

Capítulo 7

Método de los mínimos cuadrados

Este capítulo está dedicado a la resolución de problemas de aproximación de funciones y ajuste de datos a un polinomio por el método de los mínimos cuadrados.

7.1. Proyección Ortogonal

Si U es un subespacio de un espacio euclídeo E , el conjunto de vectores ortogonales a U se define como

$$U^\perp = \{w \in E \mid u \cdot w = 0 \quad \forall u \in U\}$$

$U^\perp$ es un subespacio vectorial de E denominado **suplemento ortogonal** de U respecto de E .

Así, dado un vector cualquiera $w \in E$ y conocida una base $\mathcal{B} = \{e_1, e_2, \dots, e_m\}$ de U , podemos afirmar que

$$w \in U^\perp \iff w \cdot e_i = 0 \quad \forall i = 1, 2, \dots, m$$

Por otra parte, dado un vector cualquiera $v \in E$, se denomina **proyección ortogonal** de v sobre U , al vector $p \in U$ tal que $v - p \in U^\perp$.

7.1.1. Cálculo de la proyección ortogonal

En las condiciones anteriores, para calcular dicha proyección p , tendremos que encontrar $c_1, c_2, \dots, c_m \in \mathbb{R}$, tales que el vector $p = c_1 e_1 + c_2 e_2 + \dots + c_m e_m$ verifique que:

$$v - p \perp U \iff (v - p) \cdot e_i = 0 \iff v \cdot e_i = p \cdot e_i, \quad \forall i = 1, 2, \dots, m$$

y, teniendo en cuenta que $p = c_1 e_1 + c_2 e_2 + \dots + c_m e_m$, se deduce que $c_1, c_2, \dots, c_m$, serán la solución del sistema de m ecuaciones :

$$v \cdot e_i = c_1 e_1 \cdot e_i + c_2 e_2 \cdot e_i + \dots + c_m e_m \cdot e_i \quad \forall i = 1, 2, \dots, m$$

- De la proyección ortogonal cabe destacar la siguiente propiedad:

$$\|v - p\| < \|v - w\| \quad \forall w \in U, w \neq p$$

de la que se deducirá su utilidad en la resolución de problemas de aproximación.

7.2. Aproximación mínimo cuadrática de funciones mediante polinomios

Dada $f \in \mathcal{C}[a, b]$, nos planteamos buscar el polinomio de grado n , $P_n(x) = \sum_{k=0}^n a_k x^k$, que

se ajuste lo más posible a la función en el sentido de los mínimos cuadrados, es decir, tal que

$$\|f(x) - P_n(x)\|^2 \text{ sea mínima}$$

Siendo la norma considerada, la inducida por el producto escalar

$$f \cdot g = \int_a^b f(x)g(x)dx$$

en el espacio de las funciones continuas en $[a, b]$, la condición anterior equivale a que el valor

$$\int_a^b [f(x) - P_n(x)]^2 dx \text{ sea mínimo}$$

Como ya se ha demostrado en teoría, el polinomio $P_n(x)$ solución del problema planteado, será la proyección ortogonal de f en el espacio vectorial de los polinomios de grado n , con el producto escalar indicado anteriormente.

Al valor $e = \|f - P_n(x)\|^2$ le denominaremos error cometido en la aproximación.

Ejemplo 7.2.1 *Aproxímese, por el método de mínimos cuadrados la función $f(x) = e^x$ en el intervalo $[-1, 1]$ mediante un polinomio de tercer grado, y calcúlese el error cometido en la aproximación. Dibújese el polinomio obtenido junto a la función $f(x) = e^x$ en el mismo gráfico para poder compararlas.*

