Clase 10 - Aproximación de funciones

El estudio de aproximación de funciones o, mejor dicho, de la teoría de aproximación comprende dos tipos de problemas: i) la búsqueda de parámetros óptimos de un modelo funcional propuesto que represente un conjunto de datos dados; y ii) dada una función de manera explícita, se desea encontrar un tipo o modelo más sencillo para representarla, por ejemplo un polinomio, que permita estimar valores funcionales de una manera más simple.

Aproximación discreta por cuadrados mínimos

Consideremos el problema de estimar una función desconocida a través de un conjunto de datos experimentales: (x_i, y_i), $i = 1, \ldots, m$.

Por ejemplo, consideremos los siguientes datos

x_i	1	2	3	4	5	6	7	8	9	10
y_i	1.3	3.5	4.2	5.0	7.0	8.8	10.1	12.5	13.0	15.6

Podemos representarlos gráficamente

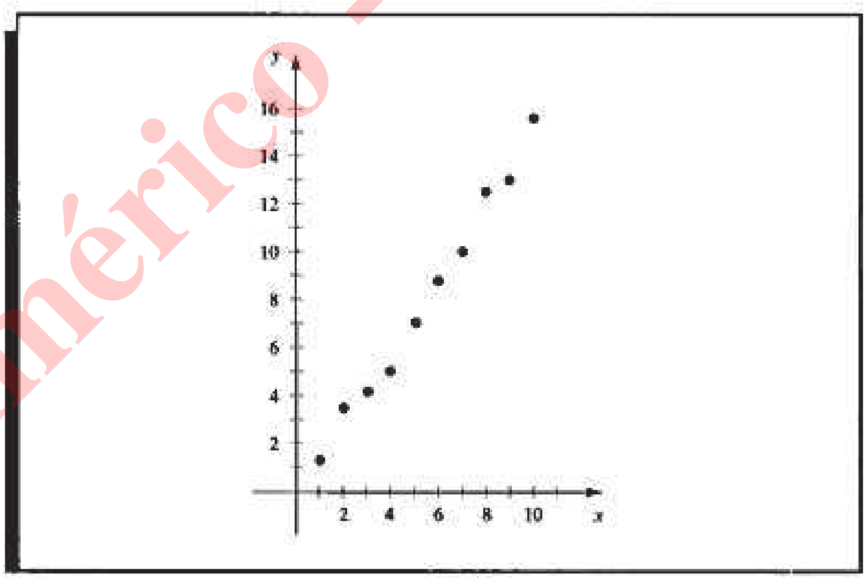

Figura 1: Datos medidos

Si se construyera un polinomio interpolante usando los m puntos se obtendría un polinomio de grado alto y ya se mencionó que esto no es conveniente. Para el caso particular de los datos del ejemplo dado donde se tienen 10 puntos se obtendría un polinomio interpolante de grado 9. Ver Figura (2).

Sin embargo, no es razonable exigir que la función pase exactamente por todos esos puntos. La Figura (1) sugiere que la relación entre x e y es lineal (polinomio de grado 1). El motivo que no exista una recta que se ajuste a estos datos, es decir que pase por todos los puntos, es que estos datos tienen errores.

