sigma^2 \sum_{i=1}^p \frac{1}{\lambda_i}$$

donde los λ_i son los valores propios de la matriz $(X'X)$. Entonces, si existe el problema de colinealidad, algunos de los λ_i serán cercanos a 0, lo que significa que $\hat{\beta}$ y β estarán muy alejados en algunas de sus entradas.

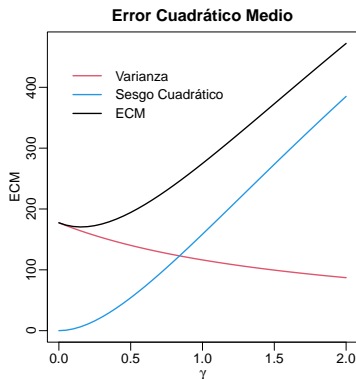
Notas

Si se cumplen las condiciones de *Gauss-Markov*, el estimador $\hat{\beta}$ es el de varianza mínima entre la clase de estimadores insesgados. Pero si se amplía la clase (permitiendo estimadores sesgados) se pueden obtener estimadores de menor varianza que el de *mínimos cuadrados ordinarios (MCO)*.

Existe un compromiso entre insesgamiento y varianza:

$$ECM(\hat{\beta}) = \text{var}(\hat{\beta}) + \text{Sesgo}^2(\hat{\beta})$$

Hay una vecindad donde el parámetro cresta o *ridge* es más eficiente que el de MCO.



Regresión risco (*ridge*)

Las ecuaciones normales se modifican de manera que para $\gamma \geq 0$

$$\begin{aligned}\hat{\beta}_\gamma &= (X'X + \gamma I)^{-1} X'y \\ &= (X'X + \gamma I)^{-1} (X'X) \hat{\beta} \\ &= Z_\gamma \hat{\beta}\end{aligned}$$

Esto es, el estimador cordillera es la transformación lineal $Z_\gamma = (X'X + \gamma I)^{-1} (X'X)$ del estimador de mínimos cuadrados ordinarios (MCO) $\hat{\beta}$. Se tiene además

$$\begin{aligned}\text{var}(\hat{\beta}_\gamma) &= \sigma^2 (X'X + \gamma I)^{-1} X'X (X'X + \gamma I)^{-1} \\ \text{ECM}(\hat{\beta}_\gamma) &= \sigma^2 \sum_{j=1}^q \frac{\lambda_j}{(\lambda_j + \gamma)^2} + \gamma^2 \beta' (X'X + \gamma I)^{-2} \beta\end{aligned}$$

Note que conforme crece γ el estimador $\hat{\beta}_\gamma$ se hace más estable pero también más sesgado.

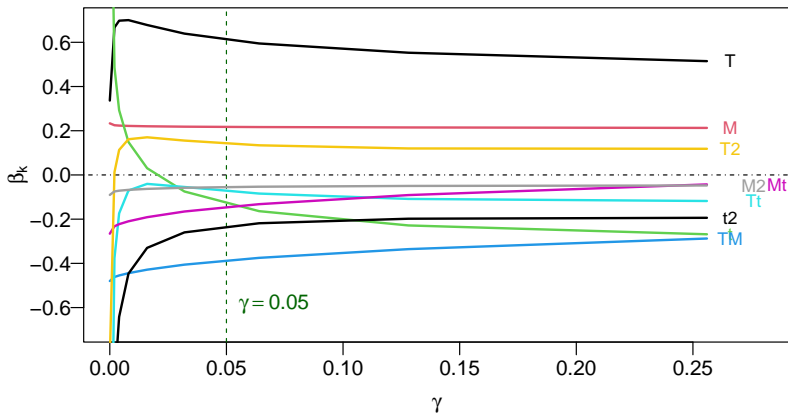
¿Qué γ usar?

Regresión riesgo: selección del parámetro de sesgo γ ¹⁰

GCV Golub et al. (1979)	Validación cruzada generalizada (datos escalados)	$\gamma = \arg \min \left[\frac{\ I - A_\gamma\ }{\text{tr}(I - A_\gamma)} \right]^2$ donde: $A_\gamma = X(X'X + \gamma I)^{-1}X'$	Buen estimador pero intensivo numéricamente.
MHKB: Hoerl et al. (1975)	Interpretación bayesiana	$\gamma = (p - 2) \frac{\sigma^2}{\hat{\beta}'\hat{\beta}}$	No recomendado en casos de colinealidad fuerte.
MLW: Lawless and Wang (1976)	Interpretación bayesiana (y y X escalados)	$\gamma = \frac{p - 2}{p} \cdot \frac{\text{tr}(X'X)}{\hat{y}'\hat{y}} \sigma^2$	Recomendado con valores propios pequeños.