Solución

Según acabamos de comentar, considerando el producto escalar

$$f \cdot g = \int_{-1}^1 f(x)g(x)dx$$

debemos obtener la proyección ortogonal P de $f(x)$ sobre $\mathbb{R}_3[x]$, lo que equivale a decir que $e^x - P$ es ortogonal a $\mathbb{R}_3[x]$. Para ello, bastará que $e^x - P$ sea ortogonal a una base de $\mathbb{R}_3[x]$, por ejemplo a la canónica

$$\begin{cases} (e^x - P) \cdot 1 = 0 \\ (e^x - P) \cdot x = 0 \\ (e^x - P) \cdot x^2 = 0 \\ (e^x - P) \cdot x^3 = 0 \end{cases} \leftrightarrow \begin{cases} e^x \cdot 1 = P \cdot 1 \\ e^x \cdot x = P \cdot x \\ e^x \cdot x^2 = P \cdot x^2 \\ e^x \cdot x^3 = P \cdot x^3 \end{cases}$$

y como $P \in \mathbb{R}_3[x]$, entonces $P = a_0 1 + a_1 x + a_2 x^2 + a_3 x^3$ y, las relaciones anteriores, son equivalentes a las siguientes

$$\begin{cases} (a_0 1 + a_1 x + a_2 x^2 + a_3 x^3) \cdot 1 = e^x \cdot 1 \\ (a_0 1 + a_1 x + a_2 x^2 + a_3 x^3) \cdot x = e^x \cdot x \\ (a_0 1 + a_1 x + a_2 x^2 + a_3 x^3) \cdot x^2 = e^x \cdot x^2 \\ (a_0 1 + a_1 x + a_2 x^2 + a_3 x^3) \cdot x^3 = e^x \cdot x^3 \end{cases}$$

o bien

$$\begin{pmatrix} 1 \cdot 1 & x \cdot 1 & x^2 \cdot 1 & x^3 \cdot 1 \\ 1 \cdot x & x \cdot x & x^2 \cdot x & x^3 \cdot x \\ 1 \cdot x^2 & x \cdot x^2 & x^2 \cdot x^2 & x^3 \cdot x^2 \\ 1 \cdot x^3 & x \cdot x^3 & x^2 \cdot x^3 & x^3 \cdot x^3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} e^x \cdot 1 \\ e^x \cdot x \\ e^x \cdot x^2 \\ e^x \cdot x^3 \end{pmatrix}$$

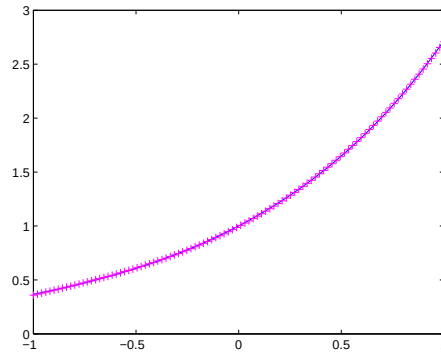
Para resolverlo con MATLAB, creamos el siguiente fichero, que llamaremos **aproximacion.m**

```
clear all
clc
close all
syms x real
n=3;           %Grado del polinomio proyección
a=-1;b=1;      %Extremos del intervalo intervalo de definición de la función
f=exp(x)       %Función a proyectar
v=x.^[0:n]     %Vector cuyas componentes son los vectores de la base canónica de Rn[x]
H=v.'*v        %Matriz de integrandos del primer miembro
G=int(H,x,a,b) %Matriz del producto escalar en Rn[x]
D=v.'*f        %Matriz de integrandos del segundo miembro
B=int(D,x,a,b) %Matriz numérica segundo miembro
E=rref([G B])  %Matriz reducida canónica de la ampliada
s=E(:,end)     %Vector solución
P=s.'*v.'      %Polinomio solución

% Las siguientes instrucciones dibujan la función y el polinomio solución
% en el intervalo [a,b]
xnum=linspace(a,b);
y=subs(f,x,xnum);
Ps=subs(P,x,xnum);
plot(xnum,y,'b')
hold on
plot(xnum,Ps,'m')

%Cálculo del error cometido

e=double(int((f-P)^2,x,a,b))
```



