PRACTICA 4b

Aproximación discreta de mínimos cuadrados.

Aproximación discreta por mínimos cuadrados. En general, los problemas que aparecen en la ciencia nos enfrentan a la observación de cantidades que cambian en el tiempo y/o el espacio. Supongamos que este cambio puede ser modelado por una relación matemática $y(x) = f(c_1, \dots, c_n; x)$ donde x es el parámetro que describe el cambio y $c_1, \dots, c_n$ son n cantidades desconocidas, llamadas *parámetros del modelo*, cuyos valores queremos determinar. En particular consideremos el caso en que el modelo es *lineal* en sus parámetros, esto es, la relación funcional f es lineal respecto de los parámetros $c_1, \dots, c_n$ y por lo tanto puede expresarse como

$$f(c_1, \dots, c_n; x) = \sum_{j=1}^{n} c_j \phi_j(x)$$

para n funciones $\phi_j(x)$ del parámetro x. Un conjunto de m observaciones de f proveerá de valores medidos y_i afectados de errores ϵ_i

$$\begin{aligned} y_i &= y(x_i) + \epsilon_i \\ &= \sum_{j=1}^{n} \phi_j(x_i)\, c_j + \epsilon_i, \quad i = 1, 2, \dots, m. \end{aligned}$$

Definiendo la *matriz de diseño*, $m \times n$, A de elementos $a_{ij} = \phi_j(x_i)$, el *vector de parámetros*, $n \times 1$, $\mathbf{x}$ de elementos c_i, el *vector de medidas*, $m \times 1$, $\mathbf{b}$ de elementos y_i, el *vector de residuos*, $m \times 1$, $\mathbf{r}$ de elementos ϵ_i, el conjunto de ecuaciones anteriores puede escribirse matricialmente como

$$A\,\mathbf{x} + \mathbf{r} = \mathbf{b}.$$

Esta relación constituye nuestro modelo lineal de las observaciones. Si $m \geq n$ (más observaciones que parámetros) se trata ahora determinar el conjunto de parámetros $\mathbf{x}$ que satisfaga mejor, en algún sentido, el modelo lineal. De acuerdo al *método de mínimos cuadrados*, los estimadores de mínimos cuadrados de los parámetros son aquellos valores que minimizan la norma euclideana de los residuos

$$\|\mathbf{r}\|_2 = \|\mathbf{b} - A\,\mathbf{x}\|_2 = \left(\mathbf{r}^t \mathbf{r}\right)^{1/2}$$

Este requisito conduce a que la solución de mínimos cuadrados debe satisfacer las *ecuaciones normales*

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

de donde se sigue que, cuando $A^t A$ es no-singular, la solución de mínimos cuadrados existe y es única. La condición *necesaria y suficiente* para la existencia de una solución única es que las columnas de la matriz A sean linealmente independientes o, dicho en forma equivalente, que el rango de la matriz A sea n. En tal caso, es fácil ver que la matriz $A^t A$ es simétrica y definida positiva[1], con lo cual un método numérico apropiado para resolver las ecuaciones normales es el *método de Choleski*. Sin embargo los problemas que se presentan en la práctica conducen a un sistema de ecuaciones normales con una matriz mal condicionada. Por tal motivo resulta fundamental recurrir a otra forma de resolución numérica del problema. En particular el *método QR* evita la formación de las ecuaciones normales y garantizan la estabilidad de la solución frente a los errores de redondeo. Este método se basa en la existencia de la factorización QR de la matriz A $m \times n$, cuando la misma es de rango n,

$$A = QR,$$

siendo Q es una matriz $m \times n$ cuyas columnas son una base ortonormal del espacio columna de A (con lo cual $Q^t Q = I_n$) y R es una matriz cuadrada de orden n triangular superior con elementos positivos sobre la diagonal. En tal caso las ecuaciones normales toman la forma

$$(QR)^t (QR)\,\mathbf{x} = (QR)^t\,\mathbf{b},$$

es decir

$$R^t R\,\mathbf{x} = R^t Q^t\,\mathbf{b}$$

y como R es no-singular, se obtiene $R\,\mathbf{x} = Q^t\,\mathbf{b}$. Así, conocida la factorización QR de A, la solución de mínimos cuadrados resulta, pues, por sustitución hacia atrás del sistema triangular superior

$$R\,\mathbf{x} = \hat{\mathbf{b}}, \qquad \hat{\mathbf{b}} = Q^t\,\mathbf{b}.$$

La subrutina `LA_GELS` de la biblioteca de rutinas de `LAPACK95` efectúa este procedimiento para resolver el problema de mínimos cuadrados.

