Resolució de sistemes d'equacions lineals

Problemes per a la ciència de dades

PID_00262432

Francesc Pozo Montero
Jordi Ripoll Missé

Universitat
Oberta
de Catalunya

Francesc Pozo Montero

Llicenciat en Matemàtiques per la Universitat de Barcelona (2000) i doctor en Matemàtica Aplicada per la Universitat Politècnica de Catalunya (2005). Ha estat professor associat a la Universitat Autònoma de Barcelona i professor associat, col·laborador i actualment professor agregat a la Universitat Politècnica de Catalunya. A més, és cofundador del Grup d'Innovació Matemàtica E-learning (GIMEL), responsable de diversos projectes d'innovació docent i autor de diverses publicacions. Com a membre del grup de recerca consolidat CoDAlab, centra la recerca en la teoria de control i les aplicacions en enginyeria mecànica i civil, com també en l'ús de la ciència de dades per al monitoratge de la integritat estructural i per al monitoratge de la condició, sobretot en turbines eòliques.

Jordi Ripoll Missé

Llicenciat en Matemàtiques i doctor en Ciències Matemàtiques per la Universitat de Barcelona (2005). Professor col·laborador de la Universitat Oberta de Catalunya des del 2011 i professor del Departament d'Informàtica, Matemàtica Aplicada i Estadística de la Universitat de Girona (UdG) des del 1996, on actualment és professor agregat i desenvolupa tasques de recerca en l'àmbit de la biologia matemàtica (models amb equacions en derivades parcials i dinàmica evolutiva). També ha estat professor i tutor de la UNED en dues etapes, primer al centre associat de Terrassa i actualment al de Girona. Ha participat en nombrosos projectes d'innovació docent, especialment pel que fa a l'aprenentatge de les matemàtiques en línia.

L'encàrrec i la creació d'aquest recurs d'aprenentatge UOC han estat coordinats per la professora: Cristina Cano Bastidas (2019)

Primera edició: febrer 2019
© Francesc Pozo Montero, Jordi Ripoll Missé
Tots els drets reservats
© d'aquesta edició, FUOC, 2019
Av. Tibidabo, 39-43, 08035 Barcelona
Disseny: Manel Andreu
Realització editorial: Oberta UOC Publishing, SL

Cap part d'aquesta publicació, incloent-hi el disseny general i la coberta, no pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, tant si és elèctric com químic, mecànic, òptic, de gravació, de fotocòpia o per altres mètodes, sense l'autorització prèvia per escrit dels titulars del copyright.

Índex

1. Problemes

Resolució de sistemes d'equacions lineals (SEL)

Temes: *definició de SEL (coeficients, incògnites, paràmetres), nombre de solucions dels SEL, teorema de Rouché-Fröbenius, matriu de coeficients i matriu ampliada del sistema, SEL homogenis (amb solució única trivial o amb infinites solucions), mètode de resolució de Gauss, mètode de resolució de Cramer, interpretació geomètrica de les equacions lineals (rectes i plans) i de les solucions del sistema (intersecció).*

1. Una de les aplicacions més interessants de la teoria dels sistemes lineals en la ciència de dades és l'ajust d'un conjunt de dades per l'anomenada *recta de regressió*, que minimitza l'error entre les dades i la predicció del model lineal.

El problema es pot formular de la manera següent: disposem de dades d'observacions (sovint, en un nombre gran n) de dues característiques d'un conjunt d'individus: $(x_1,y_1),(x_2,y_2),\ldots,(x_n,y_n)$. A continuació construïm un model lineal $y = ax+b$, on a i b són dos paràmetres per determinar, que minimitzi la distància (al quadrat) entre les prediccions del model i els valors reals:

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

La solució d'aquest problema d'optimització consisteix a resoldre el següent sistema d'equacions lineals $\frac{\partial E}{\partial a} = 2\sum_{i=1}^{n}(ax_i + b - y_i)x_i = 0$, $\frac{\partial E}{\partial b} = 2\sum_{i=1}^{n}(ax_i + b - y_i) = 0$ per a les variables a i b, que, en dividir per n les dues equacions, es transforma en:

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

En introduir la notació per les mitjanes de les dades, per exemple, $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$, $\overline{x^2} = \frac{1}{n}\sum_{i=1}^{n} x_i^2$, etc., tenim de la segona equació $b = \bar{y} - a\bar{x}$ i en substituir la primera: $a(\overline{x^2} - \bar{x}^2) = \overline{xy} - \bar{x}\bar{y}$. Per tant, hem trobat els coeficients de la recta de regressió.

a) Calculeu l'equació de la recta de regressió $y = ax + b$ a partir de les $n = 8$ dades següents:

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

Un cop calculat l'ajust lineal, dibuixeu en un mateix gràfic les dades del problema i la recta obtinguda. S'ajusten bé les dades a la recta?

b) Disposem de les següents dades d'un laboratori, sobre dues característiques dels pacients analitzats:

$$X: \quad 4.17, 5.58, 5.18, 6.11, 4.50, 4.61, 5.17, 4.53, 5.33, 5.14$$

$$Y: \quad 4.81, 4.17, 4.41, 3.59, 5.87, 3.83, 6.03, 4.89, 4.32, 4.69$$

Feu un ajust lineal $y = ax + b$ entre les dues característiques X i Y. Dibuixeu en un mateix gràfic les dades (x_i, y_i), $i = 1 \ldots 10$, la recta de regressió i les prediccions del model $(x_i, \hat{y}_i = ax_i + b)$, $i = 1 \ldots 10$.

2. Imaginem que tenim les dades de la temperatura d'una placa metàl·lica quadrada al llarg del seu perímetre:

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

Nota de l'exercici 2

En aquest exercici hem vist com es pot resoldre el fenomen de la difusió de la calor, a partir d'unes dades i utilitzant la teoria dels sistemes lineals.

A partir d'aquestes dades, volem deduir la temperatura als punts de l'interior de la placa. Per trobar una aproximació a la temperatura en els punts interiors $T_{22}, T_{23}, \ldots, T_{44}$, podem suposar que per cada punt la temperatura és igual que la temperatura mitjana dels punts que l'envolten: a dalt, a baix, a la dreta i a l'esquerra.

a) A partir de les hipòtesis d'aquest model, escriviu el sistema d'equacions lineals per a la temperatura en cada un dels nou punts interiors. Escriviu la matriu ampliada del sistema.

b) Trobeu la solució del sistema lineal. És única? En quins punts interiors s'assoleixen les temperatures màxima i mínima?

c) Amb les solucions obtingudes podríeu dir quina és la temperatura en cadascuna de les quatre cantonades de la placa: T_{11}, T_{15}, T_{51} i T_{55}? (Mitjana dels tres valors que l'envolten.)

3. En les darreres eleccions a rector d'una universitat, els tres candidats van obtenir els següents percentatges de vots: 20%, 27% i 53%, respectivament. A més, el candidat guanyador va obtenir 732 vots més que el total dels seus rivals. Quants vots va obtenir cada candidat?

