Resolución de sistemas de ecuaciones lineales

Problemas para la ciencia de datos

PID_00262433

Francesc Pozo Montero
Jordi Ripoll Missé

Universitat
Oberta
de Catalunya

Francesc Pozo Montero

Licenciado en Matemáticas por la Universidad de Barcelona (2000) y doctor en Matemática Aplicada por la Universidad Politécnica de Cataluña (2005). Ha sido profesor asociado de la Universidad Autónoma de Barcelona y profesor asociado, colaborador y actualmente profesor agregado en la Universidad Politécnica de Cataluña. Además, es cofundador del Grupo de Innovación Matemática E-learning (GIMEL), responsable de varios proyectos de innovación docente y autor de varias publicaciones. Como miembro del grupo de investigación consolidado CoDAlab, centra su investigación en la teoría de control y las aplicaciones en ingeniería mecánica y civil, así como en el uso de la ciencia de datos para la monitorización de la integridad estructural y para la monitorización de la condición, sobre todo en turbinas eólicas.

Jordi Ripoll Missé

Licenciado en Matemáticas y doctor en Ciencias Matemáticas por la Universidad de Barcelona (2005). Profesor colaborador de la Universitat Oberta de Catalunya desde 2011 y profesor del Departamento de Informática, Matemática Aplicada y Estadística de la Universidad de Girona (UdG) desde 1996, donde actualmente es profesor agregado y desarrolla tareas de investigación en el ámbito de la biología matemática (modelos con ecuaciones en derivadas parciales y dinámica evolutiva). También ha sido profesor y tutor de la UNED en dos etapas, primero en el centro asociado de Terrassa y actualmente en el de Girona. Ha participado en numerosos proyectos de innovación docente, especialmente en cuanto al aprendizaje de las matemáticas en línea.

El encargo y la creación de este recurso de aprendizaje UOC han sido coordinados por la profesora: Cristina Cano Bastidas (2019)

Primera edición: febrero 2019

Av. Tibidabo, 39-43, 08035 Barcelona
Diseño: Manel Andreu
Realización editorial: Oberta UOC Publishing, SL

Índice

1. Problemas

Resolución de sistemas de ecuaciones lineales (SEL)

Temas: *definición de SEL (coeficientes, incógnitas, parámetros), número de soluciones de los SEL, teorema de Rouché-Fröbenius, matriz de coeficientes y matriz ampliada del sistema, SEL homogéneos (con solución única trivial o con infinitas soluciones), método de resolución de Gauss, método de resolución de Cramer, interpretación geométrica de las ecuaciones lineales (rectas y planos) y de las soluciones del sistema (intersección).*

1. Una de las aplicaciones más interesantes de la teoría de los sistemas lineales en la ciencia de datos es el ajuste de un conjunto de datos por la denominada *recta de regresión*, que minimiza el error entre datos y la predicción del modelo lineal.

El problema se puede formular de la siguiente manera: disponemos de datos de observaciones (a menudo, en un número grande n) de dos características de un conjunto de individuos: $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$. A continuación construimos un modelo lineal $y = ax + b$, donde a y b son dos parámetros para determinar, que minimice la distancia (al cuadrado) entre las predicciones del modelo y los valores reales:

$$E = \sum_{i=1}^{n}(ax_i + b - y_i)^2$$

La solución a este problema de optimización consiste en resolver el siguiente sistema de ecuaciones lineales $\frac{\partial E}{\partial a} = 2\sum_{i=1}^{n}(ax_i + b - y_i)x_i = 0$, $\frac{\partial E}{\partial b} = 2\sum_{i=1}^{n}(ax_i + b - y_i) = 0$ para las variables a y b, que, al dividir por n ambas ecuaciones, se transforma en:

$$\left.\begin{array}{rcl} a\frac{1}{n}\sum_{i=1}^{n} x_i^2 + b\frac{1}{n}\sum_{i=1}^{n} x_i & = & \frac{1}{n}\sum_{i=1}^{n} x_i y_i \\ a\frac{1}{n}\sum_{i=1}^{n} x_i + b & = & \frac{1}{n}\sum_{i=1}^{n} y_i \end{array}\right\}$$

Introduciendo la notación por las medias de los datos, por ejemplo, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, etc., tenemos de la segunda ecuación $b = \bar{y} - a\bar{x}$ y, al sustituirla en la primera, tenemos $a(\overline{x^2} - \bar{x}^2) = \overline{xy} - \bar{x}\bar{y}$. Por lo tanto, hemos hallado los coeficientes de la recta de regresión.