¹⁰Brown (1993); Golub et al. (1979); Hoerl et al. (1975); Lawless and Wang (1976)

Traza cresta



Utilizar aquella γ que "estabilice" las estimaciones.

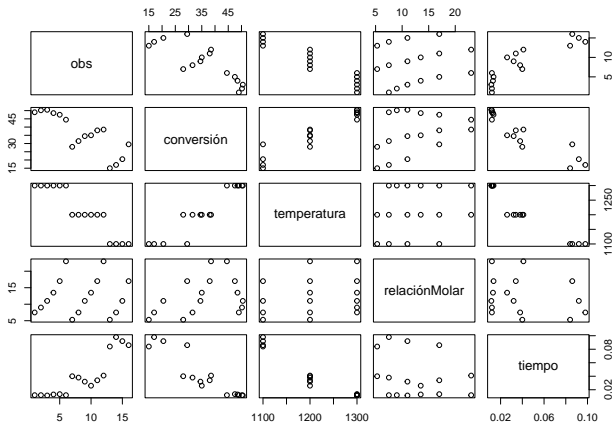
Ejemplo: Acetileno¹¹

Se desea modelar la cantidad convertida (y) en un reactor como un modelo de segundo orden en términos de la temperatura (T), la relación molar (M) y el tiempo (t).

obs	conversión	temperatura	relMolar	tiempo
i	C	T	M	t
1	49.0	1300	7.5	0.0120
2	50.2	1300	9.0	0.0120
3	50.5	1300	11.0	0.0115
4	48.5	1300	13.5	0.0130
5	47.5	1300	17.0	0.0135
6	44.5	1300	23.0	0.0120
7	28.0	1200	5.3	0.0400
8	31.5	1200	7.5	0.0380
9	34.5	1200	11.0	0.0320
10	35.0	1200	13.5	0.0260
11	38.0	1200	17.0	0.0340
12	38.5	1200	23.0	0.0410
13	15.0	1100	5.3	0.0840
14	17.0	1100	7.5	0.0980
15	20.5	1100	11.0	0.0920
16	29.5	1100	17.0	0.0860

¹¹Montgomery et al. (2001).

Ejemplo: Acetileno (cont).



Ejemplo: Acetileno (cont).

Ajuste del modelo con variables originales

Modelo Ajustado:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3617.228	3136.044	-1.153	0.293
T	5.324	4.879	1.091	0.317
M	19.244	4.303	4.472	0.004
t	13766.321	10448.318	1.318	0.236
TM	-0.014	0.003	-4.404	0.005
Tt	-10.577	8.241	-1.283	0.247
Mt	-21.035	9.241	-2.276	0.063
T2	-0.002	0.002	-1.016	0.349
M2	-0.030	0.012	-2.597	0.041
T2	-11581.683	7698.610	-1.504	0.183

Matriz de Correlacion:

	T	M	t	TM	Tt	Mt	T2	M2	t2
T	1.000	0.224	-0.958	0.346	-0.963	-0.756	1.000	0.202	-0.893
M	0.224	1.000	-0.240	0.990	-0.238	0.333	0.221	0.982	-0.248
t	-0.958	-0.240	1.000	-0.353	1.000	0.762	-0.951	-0.213	0.981
TM	0.346	0.990	-0.353	1.000	-0.352	0.205	0.344	0.974	-0.349
Tt	-0.963	-0.238	1.000	-0.352	1.000	0.764	-0.957	-0.210	0.975
Mt	-0.756	0.333	0.762	0.205	0.764	1.000	-0.753	0.319	0.723
T2	1.000	0.221	-0.951	0.344	-0.957	-0.753	1.000	0.199	-0.882
M2	0.202	0.982	-0.213	0.974	-0.210	0.319	0.199	1.000	-0.230
t2	-0.893	-0.248	0.981	-0.349	0.975	0.723	-0.882	-0.230	1.000

Ejemplo: Acetileno (cont).

