

Universidad Panamericana
Maestría en Ciencia de Datos
Econometría

Tarea RLM

Enrique Ulises Báez Gómez Tagle

9 de septiembre de 2025

Índice

1	Pregunta 1	2
2	Pregunta 2	3
3	Pregunta 3	4
4	Pregunta 4	6
5	Link al repositorio con código fuente	7

1. Pregunta 1

a) Considere los datos de la tabla 1.

Y	X2	X3
1	1	2
3	2	1
8	3	-3

Cuadro 1: Datos de la pregunta 1

b) Con base en estos datos, estime las siguientes regresiones:

$$Y_i = \alpha_1 + \alpha_2 X_{2i} + u_{1i},$$

$$Y_i = \lambda_1 + \lambda_3 X_{3i} + u_{2i},$$

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i,$$

a) ¿Es $\alpha_2 = \beta_2$? ¿Por qué?

b) ¿Es $\lambda_3 = \beta_3$? ¿Por qué?

c) ¿Qué conclusión importante obtiene de este ejercicio?

Con los datos (Y, X_2, X_3) :

$$Y = \{1, 3, 8\}, \quad X_2 = \{1, 2, 3\}, \quad X_3 = \{2, 1, -3\}.$$

Estimaciones:

$$(1) Y_i = \alpha_1 + \alpha_2 X_{2i} + u_{1i},$$

$$\hat{\alpha}_1 = -3, \hat{\alpha}_2 = 3.5.$$

$$(2) Y_i = \lambda_1 + \lambda_3 X_{3i} + u_{2i},$$

$$\hat{\lambda}_1 = 4, \hat{\lambda}_3 = -1.3571.$$

$$(3) Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i,$$

$$\hat{\beta}_1 = 2, \hat{\beta}_2 = 1, \hat{\beta}_3 = -1.$$

a) No, $\alpha_2 \neq \beta_2$. El estimador α_2 en la regresión simple está sesgado porque omite X_3 , correlacionado con X_2 . Se cumple la fórmula del sesgo por variable omitida.

b) Tampoco, $\lambda_3 \neq \beta_3$. Análogamente, al omitir X_2 , el coeficiente de X_3 se ve afectado por su correlación con X_2 .

c) Este ejercicio nos permite entender el *sesgo por variable omitida*. Los coeficientes en regresiones simples (α_2, λ_3) difieren de los verdaderos efectos parciales (β_2, β_3) que sólo se identifican en la regresión múltiple.

2. Pregunta 2

a) La demanda de rosas. En la Tabla 2 se presentan datos trimestrales (1971-III a 1975-II) sobre estas variables:

Y = cantidad de rosas vendidas (docenas);

X_2 = precio promedio al mayoreo de rosas (\$/docena);

X_3 = precio promedio al mayoreo de claveles (\$/docena);

X_4 = ingreso familiar disponible promedio semanal (\$/semana);

X_5 = variable de tendencia que toma valores de (1,2,...), durante el periodo 1971-III a 1975-II en el área metropolitana de Detroit.

Año-trim	Y	X_2	X_3	X_4	X_5
1971-III	11484	2.26	3.49	158.11	1
1971-IV	9348	2.54	2.85	173.36	2
1972-I	8429	3.07	4.06	165.26	3
1972-II	10079	2.91	3.64	172.92	4
1972-III	9240	2.73	3.21	178.46	5
1972-IV	8862	2.77	3.66	198.62	6
1973-I	6216	3.59	3.76	186.28	7
1973-II	8253	3.23	3.49	188.98	8
1973-III	8038	2.60	3.13	180.49	9
1973-IV	7476	2.89	3.20	183.33	10
1974-I	5911	3.77	3.65	181.87	11
1974-II	7950	3.64	3.60	185.00	12
1974-III	6134	2.82	2.94	184.00	13
1974-IV	5868	2.96	3.12	188.20	14
1975-I	3160	4.24	3.58	175.67	15
1975-II	5872	3.69	3.53	188.00	16

Cuadro 2: Demanda trimestral de rosas en Detroit (1971-III a 1975-II).

Se le pide considerar las siguientes funciones de demanda:

$$Y_t = \alpha_1 + \alpha_2 X_{2t} + \alpha_3 X_{3t} + \alpha_4 X_{4t} + \alpha_5 X_{5t} + u_t,$$

$$\ln Y_t = \beta_1 + \beta_2 \ln X_{2t} + \beta_3 \ln X_{3t} + \beta_4 \ln X_{4t} + \beta_5 X_{5t} + u_t.$$

(a) Estime los parámetros del modelo lineal e interprete los resultados.