a) Calculad la ecuación de la recta de regresión $y = ax + b$ a partir de los $n = 8$ datos siguientes:

$$(X,Y) : (-10,-15),(-5,1),(0,1),(5,3),(10,20),(10,7),(15,42),(20,50)$$

Una vez calculado el ajuste lineal, dibujad en un mismo gráfico los datos del problema y la recta obtenida. ¿Se ajustan bien los datos a la recta?

b) Disponemos de los siguientes datos de un laboratorio, sobre dos características de los pacientes analizados:

$$\begin{aligned} X:&\quad 4.17,5.58,5.18,6.11,4.50,4.61,5.17,4.53,5.33,5.14\\ Y:&\quad 4.81,4.17,4.41,3.59,5.87,3.83,6.03,4.89,4.32,4.69 \end{aligned}$$

Efectuad un ajuste lineal $y = ax+b$ entre las características X e Y. Dibujad en un mismo gráfico los datos (x_i,y_i), $i = 1\ldots 10$, la recta de regresión y las predicciones del modelo $(x_i,\hat{y}_i = ax_i + b)$, $i = 1\ldots 10$.

2. Imaginad que tenemos los datos de la temperatura de una placa metálica cuadrada a lo largo de su perímetro:

Nota del ejercicio 2

En este ejercicio hemos visto cómo se resuelve el fenómeno de la difusión del calor, a partir de unos datos y utilizando la teoría de los sistemas lineales.

$$T = \begin{pmatrix} & 95^\circ & 87^\circ & 63^\circ & \\ 53^\circ & T_{22} & T_{23} & T_{24} & 51^\circ \\ 29^\circ & T_{32} & T_{33} & T_{34} & 52^\circ \\ 25^\circ & T_{42} & T_{43} & T_{44} & 48^\circ \\ & 18^\circ & 21^\circ & 20^\circ & \end{pmatrix}$$

A partir de estos datos, queremos deducir la temperatura de los puntos del interior de la placa. Para obtener una aproximación de la temperatura de los puntos interiores $T_{22},T_{23},\ldots,T_{44}$, podemos suponer que para cada punto la temperatura es igual a la temperatura media de los puntos que lo rodean: arriba, abajo, a la derecha y a la izquierda.

a) A partir de las hipótesis de este modelo, escribid el sistema de ecuaciones lineales para la temperatura en cada uno de los nueve puntos interiores. Escribid la matriz ampliada del sistema.

b) Hallad la solución del sistema lineal. ¿Es única? ¿En qué puntos interiores se alcanzan las temperaturas máxima y mínima?

c) Con las soluciones obtenidas, ¿podéis calcular cuál es la temperatura en cada una de las cuatro esquinas de la placa: T_{11}, T_{15}, T_{51} y T_{55}? (Media de los tres valores que lo rodean.)

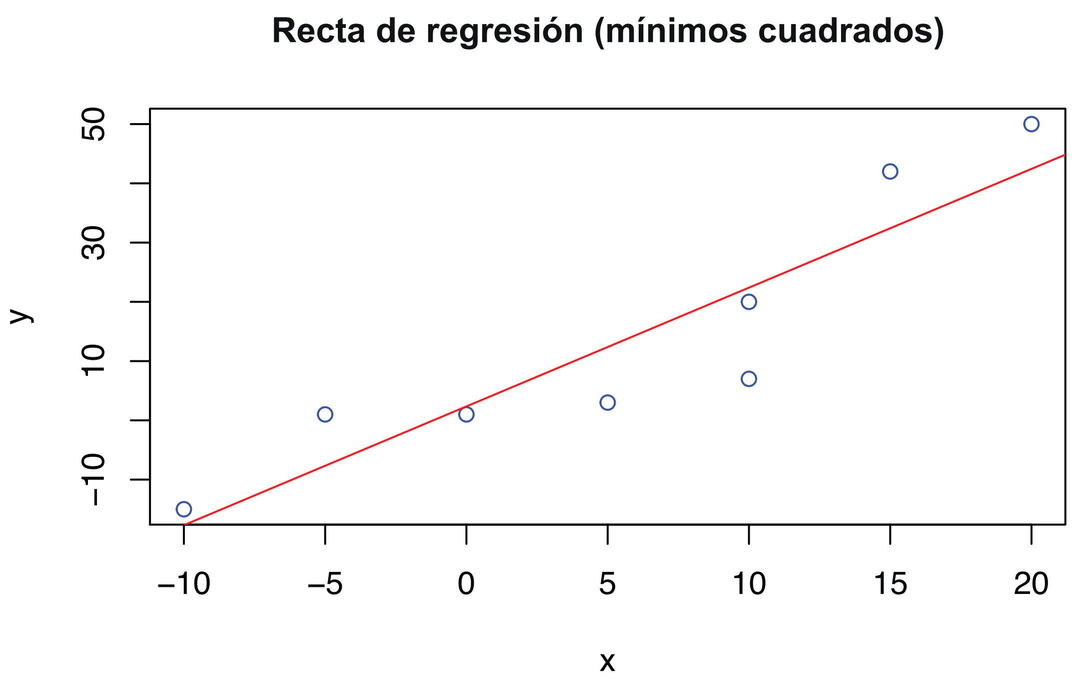
3. En las últimas elecciones a rector de una universidad, los tres candidatos obtuvieron los siguientes porcentajes de votos: 20 %, 27 % y 53 %, respectivamente. Además, el candidato ganador obtuvo 732 votos más que el total de sus rivales. ¿Cuántos votos obtuvo cada candidato?