☞ La teoría anterior puede ser aplicada en particular al caso en que la relación funcional del modelo lineal sea un polinomio de grado a lo más $(n-1)$:

$$f(c_1, \dots, c_n; x) = \sum_{j=1}^{n} c_j x^{j-1} = c_1 + c_2 x + \cdots + c_n x^{n-1}.$$

En este caso la matriz $\mathbf{A}$ tiene elementos $a_{ij} = x_i^{j-1}$. Los estimadores de mínimos cuadrados conducen así al *ajuste de los datos observacionales por un polinomio de mínimos cuadradados.*

Ejercicio 1. Mostrar que la recta $y = c_1 + c_2\,x$ que ajusta por mínimos cuadrados (llamada *recta de regresión*) un conjunto de m datos (x_i, y_i), $i = 1, 2 \dots, m$, tiene por coeficientes

$$c_1 = \frac{\sum_{i=1}^{m} x_i^2 \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} x_i y_i \sum_{i=1}^{m} x_i}{m \sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2},$$

$$c_2 = \frac{m \sum_{i=1}^{m} x_i y_i - \sum_{i=1}^{m} x_i \sum_{i=1}^{m} y_i}{m \sum_{i=1}^{m} x_i^2 - \left(\sum_{i=1}^{m} x_i\right)^2}.$$

[1] Esto es, $\mathbf{x}^t A \mathbf{x} > 0$ para todo $\mathbf{x} \neq 0$.

Ejercicio 2. Escribir un programa Fortran que calcule los coeficientes del polinomio de mínimos cuadrados de grado g que ajusta a un conjunto dado de datos (x_i, y_i), $i = 1, \ldots, m$ $(m \geq n = g+1)$. *Indicación:* Construya la matriz de diseño del problema y utilice la subrutina `LA_GELS` de `LAPACK95` para resolver el problema de mínimos cuadrados correspondiente.

Ejercicio 3. Determine los posibles polinomios de mínimos cuadrados de distintos órdenes para el siguiente conjunto de datos. En base a sus resultados, ¿puede decidir cual es el polinomio apropiado para representar el comportamiento de los datos experimentales?

i	x_i	y_i
1	-0.9	81.0
2	-0.7	50.0
3	-0.5	35.0
4	-0.3	27.0
5	-0.1	26.0
6	0.1	60.0
7	0.3	106.0
8	0.5	189.0
9	0.7	318.0
10	0.9	520.0

Ejercicio 4. El nivel de agua del Mar del Norte está determinado fundamentalmente por la denominada marea M_2 cuyo período es de alrededor de 12 horas y su expresión puede ser aproximada por una función periódica en el tiempo, de la forma:

$$H(t) = h_0 + a_1 \sin\frac{2\pi t}{12} + a_2 \cos\frac{2\pi t}{12}$$

con t medido en horas. Hallar los parámetros de este modelo que ajustan por mínimos cuadrados las siguientes mediciones.

i	1	2	3	4	5	6
t_i	0	2	4	6	8	10
H_i	1.0	1.6	1.4	0.6	0.2	0.8

Mostrar que, en este caso particular, al formar las ecuaciones normales $A^t A$ es una matriz diagonal y por lo tanto los parámetros del modelo pueden ser fácilmente calculados.

Ejercicio 5. Considere un ajuste de la forma $y = b\, e^{ax}$ para el conjunto de medidas dada en la siguiente tabla.

x	1.00	1.25	1.50	1.75	2.00
y	5.10	5.79	6.53	7.45	8.46

Se quiere determinar los coeficientes a y b que mejor ajusten a los datos.

a) Explique porqué no puede aplicar directamente el método de mínimos cuadrados al ajuste propuesto.

b) Considerando el logaritmo de la ecuación propuesta, determine los coeficientes a y b a partir de la recta de mínimos cuadrados que ajusta al logaritmo de las medidas.

c) Utilice la rutina `fit` del programa `gnuplot` para determinar el modelo no lineal original del problema. Compare el resultado con el obtenido en el punto anterior.