Variables originales: diagnóstico

VIF's:

T	M	t	TM	Tt	Mt	T2
2.856749e+06	1.095614e+04	2.017163e+06	9.802903e+03	1.428092e+06	2.403594e+02	2.501945e+06
M2	t2					
6.573359e+01	1.266710e+04					

Valores Propios:

[1]	3.543840e+13	6.843691e+08	1.133772e+05	1.156082e+04	1.371599e+03	1.79400e+00	1.00000e-02
[8]	0.000000e+00	0.000000e+00	0.000000e+00				

Valores Singulares:

[1]	5953016.142	26160.449	336.715	107.521	37.035	1.339	0.104
[8]	0.006	0.001	0.000				

Ejemplo: Acetileno (cont).

Ajuste del modelo con los regresores escalados

Modelo Ajustado:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.106	0.225	160.216	0.000
T	4.004	4.509	0.888	0.409
H	2.778	0.307	9.048	0.000
C	-8.042	6.071	-1.325	0.233
TH	-5.711	1.297	-4.404	0.005
TC	-24.202	18.856	-1.283	0.247
HC	-3.162	1.389	-2.276	0.063
T2	-9.930	9.772	-1.016	0.349
H2	-1.075	0.414	-2.597	0.041
C2	-11.909	7.916	-1.504	0.183

Matriz de Correlación:

	T	H	C	TH	TC	HC	T2	H2	C2
T	1.000	0.224	-0.958	-0.132	0.443	0.206	-0.271	0.031	-0.577
H	0.224	1.000	-0.240	0.039	0.192	-0.023	-0.148	0.498	-0.224
C	-0.958	-0.240	1.000	0.195	-0.661	-0.274	0.501	-0.018	0.765
TH	-0.132	0.039	0.195	1.000	-0.265	-0.974	0.246	0.398	0.275
TC	0.443	0.192	-0.661	-0.265	1.000	0.324	-0.972	0.126	-0.972
HC	0.206	-0.023	-0.274	-0.974	0.324	1.000	-0.279	-0.375	-0.359
T2	-0.271	-0.148	0.501	0.246	-0.972	-0.279	1.000	-0.124	0.894
H2	0.031	0.498	-0.018	0.398	0.126	-0.375	-0.124	1.000	-0.158
C2	-0.577	-0.224	0.765	0.275	-0.972	-0.359	0.894	-0.158	1.000

Ejemplo: Acetileno (cont).

Variables escaladas: diagnóstico

VIF's:

	T	H	C	TH	TC	HC	T2
	375.247759	1.740631	680.280039	31.037059	6563.345193	35.611286	1762.575365
H2							
	3.164318	1156.766284					

Valores Propios:

[1]	104.767	39.425	22.998	15.005	9.660	2.864	0.615	0.155	0.070	0.001
-----	---------	--------	--------	--------	-------	-------	-------	-------	-------	-------

Valores Singulares:

[1]	10.236	6.279	4.796	3.874	3.108	1.692	0.784	0.394	0.265	0.034
-----	--------	-------	-------	-------	-------	-------	-------	-------	-------	-------

Ejemplo: Acetileno (cont).

Ajuste de modelo con la respuesta y regresores escalados

Modelo Ajustado:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.000	0.019	0.000	1.000
T	0.336	0.379	0.888	0.409
H	0.233	0.026	9.048	0.000
C	-0.676	0.510	-1.325	0.233
TH	-0.480	0.109	-4.404	0.005
TC	-2.034	1.585	-1.283	0.247
HC	-0.266	0.117	-2.276	0.063
T2	-0.835	0.821	-1.016	0.349
H2	-0.090	0.035	-2.597	0.041
C2	-1.001	0.665	-1.504	0.183

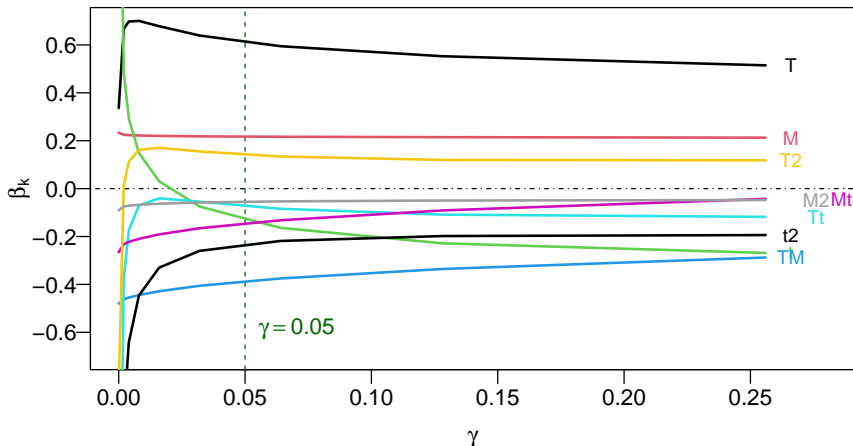
Ejemplo: Acetileno (cont).