2. Soluciones

1. a) A partir de los datos, calculamos los coeficientes del sistema lineal de dos ecuaciones con dos incógnitas (a y b) y hallamos su solución. La recta obtenida es, pues, $y = ax + b = 2.004329x + 2.350649$.
Esto también se puede hacer directamente con el siguiente código R:

```
> # Datos: dos características X e Y
> datos <- matrix(c(-10,-15,-5,1,0,1,5,3,10,20,10,7,15,42,20,50), 2,8)
> rownames(datos) <- c("X","Y")
> datos

  [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8]
X  -10   -5    0    5   10   10   15   20
Y  -15    1    1    3   20    7   42   50

> x <- datos [1,]; y <- datos[2,]
> # Ajuste a un modelo lineal (recta de regresión y=a*x+b)
> fit <- lm(y ~ x)
> summary(fit) # resumen

Call:
lm(fórmula = y ~ x)

Residuals:
    Min      1Q  Median      3Q     Max
-15.394  -4.138   0.671   7.840   9.584

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   2.3506     3.9832   0.590  0.57663
x             2.0043     0.3608   5.555  0.00144 **
---
Signif. codes:  0 ' *** ' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9.694 on 6 degrees of freedom
Multiple R-squared:  0.8372,        Adjusted R-squared:  0.8101
F-statistic: 30.86 on 1 and 6 DF,  p-value: 0.001439

> # coeficientes de la recta y=a*x+b, a=coef[2], b=coef[1]
> coef <- coefficients(fit); coef

(Intercept)           x
   2.350649    2.004329
```

```
> # Dibujo
> plot(x,y,col= "blue")     # datos (círculos)
> abline(fit,col = "red")   # recta
> title("Recta de regresión (mínimos cuadrados)")
```

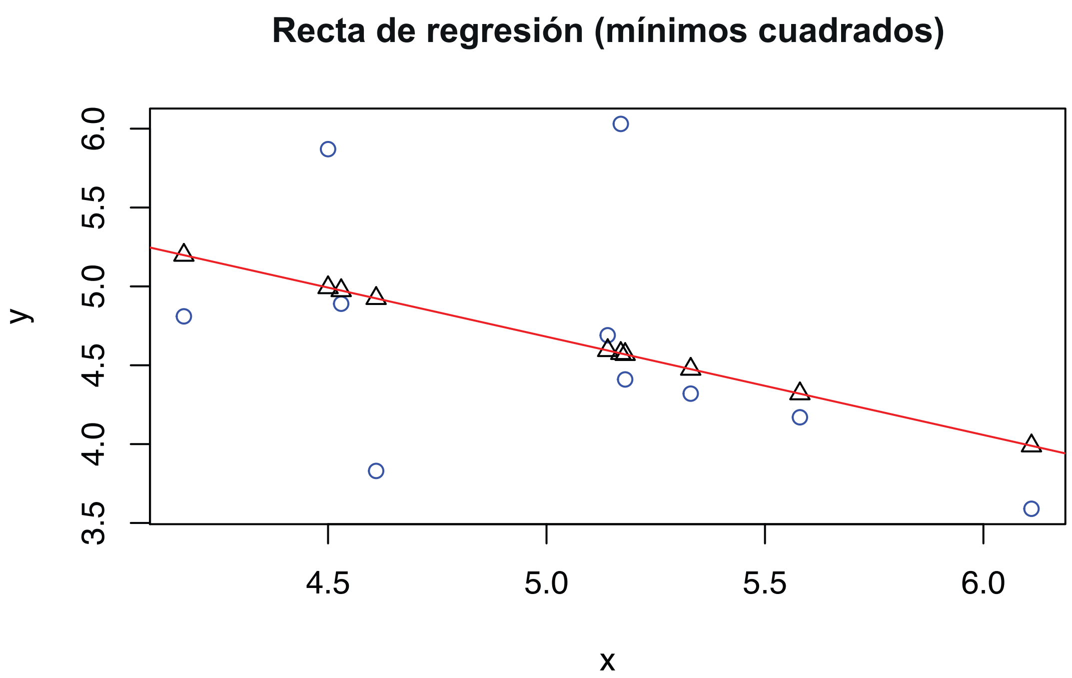


Según la gráfica, podemos decir que los datos se ajustan bastante bien a la recta.