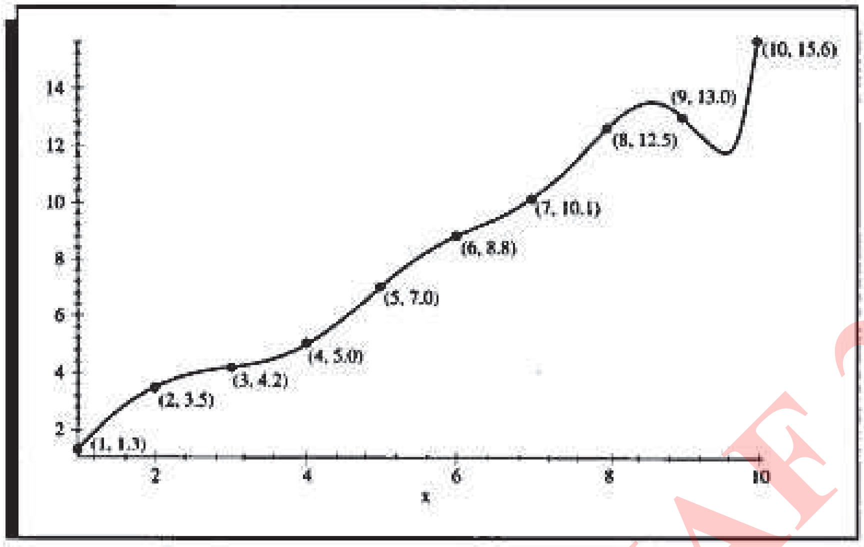

Figura 2: Polinomio de grado 9

Sería mejor pensar en un modelo lineal y encontrar la *mejor* recta (en cierto sentido), aún cuando no coincida con los datos en ningún punto. El problema ahora será determinar cuál es la mejor recta

$$y(x) = a_1 x + a_0$$

que *mejor ajusta* a esos datos, esto es, determinar los coeficientes a_1 y a_0 óptimos.

Una idea posible es determinar a_0 y a_1 tales que minimicen la función

$$E_\infty(a_0, a_1) = \max_{1 \le i \le 10} |y_i - (a_1 x_i + a_0)|.$$

Esto es conocido como el *problema minimax*.

Otra alternativa es determinar a_0 y a_1 tales que minimicen la función

$$E_1(a_0, a_1) = \sum_{i=1}^{10} |y_i - (a_1 x_i + a_0)|.$$

Esa función es conocida como *desviación absoluta*.

Por último, el **método de cuadrados mínimos** para ajustar a una recta con m datos consiste en determinar a_0 y a_1 tales que minimicen la función

$$E(a_0, a_1) = E_2(a_0, a_1) = \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2$$

con respecto a las variables a_0 y a_1. En el ejemplo anterior $m = 10$.

Una condición necesaria para tener un mínimo es que las derivadas parciales de E con respecto a a_0 y a_1 deben ser cero, esto es,

$$\begin{aligned}
\frac{\partial}{\partial a_0} E(a_0, a_1) &= \frac{\partial}{\partial a_0} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2 &= 2 \sum_{i=1}^{m} (y_i - a_1 x_i - a_0)(-1) &= 0, \\
\frac{\partial}{\partial a_1} E(a_0, a_1) &= \frac{\partial}{\partial a_1} \sum_{i=1}^{m} [y_i - (a_1 x_i + a_0)]^2 &= 2 \sum_{i=1}^{m} (y_i - a_1 x_i - a_0)(-x_i) &= 0.
\end{aligned}$$

Luego

$$
\begin{aligned}
(-2)\sum_{i=1}^{m}(y_i-a_1x_i-a_0)=0 \quad &\Rightarrow \quad \sum_{i=1}^{m}y_i-a_1\sum_{i=1}^{m}x_i-m\,a_0=0,\\
(-2)\sum_{i=1}^{m}(y_i-a_1x_i-a_0)x_i=0 \quad &\Rightarrow \quad \sum_{i=1}^{m}x_iy_i-a_1\sum_{i=1}^{m}x_i^2-a_0\sum_{i=1}^{m}x_i=0.
\end{aligned}
$$

Reordenando estas ecuaciones se puede obtener el siguiente sistema lineal de dos ecuaciones con las 2 incógnitas a_0 y a_1:

$$
\begin{cases}
a_0\,m + a_1\sum_{i=1}^{m}x_i = \sum_{i=1}^{m}y_i\\
a_0\sum_{i=1}^{m}x_i + a_1\sum_{i=1}^{m}x_i^2 = \sum_{i=1}^{m}x_iy_i.
\end{cases}
$$

Estas ecuaciones son conocidas como **ecuaciones normales**.