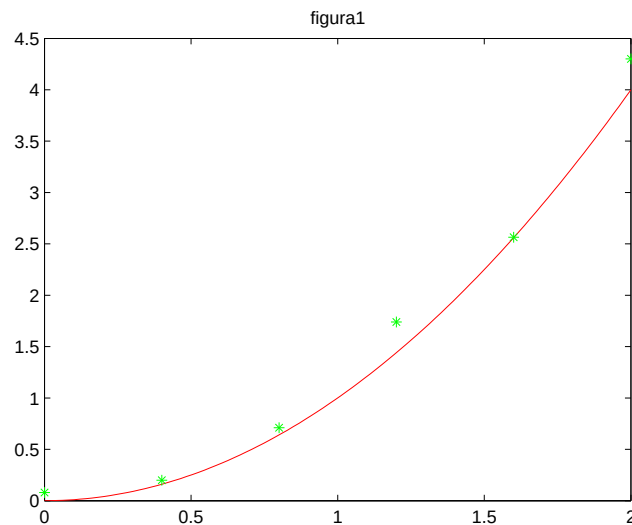
Nótese, que este fichero nos servirá para aproximar cualquier función continua f , en cualquier intervalo $[a,b]$, por un polinomio de cualquier grado n , sin más que introducir en cada caso los valores de f , a , b y n deseados.

7.3. Ajuste de datos mediante polinomios

Nos planteamos, ahora, el siguiente problema: partiendo de datos empíricos de una variable y obtenidos a partir de otra variable x

x	x_1	x_2	x_3	x_4	$\cdots$	x_n
y	y_1	y_2	y_3	y_4	$\cdots$	y_n

encontrar una función polinómica de grado k , normalmente con $k < n - 1$ prefijado, que aproxime de la mejor manera posible estos datos. La figura 1 muestra un ejemplo en el cual se representan puntos de datos experimentales a lo largo de una curva que se ajusta de manera aproximada a estos datos.



Debemos, por tanto, determinar los parámetros $c_1, c_2, c_3, \dots, c_{k+1}$ que hagan que la curva que describe el polinomio $y = c_1 + c_2x + c_3x^2 + \dots + c_{k+1}x^k$ esté lo más cercana posible a los puntos.

Si dichos puntos estuvieran sobre la curva, $c_1, c_2, c_3, \dots, c_{k+1}$ satisfarían las ecuaciones

$$\begin{cases} y(x_1) = c_1 + c_2x_1 + c_3x_1^2 + \dots + c_{k+1}x_1^k = y_1 \\ y(x_2) = c_1 + c_2x_2 + c_3x_2^2 + \dots + c_{k+1}x_2^k = y_2 \\ \vdots \\ y(x_n) = c_1 + c_2x_n + c_3x_n^2 + \dots + c_{k+1}x_n^k = y_n \end{cases}$$

es decir, el sistema

$$\begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k+1} \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

y el problema se reduciría a resolver un sistema de ecuaciones lineales compatible con matriz de

coeficientes $A = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^k \\ 1 & x_2 & x_2^2 & \cdots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^k \end{pmatrix}$, término independiente $Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ e incógnitas

$$c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k+1} \end{pmatrix}.$$

Obviamente, en la mayoría de los casos este sistema $Y = Ac$ será incompatible con lo que tendremos que conformarnos con resolverlo de forma aproximada.

Tal y como se hizo en teoría, plantearemos el problema de buscar unos valores de las incógnitas $c_1, c_2, c_3, \dots, c_{k+1}$, tales que: $\|Y - Ac\|$ sea mínima. Si consideramos los vectores linealmente independientes

$$e_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}, e_2 = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, e_3 = \begin{pmatrix} x_1^2 \\ x_2^2 \\ \vdots \\ x_n^2 \end{pmatrix}, \dots, e_{k+1} = \begin{pmatrix} x_1^k \\ x_2^k \\ \vdots \\ x_n^k \end{pmatrix}$$