b) De nuevo, a partir de los datos del ejercicio, calculamos los coeficientes del sistema lineal de dos ecuaciones con dos incógnitas (a y b) y hallamos la solución. La recta obtenida es, en este caso, $y = ax + b = -0.6229559x + 7.7957139$ y, a partir de la ecuación de la recta, calculamos las predicciones $\hat{y} = ax_i + b$, $i = 1,\ldots,10$. Esto también se puede hacer directamente con el siguiente código R:

```
> # Datos:
> x <- c(4.17,5.58,5.18,6.11,4.50,4.61,5.17,4.53,5.33,5.14)
> y <- c(4.81,4.17,4.41,3.59,5.87,3.83,6.03,4.89,4.32,4.69)
> # Ajuste a un modelo lineal (recta de regresión y=a*x+b)
> fit <- lm(y ~ x)
> summary(fit) # resumen

Call:
lm(fórmula = y ~ x)

Residuals:
     Min       1Q   Median       3Q      Max
-1.09389 -0.33069 -0.15249  0.05128  1.45497
```

```
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.7957     2.1661   3.599  0.00699 **
x            -0.6230     0.4279  -1.456  0.18351
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7485 on 8 degrees of freedom
Multiple R-squared:  0.2095,        Adjusted R-squared:  0.1106
F-statistic:  2.12 on 1 and 8 DF,  p-value: 0.1835

> # coeficientes de la recta y=a*x+b, a=coef[2], b=coef[1]
> coef <- coefficients(fit); coef

(Intercept)           x
  7.7957139  -0.6229559

> yy <- fitted(fit) # valores de las predicciones
> yy

       1        2        3        4        5        6        7        8
5.197988 4.319620 4.568803 3.989454 4.992413 4.923887 4.575032 4.973724
       9       10
4.475359 4.593721

> # Dibujo
> plot(x,y,col= "blue")     # datos (círculos)
> points(x,yy, pch=2)       # predicciones (triángulos)
> abline(fit,col = "red")   # recta
> title("Recta de regresión (mínimos cuadrados)")
```



2. a) Tenemos los datos de la temperatura en los puntos exteriores de la placa metálica. Las ecuaciones para la temperatura en los nueve puntos interiores se obtienen de la siguiente manera. Para el primer punto interior tenemos: $T_{22} = (95 + T_{32} + T_{23} + 53)/4$, es decir, $4T_{22} - T_{32} - T_{23} = 148$. Para el segundo punto interior tenemos: $T_{23} = (87 + T_{33} + T_{24} + T_{22})/4$, es decir, $4T_{23} - T_{33} - T_{24} - T_{22} = 87$, etc. La matriz ampliada del sistema es la siguiente:

$$M = (A|b) = \begin{pmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 148 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 87 \\ 0 & -1 & 4 & 0 & 0 & -1 & 0 & 0 & 0 & 114 \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 & 29 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 & 52 \\ 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 & 43 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 & 21 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & 68 \end{pmatrix}$$

b) Este sistema se puede solucionar por el método de Gauss y su única solución es: $T_{22} = 63.43304$, $T_{23} = 63.44643$, $T_{24} = 56.41518$, $T_{32} = 42.28571$, $T_{33} = 46.93750$, $T_{34} = 48.21429$, $T_{42} = 29.77232$, $T_{43} = 33.80357$, $T_{44} = 37.50446$. La temperatura máxima se alcanza en $T_{23} = 63.44643°$ y la mínima en $T_{42} = 29.77232°$.