Traza cresta:

gamma	0	0.002	0.004	0.008	0.016	0.032	0.064	0.128	0.256
T	0.336	0.572	0.631	0.667	0.676	0.663	0.633	0.597	0.564
H	0.233	0.229	0.227	0.225	0.224	0.222	0.219	0.217	0.214
C	-0.676	-0.355	-0.276	-0.227	-0.213	-0.231	-0.271	-0.318	-0.354
TH	-0.480	-0.473	-0.468	-0.460	-0.446	-0.423	-0.387	-0.339	-0.283
TC	-2.034	-0.924	-0.627	-0.408	-0.267	-0.183	-0.131	-0.099	-0.079
HC	-0.266	-0.250	-0.242	-0.232	-0.215	-0.189	-0.149	-0.097	-0.038
T2	-0.835	-0.263	-0.110	0.001	0.069	0.106	0.122	0.126	0.125
H2	-0.090	-0.082	-0.080	-0.077	-0.073	-0.068	-0.061	-0.055	-0.051
C2	-1.001	-0.538	-0.411	-0.312	-0.239	-0.180	-0.126	-0.078	-0.043

Ejemplo: Acetileno (cont).

Traza cresta



Ejemplo: Acetileno (cont).

Selección de modelo con los regresores escalados

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.1063	0.2666	135.432	< 2e-16
T	10.5194	0.2936	35.828	6.81e-12
M	2.7181	0.3318	8.192	9.56e-06
TM	-2.9614	0.3227	-9.177	3.47e-06
T2	0.9184	0.3017	3.044	0.0124
M2	-0.8389	0.3610	-2.324	0.0425

Residual standard error: 1.066 on 10 degrees of freedom

Multiple R-squared: 0.9946, Adjusted R-squared: 0.992

F-statistic: 371.5 on 5 and 10 DF, p-value: 5.13e-11

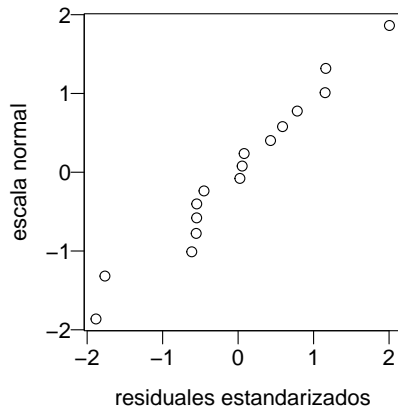
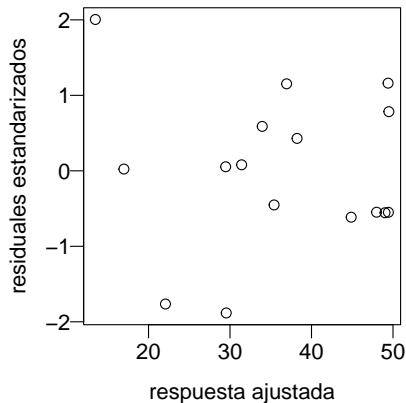
Analysis of Variance Table

Response: C

Df	Sum Sq	Mean Sq	F value	Pr(>F)
T	1 1896.68	1896.68	1667.8278	1.855e-12
M	1 56.30	56.30	49.5084	3.557e-05
TM	1 138.42	138.42	121.7211	6.413e-07
T2	1 14.80	14.80	13.0105	0.004792
M2	1 6.14	6.14	5.3993	0.042508
Residuals	10 11.37	1.14		

Ejemplo: Acetileno (cont).

Analisis de residuos



Ejemplo: Acetileno (cont).