Modelo lineal estimado por MCO (con intercepto):

$$\hat{Y}_t = \hat{\alpha}_1 + \hat{\alpha}_2 X_{2t} + \hat{\alpha}_3 X_{3t} + \hat{\alpha}_4 X_{4t} + \hat{\alpha}_5 X_{5t},$$

$$\text{con } \hat{\alpha}_1 = 10820.0, \quad \hat{\alpha}_2 = -2227.70 \ (t = -2.42, \ p = 0.034), \quad \hat{\alpha}_3 = 1251.14 \ (t = 1.08, \ p = 0.303),$$

$$\hat{\alpha}_4 = 6.283 \ (t = 0.21, \ p = 0.841), \quad \hat{\alpha}_5 = -197.40 \ (t = -1.94, \ p = 0.078).$$

Ajuste e inferencia: $R^2 = 0.835$, $R^2_{adj} = 0.775$, $F = 13.89$ ($p = 0.000281$). Durbin-Watson = 2.33.

El precio propio de las rosas (X_2) tiene signo negativo y es significativo al 5%; el precio de los claveles (X_3) es positivo pero no significativo; el ingreso (X_4) es positivo pero no significativo; la tendencia (X_5) es negativa y significativa (10%). El número de condición elevado (4.48×10^3) podría indicar *multicolinealidad* potencial.

(b) Estime los parámetros del modelo log-lineal e interprete los resultados.

Modelo log-lineal estimado por MCO:

$$\widehat{\ln Y}_t = \hat{\beta}_1 + \hat{\beta}_2 \ln X_{2t} + \hat{\beta}_3 \ln X_{3t} + \hat{\beta}_4 \ln X_{4t} + \hat{\beta}_5 X_{5t},$$

con $\hat{\beta}_1 = 3.572$, $\hat{\beta}_2 = -1.1707$ ($t = -2.40$, $p = 0.035$), $\hat{\beta}_3 = 0.7379$ ($t = 1.13$, $p = 0.282$),

$\hat{\beta}_4 = 1.1532$ ($t = 1.28$, $p = 0.227$), $\hat{\beta}_5 = -0.0301$ ($t = -1.83$, $p = 0.094$).

Ajuste e inferencia: $R^2 = 0.799$, $R^2_{adj} = 0.726$, $F = 10.92$ ($p = 0.000798$). Durbin-Watson = 2.05.

Aquí, los coeficientes $\beta_2, \beta_3, \beta_4$ son elasticidades: la demanda es elástica al precio propio (-1.17 , significativo al 5%), presenta elasticidad cruzada positiva frente al precio de claveles (0.74 , no significativa) y es normal (elasticidad ingreso 1.15 , no significativa). La tendencia es levemente decreciente (10%).

- (c) β_2 , β_3 y β_4 dan, respectivamente, las elasticidades de la demanda respecto al precio propio, precio cruzado e ingreso. ¿Cuáles son, a priori, los signos esperados de estas elasticidades? ¿Concuerdan estos resultados con las expectativas a priori?

Expectativas a priori: $\beta_2 < 0$ (ley de la demanda), $\beta_3 > 0$ si claveles son sustitutos, y $\beta_4 > 0$ si las rosas son normal.

Resultados: $\hat{\beta}_2 = -1.1707 < 0$, $\hat{\beta}_3 = 0.7379 > 0$, $\hat{\beta}_4 = 1.1532 > 0$. Los signos *concuerdan* con la teoría; sólo el efecto de precio propio es estadísticamente significativo al 5%.

- (d) ¿Cómo calcularía las elasticidades precio propio, precio cruzado e ingreso en el modelo lineal? Calculamos las elasticidades en un punto de evaluación como

$$\varepsilon_{Y, X_j} = \frac{\partial Y}{\partial X_j} \frac{\bar{X}_j}{\bar{Y}} = \hat{\alpha}_j \frac{\bar{X}_j}{\bar{Y}}, \quad j \in \{2, 3, 4\}.$$

Evaluadas en las medias, se obtienen:

$$\varepsilon_{\text{precio propio}} = -0.9053, \quad \varepsilon_{\text{precio cruzado}} = 0.5616, \quad \varepsilon_{\text{ingreso}} = 0.1484.$$

Para el modelo log-lineal, las elasticidades son constantes e iguales a los coeficientes: $\varepsilon_p = -1.1707$, $\varepsilon_{pc} = 0.7379$, $\varepsilon_y = 1.1532$.

