Problemas de Mínimos Cuadrados

- Ajuste de curvas.
- Problemas de cuadrados mínimos no lineales: Reducción a problemas lineales.

Ajuste de polinomios.

Dado un conjunto de puntos

$$(x_1, y_1), \ldots, (x_m, y_m)$$

nos proponemos encontrar el polinomio

$$p(x) = c_0 + c_1 x + \ldots + c_{n-1} x^{n-1},$$

con $n < m$ que esté **más cerca** de estos puntos en el sentido que

$$\sum_{i=1}^{m} |p(x_i) - y_i|^2,$$

sea mínima.

Esta suma de cuadrados es el cuadrado de la norma del residuo del sistema rectangular:

$$\begin{pmatrix} 1 & x_1 & & x_1^{n-1} \\ 1 & x_2 & \dots & x_2^{n-1} \\ 1 & x_3 & \dots & x_3^{n-1} \\ & \vdots & & \vdots \\ 1 & x_m & \dots & x_m^{n-1} \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \\ \vdots \\ c_{n-1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_m \end{pmatrix}.$$

Considere un sistema de ecuaciones

$$\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$$

donde $\mathbf{A} \in \mathbb{R}^{m\times n}$, con $n < m$, es una matriz rectangular de m filas y n columnas y $\boldsymbol{b} \in \mathbb{R}^m$.

Este problema, en general, no tiene solución: **sistema sobredeterminado**.

Una alternativa es buscar una solución en el sentido generalizado siguiente:

Hallar $\boldsymbol{x} \in \mathbb{R}^n$ tal que $\|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}\|_2$ **sea mínima**.

Definición. El vector $\boldsymbol{x}$ que minimiza $\|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}\|_2$ es la **solución en el sentido de mínimos cuadrados** del sistema rectangular.

Ojo! En general,

$$\mathbf{A}\boldsymbol{x} \neq \boldsymbol{b}.$$

Teorema. *Sean* $\mathbf{A} \in \mathbb{R}^{m \times n} (m \geq n)$ *y* $\boldsymbol{b} \in \mathbb{R}^m$. *Un vector* $\boldsymbol{x} \in \mathbb{R}^n$ *minimiza la norma del residuo* $\|\boldsymbol{r}\|_2 = \|\boldsymbol{b} - \mathbf{A}\boldsymbol{x}\|_2$ *si y sólo si el residuo* $\boldsymbol{r}$ *es ortogonal a la imagen de* $\mathbf{A}$*; esto es*

$$\mathbf{A}^t \boldsymbol{r} = 0,$$

donde $\mathbf{A}^t$ *es la matriz transpuesta de* $\mathbf{A}$.

Consecuencia: $\boldsymbol{x}$ debe satisfacer

$$\mathbf{A}^t \boldsymbol{r} = 0 \iff \mathbf{A}^t (\boldsymbol{b} - \mathbf{A}\boldsymbol{x}) = 0 \iff \mathbf{A}^t \mathbf{A}\boldsymbol{x} = \mathbf{A}^t b.$$

Estas últimas reciben el nombre de **ecuaciones normales**.

Observaciones

- En el caso en que $m = n$ y que la matriz $\mathbf{A}$ sea no singular, entonces las ecuaciones normales entregan como solución la solución del sistema lineal $\mathbf{A}\boldsymbol{x} = \boldsymbol{b}$.
- Las ecuaciones normales tienen solución única si y sólo si todas las columnas de $\mathbf{A}$ son l.i.; es decir, $rango(A) = n$.

Ejemplo 1.
Un problema de aproximación polinomial.

Considere la siguiente tabla de valores:

x	y
0.0	10.5000
0.5	5.4844
1.0	0.0000
1.5	-3.6094
2.0	-4.5000
2.5	-2.9531
3.0	0.0000
3.5	2.9531
4.0	4.5000
4.5	3.6094
5.0	0.0000

Se pide ajustar estos datos en el sentido de mínimos cuadrados por un polinomio de grado 3.

Solución

Nuestro problema se reduce a encontrar constantes a, b, c y d para formar el polinomio

$$p(x) = a\,x^3 + b\,x^2 + c\,x + d.$$

Evaluemos el polinomio p en los diferentes valores de x de la tabla, obteniendo así el sistema lineal rectangular con incógnitas a, b, c y d:

$$
\underbrace{\begin{pmatrix}
0.0000 & 0.0000 & 0.0000 & 1.0000 \\
0.1250 & 0.2500 & 0.5000 & 1.0000 \\
1.0000 & 1.0000 & 1.0000 & 1.0000 \\
3.3750 & 2.2500 & 1.5000 & 1.0000 \\
8.0000 & 4.0000 & 2.0000 & 1.0000 \\
15.6250 & 6.2500 & 2.5000 & 1.0000 \\
27.0000 & 9.0000 & 3.0000 & 1.0000 \\
42.8750 & 12.2500 & 3.5000 & 1.0000 \\
64.0000 & 16.0000 & 4.0000 & 1.0000 \\
91.1250 & 20.2500 & 4.5000 & 1.0000 \\
125.0000 & 25.0000 & 5.0000 & 1.0000
\end{pmatrix}}_{\mathbf{A}}
\underbrace{\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}}_{\boldsymbol{x}}
=
\underbrace{\begin{pmatrix}
10.5000 \\ 5.4844 \\ 0.0000 \\ -3.6094 \\ -4.5000 \\ -2.9531 \\ 0.0000 \\ 2.9531 \\ 4.5000 \\ 3.6094 \\ 0.0000
\end{pmatrix}}_{\boldsymbol{b}}
$$