el problema anterior equivale a encontrar la proyección ortogonal

$$p = c_1e_1 + c_2e_2 + c_3e_3 + \dots + c_{k+1}e_{k+1}$$

del vector Y sobre el subespacio $\langle e_1, e_2, \dots, e_{k+1} \rangle$, lo que nos llevaría a resolver el sistema de ecuaciones:

$$\begin{cases} Y \cdot e_1 = c_1e_1 \cdot e_1 + c_2e_2 \cdot e_1 + c_3e_3 \cdot e_1 + \dots + c_{k+1}e_{k+1} \cdot e_1 \\ Y \cdot e_2 = c_1e_1 \cdot e_2 + c_2e_2 \cdot e_2 + c_3e_3 \cdot e_2 + \dots + c_{k+1}e_{k+1} \cdot e_2 \\ \vdots \\ Y \cdot e_{k+1} = c_1e_1 \cdot e_{k+1} + c_2e_2 \cdot e_{k+1} + c_3e_3 \cdot e_{k+1} + \dots + c_{k+1}e_{k+1} \cdot e_{k+1} \end{cases}$$

de incógnitas $c_1, c_2, c_2, \dots, c_{k+1}$, que matricialmente se escribiría como:

$$\begin{pmatrix} Y \cdot e_1 \\ Y \cdot e_2 \\ \vdots \\ Y \cdot e_{k+1} \end{pmatrix} = \begin{pmatrix} e_1 \cdot e_1 & e_2 \cdot e_1 & \cdots & e_{k+1} \cdot e_1 \\ e_1 \cdot e_2 & e_2 \cdot e_2 & \cdots & e_{k+1} \cdot e_2 \\ \vdots & \vdots & \ddots & \vdots \\ e_1 \cdot e_{k+1} & e_2 \cdot e_{k+1} & \cdots & e_{k+1} \cdot e_{k+1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_{k+1} \end{pmatrix}$$

Este sistema recibe el nombre de sistema de ecuaciones normales y se trata de un sistema compatible y determinado.

Considerando la norma euclídea, el valor $e = \|Y - p\|^2$, se denomina error de mínimos cuadrados de la aproximación. El error de mínimos cuadrados cometido depende, obviamente, del número de puntos considerados, así que otro valor interesante de calcular es el llamado error cuadrático medio, que viene dado por $emed = \frac{\sqrt{e}}{n}$, donde n es el número de medidas efectuadas.

Nota: Si denotamos, finalmente, por $P(x) = c_1 + c_2x + \dots + c_{k+1}x^k$, al polinomio aproximador

obtenido, tendremos que $p = \begin{pmatrix} P(x_1) \\ P(x_2) \\ \vdots \\ P(x_n) \end{pmatrix}$ con lo que $e = \sum_{i=1}^n (y_i - P(x_i))^2$.

Ejemplo 7.3.1 Encuéntrese la parábola que ofrece la mejor aproximación por mínimos cuadrados a los puntos $(-1,1), (0,-1), (1,0)$ y $(2,2)$.

La tabla de datos que nos proporciona el enunciado sería, en este caso, la siguiente:

x	-1	0	1	2
y	1	-1	0	2

De acuerdo con la notación anterior tendremos que resolver el sistema compatible y determinado:

$$\begin{pmatrix} Y \cdot e_1 \\ Y \cdot e_2 \\ Y \cdot e_3 \end{pmatrix} = \begin{pmatrix} e_1 \cdot e_1 & e_2 \cdot e_1 & e_3 \cdot e_1 \\ e_1 \cdot e_2 & e_2 \cdot e_2 & e_3 \cdot e_2 \\ e_1 \cdot e_3 & e_2 \cdot e_3 & e_3 \cdot e_3 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$$

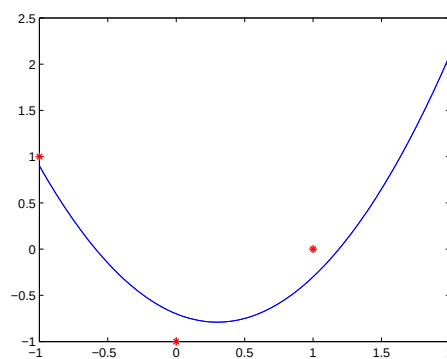
donde Y es el vector con las mediciones de la variable y , e_2 con las de la variable x , $e_1 = \text{ones}(3,1)$ y $e_3 = e_2^2$.