c) Al hacer las medias de los tres puntos que rodean cada una de las cuatro esquinas obtenemos:
$T_{11} = 70.47768$, $T_{15} = 56.80506$, $T_{51} = 24.25744$ y $T_{55} = 35.16815$.
La resolución de este problema se puede efectuar con el siguiente código R:

```
> T<- matrix(c(NA,53,29,25,NA,95,NA,NA,NA,18,
+              87,NA,NA,NA,21,63,NA,NA,NA,20,NA,51,52,48,NA), 5,5)
> T

     [,1] [,2] [,3] [,4] [,5]
[1,]   NA   95   87   63   NA
[2,]   53   NA   NA   NA   51
[3,]   29   NA   NA   NA   52
[4,]   25   NA   NA   NA   48
[5,]   NA   18   21   20   NA

> A <- matrix(c(4, -1, 0, -1, 0, 0, 0, 0, 0,-1, 4, -1, 0, -1, 0, 0, 0, 0,
+               0, -1, 4, 0, 0, -1, 0, 0, 0,-1, 0, 0, 4, -1, 0, -1, 0, 0,
+               0, -1, 0, -1, 4, -1, 0, -1, 0,0, 0, -1, 0, -1, 4, 0, 0, -1,
+               0, 0, 0, -1, 0, 0, 4, -1, 0,0, 0, 0, 0, -1, 0, -1, 4, -1,
+               0, 0, 0, 0, 0, -1, 0, -1, 4),9,9)
```

```
> b <- matrix(c(148,87,114,29,0,52,43,21,68),9,1)
> cbind(A,b) # matriz ampliada del sistema lineal A*x=b

      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10]
 [1,]    4   -1    0   -1    0    0    0    0    0   148
 [2,]   -1    4   -1    0   -1    0    0    0    0    87
 [3,]    0   -1    4    0    0   -1    0    0    0   114
 [4,]   -1    0    0    4   -1    0   -1    0    0    29
 [5,]    0   -1    0   -1    4   -1    0   -1    0     0
 [6,]    0    0   -1    0   -1    4    0    0   -1    52
 [7,]    0    0    0   -1    0    0    4   -1    0    43
 [8,]    0    0    0    0   -1    0   -1    4   -1    21
 [9,]    0    0    0    0    0   -1    0   -1    4    68

> temp <- solve(A,b) # solución del sistema lineal
> temp

          [,1]
 [1,] 63.43304
 [2,] 63.44643
 [3,] 56.41518
 [4,] 42.28571
 [5,] 46.93750
 [6,] 48.21429
 [7,] 29.77232
 [8,] 33.80357
 [9,] 37.50446

> max(temp) # temperatura máxima

[1] 63.44643

> min(temp) # temperatura mínima

[1] 29.77232

> T[2,2] <- temp[1]; T[2,3] <- temp[2]; T[2,4] <- temp[3]; # puntos interiores
> T[3,2] <- temp[4]; T[3,3] <- temp[5]; T[3,4] <- temp[6];
> T[4,2] <- temp[7]; T[4,3] <- temp[8]; T[4,4] <- temp[9];
> T[1,1] <- (T[1,2]+T[2,1]+T[2,2])/3; # esquinas
> T[1,5] <- (T[1,4]+T[2,5]+T[2,4])/3;
> T[5,1] <- (T[4,1]+T[5,2]+T[4,2])/3;
> T[5,5] <- (T[4,5]+T[5,4]+T[4,4])/3;
> T

         [,1]     [,2]     [,3]     [,4]     [,5]
[1,] 70.47768 95.00000 87.00000 63.00000 56.80506
[2,] 53.00000 63.43304 63.44643 56.41518 51.00000
[3,] 29.00000 42.28571 46.93750 48.21429 52.00000
[4,] 25.00000 29.77232 33.80357 37.50446 48.00000
[5,] 24.25744 18.00000 21.00000 20.00000 35.16815
```

3. Tenemos los datos de los porcentajes de voto en las últimas elecciones a rector de universidad. Si x,y,z corresponde al número de votos de cada candidato y T al total de votos, entonces el sistema lineal que hay que resolver es: $x = 20/100T$, $y = 27/100T$, $z = 53/100T$, $z = x+y+732$. Al sustituir las tres primeras ecuaciones en la cuarta, podemos calcular fácilmente el total de votos T = 12.200 y, de aquí, el número de votos de cada candidato: x = 2.440, y = 3.294 y z = 6.466.
El código R para resolver el problema es el siguiente:

```
> A <- matrix(c(1, 0, 0, -1/5,0, 1, 0, -27/100,0,
+               0, 1, -53/100,-1, -1, 1, 0), 4,4, byrow = T)
> b <- matrix(c(0,0,0,732), 4,1)
> cbind(A,b) # matriz ampliada del sistema lineal A*x=b

     [,1] [,2] [,3]  [,4] [,5]
[1,]    1    0    0 -0.20    0
[2,]    0    1    0 -0.27    0
[3,]    0    0    1 -0.53    0
[4,]   -1   -1    1  0.00  732

> solve(A,b) # solución del sistema lineal

      [,1]
[1,]  2440
[2,]  3294
[3,]  6466
[4,] 12200
```