# Pràctica 7
# Mètode dels mínims quadrats

## Índex



## 1 Ajust per mínims quadrats

El terme *mínims quadrats* descriu el problema molt freqüent de resoldre sistemes d'equacions lineals sobre-determinats, és a dir, sistemes lineals amb més equacions que incògnites. En tal cas, en comptes de resoldre les equacions de manera exacta, com que normalment no existeix tal solució, se cerca solament una solució que faça $\mathsf{A}\vec{x}$ tan proper a $\vec{b}$ com siga possible.

- Una *solució per mínims quadrats* del sistema d'equacions lineals $\mathsf{A}\vec{x} = \vec{b}$ és un vector $\vec{x}_M$ en $\mathbb{R}^n$ tal que $\|\vec{b} - \mathsf{A}\vec{x}_M\| \leq \|\vec{b} - \mathsf{A}\vec{x}\|$ per a tot $\vec{x} \in \mathbb{R}^n$.

  Si apliquem el teorema de l'aproximació òptima al subespai $\mathrm{Col}\,\mathsf{A}$ i anomenem $\vec{b}_p$ a la projecció ortogonal de $\vec{b}$ sobre $\mathrm{Col}\,\mathsf{A}$, sabem que $\vec{b}_p$ és el vector de $\mathrm{Col}\,\mathsf{A}$ més pròxim a $\vec{b}$. Aleshores hi haurà un vector $\vec{x}_M$ que serà la solució de $\mathsf{A}\vec{x} = \vec{b}_p$. Tindrem, doncs, que $\mathsf{A}\vec{x}_M = \vec{b}_p$.

  Com que també sabem que $\vec{b} - \vec{b}_p \in (\mathrm{Col}\,\mathsf{A})^\perp$, aleshores $\vec{b} - \mathsf{A}\vec{x}_M \in (\mathrm{Col}\,\mathsf{A})^\perp = \mathrm{Nul}\,\mathsf{A}^t$, és a dir,
  $$\mathsf{A}^t(\vec{b} - \mathsf{A}\vec{x}_M) = 0,$$
  d'on
  $$\mathsf{A}^t\mathsf{A}\vec{x}_M = \mathsf{A}^t\vec{b}.$$

- L'equació matricial $\mathsf{A}^t\mathsf{A}\vec{x}_M = \mathsf{A}^t\vec{b}$ es coneix com *sistema d'equacions normals* per a $\mathsf{x}_M$.

- Si les columnes de A són linealment independents, aleshores la matriu $\mathsf{A}^t\mathsf{A}$ és invertible i la solució per mínims quadrats $\vec{x}_M = (\mathsf{A}^t\mathsf{A})^{-1}\mathsf{A}^t\vec{b}$ serà única.

- S'anomena *error de mínims quadrats* a $\|\vec{b} - \mathsf{A}\vec{x}_M\|$.

**Exemple 1.** Determinem les solucions per mínims quadrats de $\mathsf{A}\vec{x} = \vec{b}$, sent

$$\mathsf{A} = \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix}.$$

Hem de resoldre $\mathsf{A}^t\mathsf{A}\vec{x}_M = \mathsf{A}^t\vec{b}$.

$$\mathsf{A}^t\mathsf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix}, \qquad \mathsf{A}^t\vec{b} = \begin{bmatrix} 1 & 1 & 1 \\ 3 & -1 & 1 \end{bmatrix} \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}.$$

El sistema

$$\begin{bmatrix} 3 & 3 \\ 3 & 11 \end{bmatrix} \begin{bmatrix} x_{1M} \\ x_{2M} \end{bmatrix} = \begin{bmatrix} 6 \\ 14 \end{bmatrix}$$

té com a única solució

$$\begin{bmatrix} x_{1M} \\ x_{2M} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

L'error de mínims quadrats serà

$$\|\vec{b} - \mathsf{A}\vec{x}_M\| = \left\| \begin{bmatrix} 5 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 & 3 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\| = \sqrt{6}$$

Amb Scilab:

Podem trobar la solució per mínims quadrats de dues maneres: calculant la solució del sistema $\mathsf{A}^t\mathsf{A}\vec{x}_M = \mathsf{A}^t\vec{b}$ utilitzant **rref** (o \) o tenint en compte que en Scilab la comanda **x=A\b** proporciona una solució per mínims quadrats en el cas que A no siga una matriu quadrada. A més a més, si A és de rang complet per columnes, la solució serà única. Tanmateix, quan A no és de rang complet per columnes, la solució no és única.

```
-->A=[1 3;1 -1;1 1];b=[5;1;0];

-->Am=A'*A; bm=A'*b;

-->rref([Am,bm])
 ans  =

    1.    0.    1.
    0.    1.    1.

  -->Am\bm
```

```
 ans  =

    1.
    1.

-->A\b
 ans  =

    1.
    1.
```

i hem obtingut la mateixa solució per tots els mètodes perquè A és de rang complet per columnes:

```
-->rank(A)
ans  =

    2.
```

# 2 Aplicacions de l'ajust per mínims quadrats

En ciències i enginyeria, els experiments que es realitzen produeixen un conjunt de dades $(x_1, y_1)$, $(x_2, y_2), \dots,$ $(x_n, y_n)$ amb les abscisses diferents. El problema que es planteja és trobar una funció $y = f(x)$ que relacione les dades el millor possible. Cercar aquesta solució implicarà resoldre un problema de mínims quadrats.

## 2.1 Ajust de rectes

La relació més senzilla entre dues variables $x$ i $y$ és l'equació lineal $y = \beta_0 + \beta_1 x$. Sovint, en representar gràficament el conjunt de dades experimentals fa l'efecte que aquests queden a prop d'una recta. Volem determinar els paràmetres $\beta_0$ i $\beta_1$ que facen la recta tan «pròxima» als punts com siga possible.

La *recta de regressió per mínims quadrats* és aquella recta $y = \beta_0 + \beta_1 x$ que minimitza l'error residual.

Si les dades estiguessen sobre la recta, els paràmetres $\beta_0$ i $\beta_1$ satisfarien les equacions

$$\beta_0 + \beta_1 x_i = y_i \qquad \text{per a } i \in \{1, 2, \dots, n\}.$$

En forma matricial tindrem

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}.$$

Anomenarem *matriu de disseny* a

$$\mathsf{X} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \cdots & \cdots \\ 1 & x_n \end{bmatrix},$$

*vector paràmetre* a

$$\vec{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}$$

i *vector observació* a

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}$$

S'anomena *vector residual* a

$$\vec{\varepsilon} = \vec{y} - \mathsf{X}\vec{\beta}.$$

Per determinar el vector paràmetre n'hi ha prou amb resoldre per mínims quadrats el sistema $\mathsf{X}\vec{\beta} = \mathsf{y}$, és a dir, trobar la solució que minimitze la norma del vector residual.

**Exemple 2.** Trobarem la solució de l'equació $y = \beta_0 + \beta_1 x$ de la recta de mínims quadrats que s'ajuste als punts $(0,1)$, $(1,1)$, $(2,2)$ i $(3,2)$. Construïm la matriu de disseny X i calculem $\mathsf{X}^t\mathsf{X}$.

$$\mathsf{X} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \qquad \mathsf{X}^t\mathsf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix}.$$

Obtenim

$$\mathsf{X}^t\vec{y} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}$$

Resolem

$$\begin{bmatrix} 4 & 6 \\ 6 & 14 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 6 \\ 11 \end{bmatrix}, \quad \text{la solució del qual és} \quad \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 9/10 \\ 2/5 \end{bmatrix}$$

Per tant, la recta d'ajust per mínims quadrats serà

$$y = \frac{9}{10} + \frac{2}{5}x.$$

Amb Scilab:

```
-->X=[1 0;1 1;1 2;1 3],y=[1;1;2;2]
 X  =

    1.    0.
    1.    1.
    1.    2.
    1.    3.
 y  =

    1.
    1.
    2.
    2.

-->X1=X'*X
 X1  =

    4.    6.
    6.    14.

-->y1=X'*y
 y1  =

    6.
    11.

-->b=X1\y1
 b  =

    0.9
    0.4
```

Calculem ara la norma del vector residual $\|\vec{y} - \mathsf{X}\vec{\beta}\|$:

```
-->norm(y-X*b)
 ans  =

    0.4472136
```