- (e) Con base en su análisis, ¿cuál modelo, si existe, escogería y por qué?

Comparación: el modelo lineal exhibe mayor $R^2_{adj} = 0.775$ que el log-lineal (0.726), pero el log-lineal se ve favorecido por los criterios de información ($AIC = -9.08$, $BIC = -5.22$ frente a 269.48 y 273.34). Además, el log-lineal entrega elasticidades directamente interpretables y suele capturar mejor relaciones proporcionales.

Con base en AIC/BIC de OLS bajo normalidad y en la interpretación económica (elasticidades), **elegimos el modelo log-lineal**. No obstante, la muestra es pequeña ($n = 16$) y hay indicios de multicolinealidad; es recomendable revisar los resultados con precaución.

3. Pregunta 3

- a) Desembolsos del presupuesto de defensa de Estados Unidos, 1962–1981. Para explicar el presupuesto de defensa, considere el siguiente modelo:

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + \beta_4 X_{4t} + \beta_5 X_{5t} + u_t.$$

Donde:

Y_t = desembolsos del presupuesto de defensa durante el año t , \$ miles de millones.

X_{2t} = PNB durante el año t , \$ miles de millones.

X_{3t} = ventas militares de Estados Unidos/ayuda en el año t , \$ miles de millones.

X_{4t} = ventas de la industria aeroespacial, \$ miles de millones.

X_{5t} = conflictos militares que implican a más de 100 000 soldados. Esta variable adquiere el valor de 1 cuando participan 100 000 soldados o más, y es igual a cero cuando el número de soldados no llega a 100 000.

Año	Y	X ₂	X ₃	X ₄	X ₅
1962	51.1	560.3	0.6	16.0	0
1963	52.3	590.5	0.9	16.4	0
1964	53.6	632.4	1.1	16.7	0
1965	49.6	684.9	1.4	17.0	1
1966	56.8	749.9	1.6	20.2	1
1967	70.1	793.0	1.0	23.1	1
1968	80.5	865.0	0.8	25.6	1
1969	81.2	931.4	1.5	24.6	1
1970	80.3	992.7	1.0	24.8	1
1971	77.7	1077.6	1.5	27.1	1
1972	78.3	1185.9	2.95	21.5	1
1973	74.5	1326.4	4.8	24.3	0
1974	77.8	1434.2	10.3	26.8	0
1975	85.6	1549.2	16.0	29.5	0
1976	89.4	1748.0	14.7	30.4	0
1977	97.5	1918.3	8.3	33.3	0
1978	105.2	2163.9	11.0	38.0	0
1979	117.7	2417.8	13.0	46.2	0
1980	135.9	2633.1	15.3	57.6	0
1981	162.1	2937.7	18.0	68.9	0

Cuadro 3: EE. UU.: Presupuesto de defensa y variables explicativas (1962–1981).

b) Con base en la Tabla 3, responda:

- (a) Estime los parámetros del modelo lineal y sus errores estándar, y obtenga R^2 y R^2 ajustada. El modelo estimado por MCO (entre paréntesis se reportan los errores estándar) es

$$\hat{Y}_t = \hat{\beta}_1 + \hat{\beta}_2 X_{2t} + \hat{\beta}_3 X_{3t} + \hat{\beta}_4 X_{4t} + \hat{\beta}_5 X_{5t},$$

$$\text{con } \hat{\beta}_1 = 19.7122 \text{ (3.3509)}, \quad \hat{\beta}_2 = 0.0164 \text{ (0.0065)}, \quad \hat{\beta}_3 = -0.2261 \text{ (0.4556)}, \\ \hat{\beta}_4 = 1.3967 \text{ (0.2608)}, \quad \hat{\beta}_5 = 5.3564 \text{ (3.0201)},$$

Métricas de ajuste: $R^2 = 0.9784$, $R^2_{adj} = 0.9726$, $F = 169.5$ ($p = 2.73 \times 10^{-12}$). Durbin–Watson = 1.169.

Ecuación en niveles: $\hat{Y}_t = 19.7122 + 0.0164 X_{2t} - 0.2261 X_{3t} + 1.3967 X_{4t} + 5.3564 X_{5t}$.

