Tema 5 (III) - El método de aproximación por mínimos cuadrados

Resumen: Dado un sistema incompatible $A\mathbf{x} = \mathbf{b}$, esto es $\mathbf{b} \notin \text{col}(A)$, el *Teorema de la Proyección* asegura que el vector de $\text{col}(A)$ más cercano al vector $\mathbf{b}$ es $\mathsf{P}_{\text{col}(A)}(\mathbf{b})$. La aproximación por mínimos cuadrados consiste en cambiar el sistema original, que es incompatible, por el sistema $A\mathbf{x} = \mathsf{P}_{\text{col}(A)}(\mathbf{b})$ que es compatible y es "*el más próximo*" al sistema original. Hay, en la práctica, dos formas equivalentes de proceder:

(1) - Abordaje "*directo*", esto es, hallar $\mathsf{P}_{\text{col}(A)}(\mathbf{b})$ y, a continuación, resolver el sistema $A\mathbf{x} = \mathsf{P}_{\text{col}(A)}(\mathbf{b})$.

(2) - Resolver las *ecuaciones normales* $(A^tA)\mathbf{x} = A^t\mathbf{b}$.

1. En cada caso, probar que el sistema $A\mathbf{x} = \mathbf{b}$ es incompatible y hallar la "*mejor*" solución aproximada usando los procedimientos (1) y (2) descritos en el resumen. Comparar la eficiencia de los métodos.

$$A = \begin{pmatrix} 1 & 2 \\ -1 & 0 \\ 0 & -2 \end{pmatrix} \mathbf{b} = \begin{pmatrix} -2 \\ 2 \\ -3 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 1 \\ -1 & 0 & -1 \end{pmatrix} \mathbf{b} = \begin{pmatrix} 2 \\ 3 \\ -2 \end{pmatrix} \quad A = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 1 \end{pmatrix} \mathbf{b} = \begin{pmatrix} 3 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

2. Los salarios medios (en millones de \$) pagados en la *Mayor League Baseball* desde 1990 a 1999 son:

Año	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Salario	0,60	0,85	1,03	1,08	1,17	1,11	1,12	1,34	1,40	1,61

a) Representar la tabla como una nube de puntos (empezar con $x_0 = 0$ para 1990) y comprobar que una recta serviría como modelo "*razonable*" para aproximar esta nube de puntos.

b) Usar el método de los mínimos cuadrados para ajustar una recta (recta de regresión) a la nube.

c) Estimar cuáles han sido los salarios medios desde 2000 a 2010. Contrastar la estimación con la fuente original: http://www.mlbplayers.com/pdf9/4923609.pdf

3. Encontrar la recta de regresión para los conjuntos de puntos siguientes:

a) $(1,1)$, $(2,2)$, $(3,2)$, $(4,3)$

b) $(-3,-3)$, $(-2,-2)$, $(0,0)$, $(1,2)$

4. Encontrar la parábola que mejor se ajuste, en el sentido de los mínimos cuadrados, a la nube de puntos:

a) $(-1,1)$, $(0,-1)$, $(1,0)$, $(2,2)$

b) $(-2,4)$, $(-1,7)$, $(0,3)$, $(1,0)$, $(2,-1)$

5. Un investigador tiene una serie de parejas de datos experimentales (x_i, y_i) de los que se supone que, aproximadamente, será $y_i = p(x_i)$ siendo p un polinomio de segundo grado. Por ensayo y error, conjetura que $p(x) = x^2 - 4x + 1$ ¿Puede mejorarse la conjetura? En la tabla siguiente se resumen los datos.

x_i	1	2	3	4	5
y_i DATOS REALES	$-1,8$	$-2,9$	$-2,1$	$1,1$	$5,9$
$\hat{y}_i = x_i^2 - 4x_i + 1$	-2	-3	-2	1	6

6. Probar que, para toda matriz $A_{m\times n}$, se tiene que $\text{rango}(A) = \text{rango}(A^tA)$. Indicaciones:

a) Probar que si $\mathbf{x} \in \mathbb{R}^n$ es tal que $A\mathbf{x} = \vec{\mathbf{0}}$, entonces $A^tA\mathbf{x} = \vec{\mathbf{0}}$ lo cual, implica que $\text{nulo}(A) \subseteq \text{nulo}(A^tA)$.

b) Viceversa: $A^tA\mathbf{x} = \vec{\mathbf{0}} \Rightarrow \mathbf{x}^tA^tA\mathbf{x} = 0 \Rightarrow (A\mathbf{x})^t(A\mathbf{x}) = \|A\mathbf{x}\|^2 = 0 \Rightarrow A\mathbf{x} = \vec{\mathbf{0}} \Rightarrow \text{nulo}(A^tA) \subseteq \text{nulo}(A)$.

c) Aplicar ahora el hecho de que $\text{rango}(A) + \dim(\text{nulo}(A)) = n$.

7. Si las columnas de A son linealmente independientes, podemos usar las ecuaciones normales para hallar la matriz de la proyección $\mathsf{P}_{\text{col}(A)}$ sin encontrar una base ortonormal de $\text{col}(A)$. Pasos a seguir:

a) Probar que para todo $\mathbf{b} \in \mathbb{R}^n$, la solución del sistema $A\hat{\mathbf{x}} = \mathsf{P}_{\text{col}(A)}(\mathbf{b})$ es $\hat{\mathbf{x}} = (A^tA)^{-1}A^t\mathbf{b}$. $(*)$

b) Usar el hecho anterior para probar que la matriz de $\mathsf{P}_{\text{col}(A)}$ es $A(A^tA)^{-1}A^t$.

c) Para la matriz A dada, hallar la matriz de $\mathsf{P}_{\text{col}(A)}$ de dos formas diferentes. $A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}$

$(*)$ - **Nota:** A la matriz $(A^tA)^{-1}A^t$ se le denomina *pseudoinversa* de A.

8. Hallar $\min\left\{(2-x+y)^2 + (1+x+y)^2 + (3-x-2y)^2 \mid x, y \in \mathbb{R}\right\}$