Resolviendo este sistema se obtienen los coeficientes buscados

$$
a_0=\frac{\left(\sum_{i=1}^{m}x_i^2\right)\left(\sum_{i=1}^{m}y_i\right)-\left(\sum_{i=1}^{m}x_iy_i\right)\left(\sum_{i=1}^{m}x_i\right)}{m\left(\sum_{i=1}^{m}x_i^2\right)-\left(\sum_{i=1}^{m}x_i\right)^2},\quad a_1=\frac{m\left(\sum_{i=1}^{m}x_iy_i\right)-\left(\sum_{i=1}^{m}x_i\right)\left(\sum_{i=1}^{m}y_i\right)}{m\left(\sum_{i=1}^{m}x_i^2\right)-\left(\sum_{i=1}^{m}x_i\right)^2}.
$$

Para los datos del problema inicial se obtienen $a_0=-0.360$ y $a_1=1.538$, y el gráfico de la aproximación lineal puede verse en la Figura (3).

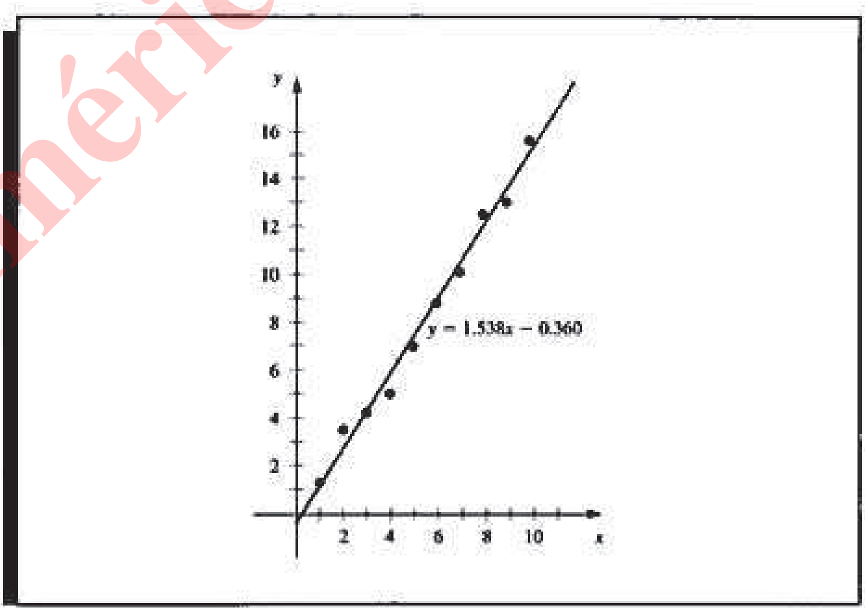

Figura 3: Modelo lineal

Ahora consideremos el caso general donde se tienen los puntos $\{(x_i,y_i)\}$ para $i=1,\ldots,m$ y se propone un modelo polinomial de grado (menor o igual a) n, con $n<m-1$

$$
P_n(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x^1+a_0=\sum_{j=0}^{n}a_jx_i^j.
$$

Al igual que en el caso de dos coeficientes, se deben determinar $a_0, a_1, \ldots, a_n$ que minimizan

$$E(a_0,\ldots,a_n) = \sum_{i=1}^{m}[y_i - P_n(x_i)]^2.$$

Podemos reescribir esta función convenientemente

$$\begin{aligned}
E(a_0,\ldots,a_n) &= \sum_{i=1}^{m}[y_i - P_n(x_i)]^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m} P_n(x_i)y_i + \sum_{i=1}^{m}(P_n(x_i))^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right) y_i + \sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right)^2 \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{i=1}^{m} y_i\left(\sum_{j=0}^{n} a_j x_i^j\right) + \sum_{i=1}^{m}\left(\sum_{j=0}^{n} a_j x_i^j\right)\left(\sum_{k=0}^{n} a_k x_i^k\right) \\
&= \sum_{i=1}^{m} y_i^2 - 2\sum_{j=0}^{n} a_j\left(\sum_{i=1}^{m} y_i x_i^j\right) + \sum_{j=0}^{n}\sum_{k=0}^{n} a_j a_k\left(\sum_{i=1}^{m} x_i^{j+k}\right)
\end{aligned}$$