- (b) Comente los resultados, considerando cualquier expectativa *a priori* que tenga sobre la relación entre Y y las diversas variables X.

Signos esperados: se anticipa efecto positivo de PNB (X_2), ventas militares/ayuda (X_3), ventas aeroespaciales (X_4) y de la dummy de conflicto (X_5).

Resultados: $\hat{\beta}_2 > 0$ y significativo al 5 % ($t = 2.51$); $\hat{\beta}_4 > 0$ y altamente significativo ($t = 5.36$); $\hat{\beta}_5 > 0$ y marginal al 10 % ($t = 1.77$); $\hat{\beta}_3 < 0$ y no significativo.

En promedio, manteniendo todo lo demás constante:

- Un aumento de \$1 mil millones en el PNB se asocia con 0.016 mil millones adicionales en defensa.
- Un aumento de \$1 mil millón en ventas aeroespaciales se asocia con 1.397 mil millones adicionales en defensa.
- La presencia de un conflicto $X_5 = 1$ eleva el gasto en ≈ 5.36 mil millones.

Diagnóstico: el número de condición ($\approx 5.53 \times 10^3$) y VIF (sin intercepto) altos obtenidos (p.ej. $VIF_{X_2} \approx 80$, $VIF_{X_4} \approx 62$) sugieren una *multicolinealidad* severa entre regresores macro, lo cual puede inflar errores estándar y volver inestables algunos signos (como X_3). Durbin–Watson = 1.17 sugiere posible autocorrelación positiva de primer orden en residuos (series anuales).

(c) ¿Qué otra(s) variable(s) incluiría en el modelo y por qué?

Incluiría variables para (i) trabajar en términos reales y (ii) capturar dinámica/geopolítica:

- Deflactor del gasto de defensa o CPI (para expresar todas las series en términos reales) y una tendencia temporal.
- Petróleo y/o choques energéticos 1973–79; tasa de inflación o interés (política macro).
- Gasto/PNB rezagado o Y_{t-1} (inercia presupuestal) y rezagos de X_2, X_4 .
- Dummies geopolíticas (p.ej., Vietnam 1965–73) o un indicador de tensiones internacionales adicional a X_5 .

Con estas variables, se podrían mitigar sesgos por omisión y reducir la multicolinealidad al separar tendencias comunes entre X_2 y X_4 .

4. Pregunta 4

La tabla 7.12 presenta datos del gasto de consumo real, ingreso real, riqueza real y tasas de interés reales de Estados Unidos de 1947 a 2000.

Año	C	Yd	Riqueza	Tasa de Interés
1947	976.4	1035.2	5166.8	-10.351
1948	998.1	1090.0	5280.8	-4.720
1949	1025.3	1095.6	5607.4	1.044
1950	1090.9	1192.7	5759.5	0.407
1951	1107.7	1227.0	6081.6	-5.283
1952	1142.4	1266.8	6243.9	-0.277
1953	1221.4	1327.5	6355.6	0.561
1954	1277.2	1344.0	6797.4	-0.138
1955	1314.0	1433.8	7172.2	0.262
1956	1348.8	1502.3	7375.2	-0.736
1957	1381.8	1539.5	7315.3	-0.261
1958	1393.0	1553.7	7870.0	-0.575
1959	1470.7	1623.8	8188.1	2.296
1960	1516.0	1664.8	8351.8	1.511
1961	1541.2	1720.0	8971.9	1.296
1962	1617.3	1803.5	9091.5	1.396
1963	1684.8	1871.5	9436.1	2.085
1964	1784.8	2006.9	10004.4	2.027
1965	1897.6	2131.0	10562.8	2.112
1966	2066.2	2244.6	11502.0	2.220
1967	2066.2	2340.5	12341.0	2.120
1968	2264.8	2448.2	12145.4	1.055
1969	2314.5	2524.3	11672.3	1.732
1970	2405.2	2630.0	11650.8	1.176
1971	2505.5	2745.3	12312.9	-0.712
1972	2650.5	2874.3	13499.9	-0.156
1973	2675.9	3072.3	13081.0	1.414
1974	2653.7	3051.9	11868.8	-1.043
1975	2710.9	3108.5	12634.4	-3.534
1976	2868.9	3243.5	13456.8	-0.657
1977	2992.1	3360.7	13786.3	-1.190
1978	3124.7	3527.5	14450.5	0.113
1979	3203.2	3628.6	15340.0	1.704
1980	3193.0	3658.0	15965.0	2.298
1981	3236.0	3741.1	15965.0	4.704
1982	3275.5	3791.7	16312.5	4.449
1983	3454.3	3906.9	16944.8	5.691
1984	3640.6	4207.6	17526.7	5.848
1985	3820.9	4347.8	19068.3	4.331
1986	3981.2	4486.6	20530.0	3.768
1987	4113.4	4586.5	21235.7	2.819
1988	4279.5	4784.1	22332.0	3.287
1989	4393.7	4906.5	23659.8	4.318
1990	4474.5	5014.2	23105.1	3.595
1991	4466.6	5033.0	24050.2	1.803
1992	4594.5	5189.3	24418.2	1.007
1993	4748.9	5261.3	25092.3	0.625
1994	4928.1	5397.2	25218.6	2.206
1995	5075.6	5539.1	27439.7	3.333
1996	5237.5	5677.7	29448.2	3.083
1997	5423.9	5854.5	32664.1	3.120
1998	5683.7	6168.6	35887.0	3.584
1999	5968.4	6320.0	39591.3	3.245
2000	6257.8	6539.2	38167.7	3.576