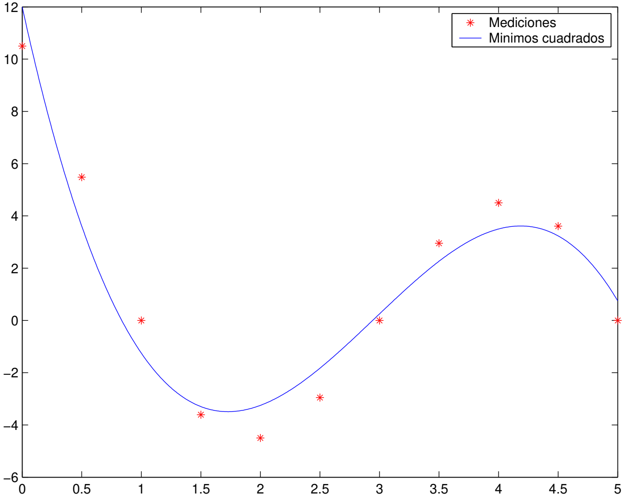
es decir,

$$\mathbf{A}\,\boldsymbol{x} = \boldsymbol{b}.$$

Para resolver multiplicamos por $\mathbf{A}^t$ obteniendo las ecuaciones normales

$$\underbrace{\begin{pmatrix} 30913 & 6901 & 1583 & 378 \\ 6901 & 1583 & 378 & 96 \\ 1583 & 378 & 96 & 27 \\ 378 & 96 & 27 & 11 \end{pmatrix}}_{\mathbf{A}^t\mathbf{A}} \underbrace{\begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{pmatrix} 649.8809 \\ 138.0586 \\ 25.5234 \\ 15.9844 \end{pmatrix}}_{\mathbf{A}^t\boldsymbol{y}}$$

De aquí, el polinomio es: $p(x) = -0.9583x^3 + 8.5x^2 - 20.7917x + 12.$



Ejemplo 2. Un problema no lineal reducible a lineal.

Considere la siguiente tabla de valores

x	0.0000	0.4000	0.8000	1.2000	1.6000	2.0000
y	3.1437	4.4169	6.0203	8.6512	11.0078	16.2161

Se quiere ajustar una función

$$f(x) = ae^{bx}.$$

a estos datos en el sentido de mínimos cuadrados.

Solución

Tomando logaritmos se transforma en un problema lineal de cuadrados mínimos:

$$z = \ln(y) = \ln(f(x)) = \ln(a) + bx.$$

x	0.0000	0.4000	0.8000	1.2000	1.6000	2.0000
$z = \ln(y)$	1.1454	1.4854	1.7951	2.1577	2.3986	2.7860

$$\underbrace{\begin{pmatrix} 1 & 0.0000 \\ 1 & 0.4000 \\ 1 & 0.8000 \\ 1 & 1.2000 \\ 1 & 1.6000 \\ 1 & 2.0000 \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} \ln(a) \\ b \end{pmatrix}}_{\boldsymbol{x}} = \underbrace{\begin{pmatrix} 1.1454 \\ 1.4854 \\ 1.7951 \\ 2.1577 \\ 2.3986 \\ 2.7860 \end{pmatrix}}_{\boldsymbol{b}}.$$

Resolvemos el problema de cuadrados mínimos

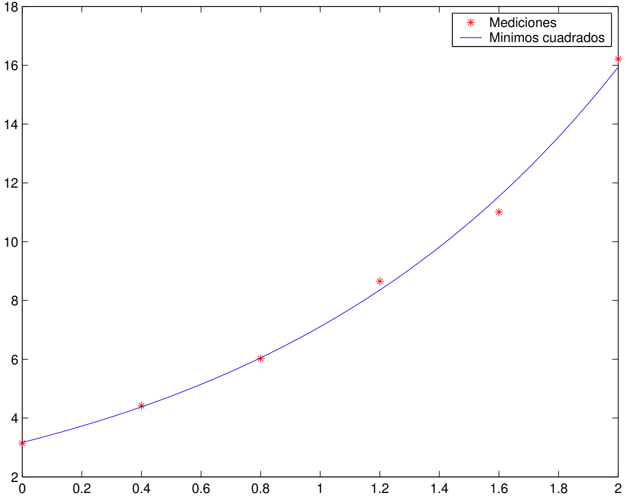
$$(\mathbf{A}^t \mathbf{A})\boldsymbol{x} = \mathbf{A}^t \boldsymbol{b}$$

y obtenemos

$$\ln(a) = 1.1539 \implies a = e^{1.1539} = 3.1705 \qquad \text{y} \qquad b = 0.8075$$

Por lo tanto

$$f(x) = 3.1705e^{0.8075x}$$



Otros ejemplos de modelos no lineales reducibles a lineales

- $f(t) = c\,e^{at-bt^2}$: En este caso, se aplica logaritmo para obtener:

$$\ln(f(t)) = \ln(c) + a\,t - bt^2$$

- $f(t) = \dfrac{a}{b+t}$: En este caso se toman los recíprocos para obtener

$$\frac{1}{f(t)} = \frac{b}{a} + \frac{1}{a}t$$

- $f(t) = \dfrac{1}{1+ae^{-ct}}$. En este caso se toman recíprocos, se resta 1 y después se aplica logaritmo para obtener:

$$\ln\left(\frac{1}{f(t)} - 1\right) = \ln(a) - c\,t.$$