## 2.2 Ajust de corbes per mínims quadrats

Si representem gràficament el conjunt de dades i observem que no són a prop d'una recta, cercarem la corba que millor els ajuste

$$y = \beta_0 + \beta_1 f_1(x) + \cdots + \beta_k f_k(x),$$

on $f_1$, $f_2, \ldots$, $f_n$ són funcions conegudes i $\beta_0$, $\beta_1, \ldots$, $\beta_n$ els paràmetres que hem de determinar per construir la corba. Procedint de manera anàloga a com ho hem fet en la regressió lineal, hem de trobar un vector paràmetre que faça la norma del vector residual mínima. En aquest cas, el sistema escrit en forma matricial serà

$$\begin{bmatrix} 1 & f_1(x_1) & f_2(x_1) & \ldots & f_k(x_1) \\ 1 & f_1(x_2) & f_2(x_2) & \ldots & f_k(x_1) \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 1 & f_1(x_n) & f_2(x_n) & \ldots & f_k(x_n) \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \ldots \\ \beta_k \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \ldots \\ y_n \end{bmatrix}.$$

La matriu d'aquest sistema és la matriu de disseny.

**Nota 1.** Si la funció no té terme independent, en construir la matriu de disseny prescindim de la primera columna, la dels uns.

**Exemple 3.** Amb les dades $(1, 1{,}8)$, $(2, 2{,}7)$, $(3, 3{,}4)$, $(4, 3{,}8)$, $(5, 3{,}9)$, determinarem el vector paràmetre, el vector residual i la corba de mínims quadrats associada amb les dades per a la funció $y = \beta_0 + \beta_1 x + \beta_2 x^2$.

1. Haurem de resoldre pel mètode de mínims quadrats el sistema d'equacions lineals $\vec{y} = \mathsf{X}\vec{\beta}$:

$$\begin{bmatrix} 1{,}8 \\ 2{,}7 \\ 3{,}4 \\ 3{,}8 \\ 3{,}9 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}.$$

2. Calcularem

$$\mathsf{X}^t\mathsf{X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 3 & 9 \\ 1 & 4 & 16 \\ 1 & 5 & 25 \end{bmatrix} = \begin{bmatrix} 5 & 15 & 55 \\ 15 & 55 & 225 \\ 55 & 225 & 979 \end{bmatrix}$$

$$\mathsf{X}^t\vec{y} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 5 \\ 1 & 4 & 9 & 16 & 25 \end{bmatrix} \begin{bmatrix} 1{,}8 \\ 2{,}7 \\ 3{,}4 \\ 3{,}8 \\ 3{,}9 \end{bmatrix} = \begin{bmatrix} 15{,}6 \\ 52{,}1 \\ 201{,}5 \end{bmatrix}.$$

3. Resoldrem $\mathsf{X}^t\mathsf{X}\vec{\beta} = \mathsf{X}^t\vec{y}$ i obtindrem el vector paràmetre

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} 0{,}58 \\ 1{,}34 \\ -0{,}136 \end{bmatrix}.$$

Per tant la funció serà $y = 0{,}58 + 1{,}34x - 0{,}136x^2$. Si ara calculem $\vec{\varepsilon} = \vec{y} - \mathsf{X}\vec{\beta}$, obtindrem el vector residual

$$\begin{bmatrix} 0{,}114 \\ -0{,}026 \\ 0{,}008 \\ 0{,}014 \\ 0{,}008 \end{bmatrix}.$$

Amb Scilab:

Introduïm les matrius X, $\vec{y}$:

```
-->X=[1 1 1;1 2 4;1 3 9;1 4 16;1 5 25], y=[1.8;2.7;3.4;3.8;3.9]
 X  =

    1.    1.    1.
    1.    2.    4.
    1.    3.    9.
    1.    4.    16.
    1.    5.    25.
 y  =

    1.8
    2.7
    3.4
    3.8
    3.9
```

Calculem $\mathsf{X}_1 = \mathsf{X}^t\mathsf{X}$, $\vec{y}_1 = \mathsf{X}^t\mathsf{y}$:

```
-->X1=X'*X
 X1  =

    5.     15.     55.
    15.    55.     225.
    55.    225.    979.

-->y1=X'*y
 y1  =

    15.6
    52.1
    201.5
```

Resolem $\mathsf{X}_1\vec{b} = \vec{y}_1$:

```
-->b=X1\y1
 b  =

    0.58
    1.3442857
  - 0.1357143
```

El vector residual serà

```
-->E=y-X*b
 E  =

    0.0114286
  - 0.0257143
    0.0085714
    0.0142857
  - 0.0085714
```