Cuadro 4: Gasto de consumo real, ingreso real, riqueza real y tasas de interés reales de Estados Unidos (1947–2000).

(a) Con los datos de la tabla, estime la función de consumo lineal usando los datos de ingreso, riqueza y tasa de interés. ¿Cuál es la ecuación ajustada?

Modelo estimado por MCO (errores estándar entre paréntesis):

$$\hat{C}_t = \hat{\gamma}_1 + \hat{\gamma}_2 Yd_t + \hat{\gamma}_3 Riqueza_t + \hat{\gamma}_4 Tasa_t,$$

con $\hat{\gamma}_1 = -3.0103 (15.014)$, $\hat{\gamma}_2 = 0.7344 (0.0160)$, $\hat{\gamma}_3 = 0.0354 (0.0030)$, $\hat{\gamma}_4 = -5.7072 (2.679)$.
 Métricas de ajuste: $R^2 = 0.9992$, $R^2_{adj} = 0.9991$, $F = 1.975 \times 10^4$ ($p = 8.05 \times 10^{-77}$), $n = 54$.
 Durbin-Watson = 1.313.
 Ecuación: $\hat{C}_t = -3.010 + 0.7344 Yd_t + 0.0354 Riqueza_t - 5.7072 Tasa_t$.

- (b) ¿Qué indican los coeficientes estimados sobre las relaciones entre las variables y el gasto de consumo?
Signos esperados vs. estimados: $\hat{\gamma}_2 > 0$ (cumple), $\hat{\gamma}_3 > 0$ (cumple) y $\hat{\gamma}_4 < 0$ (cumple).

Interpretación marginal:

- Propensión marginal a consumir del ingreso disponible: $\partial C / \partial Yd = \hat{\gamma}_2 = 0.7344$. Un incremento de 1 unidad en Yd aumenta el consumo en 0.734 unidades, ceteris paribus.
- Efecto riqueza: $\partial C / \partial W = \hat{\gamma}_3 = 0.0354$. El consumo crece con la riqueza, aunque el impacto unitario es menor que el del ingreso.
- Efecto de la tasa real: $\partial C / \partial r = \hat{\gamma}_4 = -5.7072$. Tasas más altas reducen el consumo (sustitución intertemporal); el coeficiente es significativo al 5 %.

Magnitudes relativas (betas estandarizados): ingreso 0.802, riqueza 0.209, tasa -0.011 : el ingreso explica la mayor parte de la variación contemporánea del consumo, seguido por la riqueza; la tasa tiene efecto pequeño pero de signo teórico. *Multicolinealidad - VIF (sin intercepto)* : $VIF_{Yd} \approx 65.9$, $VIF_{Riqueza} \approx 65.8$, $VIF_{Tasa} \approx 1.56$ y un número de condición $\approx 4.48 \times 10^4$ sugieren fuerte colinealidad entre ingreso y riqueza; esto puede inflar errores estándar y volver sensibles las estimaciones a la especificación. Para robustez, puede considerarse trabajar en logaritmos/tasas de crecimiento, usar componentes permanentes o rezagos.

5. Link al repositorio con código fuente

<https://github.com/enriquegomeztagle/MCD-Econometria/tree/main/HWs/MLR-Homework>