Matriz de Correlacion:

T	M	TM	T2	M2	
T	1.000	0.224	-0.132	-0.271	0.031
M	0.224	1.000	0.039	-0.148	0.498
TM	-0.132	0.039	1.000	0.246	0.398
T2	-0.271	-0.148	0.246	1.000	-0.124
M2	0.031	0.498	0.398	-0.124	1.000

VIF's:

T	M	TM	T2	M2
1.137070	1.452181	1.373653	1.200615	1.719028

Valores Propios:

[1]	25.482	22.274	16.000	11.466	10.600	5.177
-----	--------	--------	--------	--------	--------	-------

Valores Singulares:

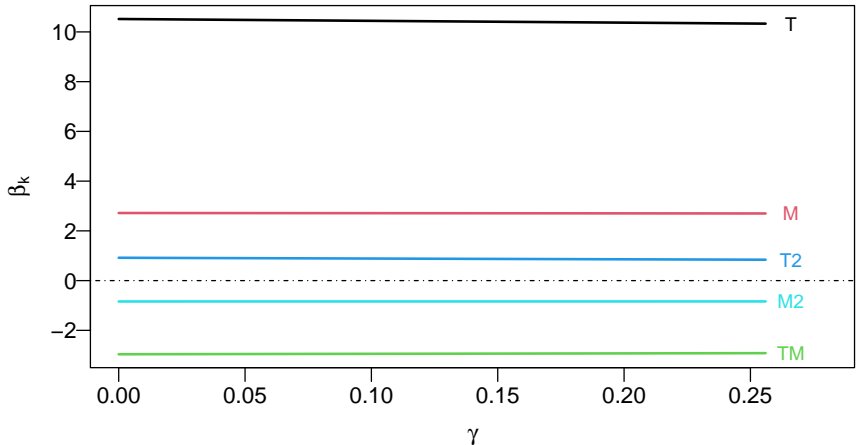
[1]	5.048	4.720	4.000	3.386	3.256	2.275
-----	-------	-------	-------	-------	-------	-------

Traza cresta:

	0	0.002	0.004	0.008	0.016	0.032	0.064	0.128	0.256
T	10.519	10.518	10.516	10.513	10.507	10.496	10.472	10.425	10.332
M	2.718	2.718	2.718	2.718	2.717	2.716	2.714	2.709	2.700
TM	-2.961	-2.961	-2.961	-2.960	-2.959	-2.956	-2.950	-2.939	-2.917
T2	0.918	0.918	0.917	0.916	0.913	0.908	0.898	0.878	0.840
M2	-0.839	-0.839	-0.839	-0.839	-0.839	-0.839	-0.839	-0.838	-0.837

Ejemplo: Acetileno (cont).

Traza cresta



Ejemplo: Acetileno (cont).

Modelo final (regresores escalados)

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36.1063	0.2666	135.432	< 2e-16
T	10.5194	0.2936	35.828	6.81e-12
M	2.7181	0.3318	8.192	9.56e-06
TM	-2.9614	0.3227	-9.177	3.47e-06
T2	0.9184	0.3017	3.044	0.0124
M2	-0.8389	0.3610	-2.324	0.0425

Residual standard error: 1.066 on 10 degrees of freedom

Multiple R-squared: 0.9946, Adjusted R-squared: 0.992

F-statistic: 371.5 on 5 and 10 DF, p-value: 5.13e-11

Referencias

- Box, G. E. P. and D. R. Cox (1964).
An Analysis of Transformations.
Journal of the Royal Statistical Society, Series B 26(2), 211–252.
- Box, G. E. P. and P. Newbold (1971).
Some comments on a paper of Coen, Gomme and Kendall.
Journal of the Royal Statistical Society Series A 134, 229–240.
- Breusch, T. and A. Pagan (1979).
A Simple Test for Heteroscedasticity and Random Coefficient Variation.
Econometrica 47(5), 1287–1294.
- Brown, P. J. (1993).
Measurement, Regression and Calibration.
Oxford: Oxford University Press.
- Chatterjee, S., A. S. Hadi, and B. Price (2000).
Regression Analysis by Example (3 ed.).
New York: Wiley.
- Coen, P. J., E. D. Gomme, and M. G. Kendall (1969).
Lagged Relationships in Economic Forecasting.
Journal of the Royal Statistical Society Series A 132(2), 133–163.
- Durbin, J. and G. Watson (1950).
Testing for Serial Correlation in Least Squares Regression, I.
Biometrika 38(3-4), 409–428.
- Durbin, J. and G. Watson (1951).
Testing for Serial Correlation in Least Squares Regression, II.
Biometrika 38(1-2), 159–179.
- Durbin, J. and G. Watson (1971).
Testing for Serial Correlation in Least Squares Regression, III.
Biometrika 58(1), 1–19.