Para minimizar E se deben calcular las derivadas parciales $\partial E/\partial a_j$ para $j = 0,\ldots,n$ e igualar a cero:

$$\frac{\partial E}{\partial a_j} = -2\sum_{i=1}^{m} y_i x_i^j + 2\sum_{k=0}^{n} a_k\left(\sum_{i=1}^{m} x_i^{j+k}\right) = 0, \qquad \text{para} \qquad j = 0,\ldots,n.$$

Así tenemos las $n+1$ **ecuaciones normales** en las $n+1$ incógnitas $a_0,\ldots,a_n$:

$$\sum_{k=0}^{n} a_k \sum_{i=1}^{m} x_i^{j+k} = \sum_{i=1}^{m} y_i x_i^j, \qquad \text{para} \qquad j = 0,\ldots,n.$$

Es decir, se obtiene el siguiente sistema lineal

$$\left\{\begin{array}{ccccccccccc}
a_0\sum_{i=1}^{m} x_i^0 & + & a_1\sum_{i=1}^{m} x_i^1 & + & a_2\sum_{i=1}^{m} x_i^2 & + & \ldots & + & a_n\sum_{i=1}^{m} x_i^n & = & \sum_{i=1}^{m} y_i x_i^0, \\
a_0\sum_{i=1}^{m} x_i^1 & + & a_1\sum_{i=1}^{m} x_i^2 & + & a_2\sum_{i=1}^{m} x_i^3 & + & \ldots & + & a_n\sum_{i=1}^{m} x_i^{n+1} & = & \sum_{i=1}^{m} y_i x_i^1, \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
a_0\sum_{i=1}^{m} x_i^n & + & a_1\sum_{i=1}^{m} x_i^{n+1} & + & a_2\sum_{i=1}^{m} x_i^{n+2} & + & \ldots & + & a_n\sum_{i=1}^{m} x_i^{2n} & = & \sum_{i=1}^{m} y_i x_i^n.
\end{array}\right.$$

Este sistema lineal puede expresarse matricialmente, y es posible probar que si las x_i son todas distintas entonces las ecuaciones normales admiten una única solución.

Observación: a veces el modelo propuesto no es polinomial. Por ejemplo $i)$ $y = be^{ax}$ o $ii)$ $y = bx^a$, donde a y b son los coeficientes a determinar.

Lamentablemente, si se repitiera el procedimiento anterior cambiando el polinomio por los modelos $i)$ o $ii)$ no se obtiene un sistema lineal, sino un sistema no lineal el cual no se puede resolver por métodos sencillos.

Un método más simple consiste en aplicar logaritmo natural en los modelos i) o ii):

$$\ln y = lnb + ax \qquad \text{o} \qquad \ln y = lnb + a\ln x,$$

de esta manera es posible usar nuevamente un modelo lineal.

Ejemplo: supongamos que se conocen los siguientes datos

x_i	1.00	1.25	1.50	1.75	2.00
y_i	5.10	5.79	6.53	7.45	8.46

y se sabe que corresponden a un modelo de la forma $y = be^{ax}$. Aplicando logaritmo se obtiene $\ln y = \ln b + ax$, un modelo lineal $\tilde{y} = \tilde{b} + ax$. Antes de aplicar las fórmulas de cuadrados mínimos para el modelo lineal conviene reescribir la tabla anterior:

x_i	1.00	1.25	1.50	1.75	2.00
$\tilde{y}_i$	1.629	1.756	1.876	2.008	2.135

Así se obtienen $a = 0.5056$ y $\tilde{b} = 1.122$ y por lo tanto $b = e^{\tilde{b}} = e^{1.122} = 3.071$.