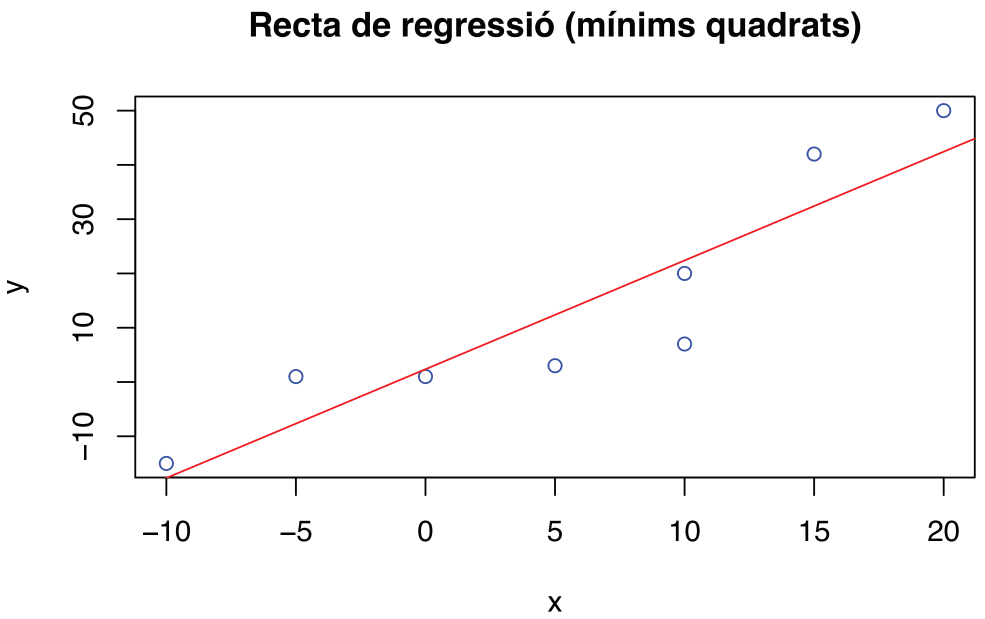
2. Solucions

1. a) A partir de les dades, calculem els coeficients del sistema lineal de dues equacions amb dues incògnites (a i b) i trobem la seva solució. La recta obtinguda és, doncs, $y = ax + b = 2.004329x + 2.350649$.

Això també ho podem fer directament amb el codi R següent:

```
> # Dades: dues característiques X i Y
> dades <- matrix(c(-10,-15,-5,1,0,1,5,3,10,20,10,7,15,42,20,50), 2,8)
> rownames(dades) <- c("X","Y")
> dades

  [,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 <- dades[1,]; y <- dades[2,]
> # Ajust a un model lineal (recta de regressió y=a*x+b)
> fit <- lm(y ~ x)
> summary(fit) # resum

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

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

(Intercept)           x
   2.350649    2.004329
```

```
> # Dibuix
> plot(x,y,col= "blue")     # dades (cercles)
> abline(fit,col = "red")   # recta
> title("Recta de regressió (mínims quadrats)")
```



Segons la gràfica, podem dir que les dades s'ajusten bastant bé a la recta.

b) De nou, a partir de les dades de l'exercici, calculem els coeficients del sistema lineal de dues equacions amb dues incògnites (a i b) i en trobem la solució. La recta obtinguda és, en aquest cas, $y = ax + b = -0.6229559x + 7.7957139$ i, a partir de l'equació de la recta, calculem les prediccions $\hat{y} = ax_i + b$, $i = 1,...,10$. Això també ho podem fer directament amb el codi R següent:

```
> # Dades:
> 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)
> # Ajust a un model lineal (recta de regressió y=a*x+b)
> fit <- lm(y ~ x)
> summary(fit) # resum

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
```