**Exemple 4.** Amb les dades de l'exemple anterior, determinarem els paràmetres que ajusten la funció $y = \beta_1 \cos(\pi x/3) + \beta_2 \sin(\pi x/3)$.

$$\mathsf{X} = \begin{bmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ \cos\frac{2\pi}{3} & \sin\frac{2\pi}{3} \\ \cos\frac{3\pi}{3} & \sin\frac{3\pi}{3} \\ \cos\frac{4\pi}{3} & \sin\frac{4\pi}{3} \\ \cos\frac{5\pi}{3} & \sin\frac{5\pi}{3} \end{bmatrix}$$

$$\mathsf{X}^t\mathsf{X} = \begin{bmatrix} \cos\frac{\pi}{3} & \cos\frac{2\pi}{3} & \cos\frac{3\pi}{3} & \cos\frac{4\pi}{3} & \cos\frac{5\pi}{3} \\ \sin\frac{\pi}{3} & \sin\frac{2\pi}{3} & \sin\frac{3\pi}{3} & \sin\frac{4\pi}{3} & \sin\frac{5\pi}{3} \end{bmatrix} \begin{bmatrix} \cos\frac{\pi}{3} & \sin\frac{\pi}{3} \\ \cos\frac{2\pi}{3} & \sin\frac{2\pi}{3} \\ \cos\frac{3\pi}{3} & \sin\frac{3\pi}{3} \\ \cos\frac{4\pi}{3} & \sin\frac{4\pi}{3} \\ \cos\frac{5\pi}{3} & \sin\frac{5\pi}{3} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

$$\mathsf{X}^t\vec{y} = \begin{bmatrix} \cos\frac{\pi}{3} & \cos\frac{2\pi}{3} & \cos\frac{3\pi}{3} & \cos\frac{4\pi}{3} & \cos\frac{5\pi}{3} \\ \sin\frac{\pi}{3} & \sin\frac{2\pi}{3} & \sin\frac{3\pi}{3} & \sin\frac{4\pi}{3} & \sin\frac{5\pi}{3} \end{bmatrix} \begin{bmatrix} 1{,}8 \\ 2{,}7 \\ 3{,}4 \\ 3{,}8 \\ 3{,}9 \end{bmatrix} = \begin{bmatrix} -3{,}8 \\ -2{,}8 \end{bmatrix}$$

Resolem el sistema i obtenim el vector paràmetre.

$$\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} -3{,}8 \\ -2{,}8 \end{bmatrix}, \quad \text{la solució del qual és} \quad \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} = \begin{bmatrix} -1{,}9 \\ -0{,}9 \end{bmatrix}.$$

Per tant, la funció demanada serà $y = -1{,}9\cos(\pi x/3) - 0{,}9\sin(\pi x/3)$, i el vector residual és

$$\vec{\varepsilon} = \vec{y} - \mathsf{X}\vec{\beta} = \begin{bmatrix} 3{,}5 \\ 2{,}5 \\ 1{,}5 \\ 2 \\ 4 \end{bmatrix}.$$

Observem que el vector residual és menor quan hem fet servir la primera funció per ajustar les dades que quan fem servir la segona.

Amb Scilab:

Introduïm la matriu de disseny X:

```
-->X=[cos(%pi/3),   sin(%pi/3)
-->   cos(%pi*2/3), sin(%pi*2/3)
-->   cos(%pi*3/3), sin(%pi*3/3)
-->   cos(%pi*4/3), sin(%pi*4/3)
-->   cos(%pi*5/3), sin(%pi*5/3)]
 X  =

    0.5     0.8660254
  - 0.5     0.8660254
  - 1.      1.225D-16
  - 0.5   - 0.8660254
    0.5   - 0.8660254
```

Calculem $\mathsf{X}^t\mathsf{X}$ i $\mathsf{X}^t\vec{y}$:

```
-->X1=X'*X
 X1  =

    2.            2.924D-16
    2.924D-16     3.

-->y1=X'*y
 y1  =

  - 3.8
  - 2.7712813
```

Resolem $\mathsf{X}_1\vec{b} = \vec{y}_1$:

```
-->b1=X1\y1
 b1  =

  - 1.9
  - 0.9237604
```

La segona component no és la mateixa que la d'abans per errors d'arrodoniment degudes a l'ús de les funcions trigonomètriques.

Per determinar el vector residual fem

```
-->E=y-X*b1
 E  =

    3.55
    2.55
    1.5
    2.05
    4.05
```