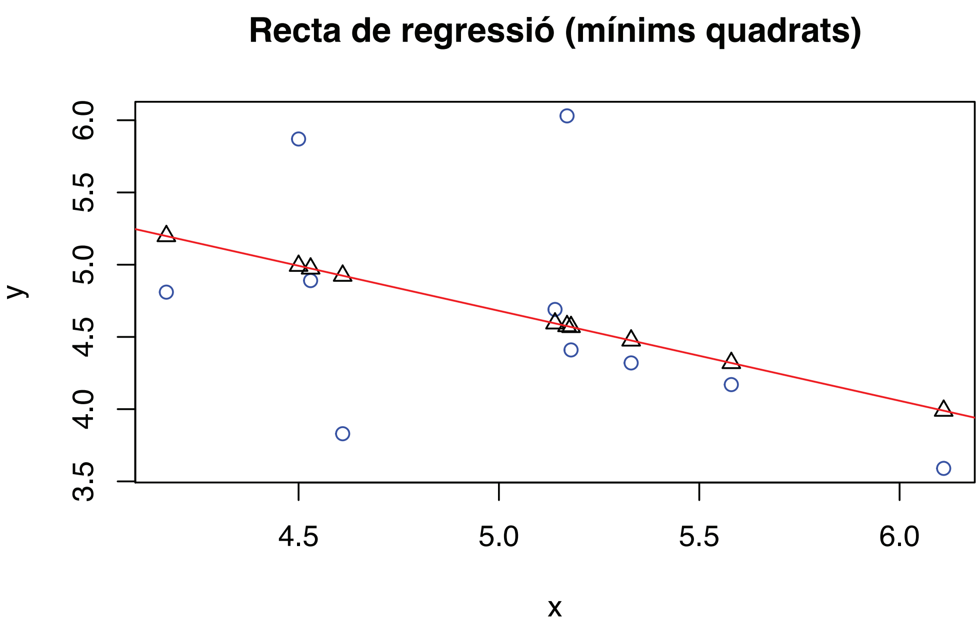
```
> # coeficients de la recta y=a*x+b, a=coef[2], b=coef[1]
> coef <- coefficients(fit); coef
```

```
(Intercept)           x
  7.7957139  -0.6229559
```

```
> yy <- fitted(fit) # valors de les prediccions
> 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
```

```
> # Dibuix
> plot(x,y,col= "blue")     # dades (cercles)
> points(x,yy, pch=2)       # prediccions (triangles)
> abline(fit,col = "red")   # recta
> title("Recta de regressió (mínims quadrats)")
```



2. a) Tenim les dades de la temperatura en els punts exteriors de la placa metàl·lica. Les equacions per a la temperatura en els nou punts interiors s'obtenen de la següent manera. Per al primer punt interior tenim: $T_{22} = (95+T_{32}+T_{23}+53)/4$, és a dir, $4T_{22}-T_{32}-T_{23} = 148$. Per al segon punt interior tenim: $T_{23} = (87 + T_{33} + T_{24} + T_{22})/4$, és a dir, $4T_{23} - T_{33} - T_{24} - T_{22} = 87$, etc. La matriu ampliada del sistema és la següent:

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

b) Aquest sistema es pot solucionar pel mètode de Gauss i la seva única solució és: $T_{22} = 63.43304$, $T_{23} = 63.44643$, $T_{24} = 56.41518$, $T_{32} = 42.28571$, $T_{33} = 46.93750$, $T_{34} = 48.21429$, $T_{42} = 29.77232$, $T_{43} = 33.80357$, $T_{44} = 37.50446$. La temperatura màxima s'assoleix a $T_{23} = 63.44643°$ i la mínima a $T_{42} = 29.77232°$.

c) En fer les mitjanes dels tres punts que envolten cadascuna de les quatre cantonades obtenim:
$T_{11} = 70.47768$, $T_{15} = 56.80506$, $T_{51} = 24.25744$ i $T_{55} = 35.16815$.
La resolució d'aquest problema es pot fer amb el codi R següent:

```
> 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) # matriu 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) # solució 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 màxima

[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]; # punts interiors
> 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; # cantonades
> 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. Tenim les dades dels percentatges de vot en les darreres eleccions a rector d'universitat. Si x,y,z correspon al nombre de vots de cada candidat i T al total de vots, aleshores el sistema lineal que cal resoldre és: $x = 20/100T$, $y = 27/100T$, $z = 53/100T$, $z = x+y+732$. En substituir les tres primeres equacions en la quarta, podem calcular fàcilment el total de vots $T = 12.200$ i, d'aquí, el nombre de vots de cada candidat: $x = 2.440$, $y = 3.294$ i $z = 6.466$. El codi R per resoldre el problema és el següent:

```
> 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) # matriu 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) # solució del sistema lineal

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