Lo resolvemos en el fichero **ajuste1.m** siguiente:

```
% Ajuste por un polinomio de grado 2 ; y=c1+c2x+c3x^2
% Mediciones de la variable y en el vector Y
Y=[1;-1;0;2];
% Valores de la variable x en el vector e2
e2=[-1;0;1;2];n=size(e2,1);
e1=ones(n,1);e3=e2.^2;
B=[e1 e2 e3].'*Y; % Vector de términos independientes
G=[e1 e2 e3].'*[e1 e2 e3]; % Matriz de Gram de la base
E=rref([G B]);
s=E(:,end); % Vector de soluciones c1,c2,c3.
syms x
P=s.*[1 x x^2].'
```

```
% Para calcular el error
Pe=subs(P,x,e2);% Vector [P(x1),P(x2),P(x3),P(x4)] .'
e=norm(Y-Pe)^2
emed=sqrt(e)/n
% Para dibujar
plot(e2,Y,'r*')
hold on
xnum=linspace(-1,2);
Pnum=subs(P,x,xnum);
plot(xnum,Pnum);
```

La gráfica resultante sería:



Nota: Hemos visto que el problema de mínimos cuadrados planteado equivale a buscar los valores de c_1 , c_2 y c_3 tales que $\|Y - Ac\|$ sea mínima, siendo A la matriz cuyas columnas son los vectores e_1 , e_2 y e_3 y $c = (c_1, c_2, c_3)^t$. Dichos valores, según hemos visto, pueden obtenerse con MatLab mediante la orden $A \backslash Y$. Tendríamos, por tanto, otro método para resolver estos problemas, si bien nosotros no lo utilizaremos en este primer curso.

Ejemplo 7.3.2 La tabla dada proporciona las expectativas de vida de personas nacidas en España en los años que allí se indican. Determínese la recta de aproximación por mínimos cuadrados de estos datos y utilícese para predecir la expectativa de vida de alguien nacido en el año 2000.

Año	1920	1930	1940	1950	1960	1970	1980	1990
Expectativa de vida (en años)	54,1	59,7	62,9	68,2	69,7	70,8	73,7	75,4

Solución

Lo resolvemos en el fichero **ajuste2.m** siguiente:

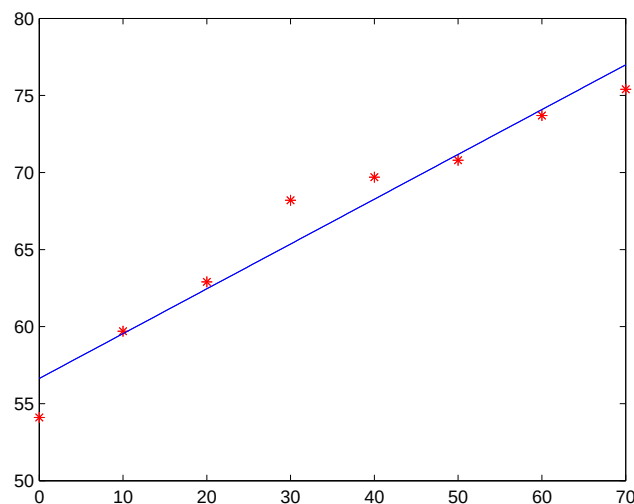
```
% Ajuste por un polinomio de grado 1; y=a+ b t;
% Mediciones de la variable y en el vector Y
Y=[54.1;59.7;62.9;68.2;69.7;70.8;73.7;75.4] ;
% Valores de la variable t en el vector e2
e2=[1920:10:1990] .';n=size(e2,1);
```

```

e1=ones(n,1);
B=[e1 e2].'*Y; % Vector de términos independientes
G=[e1 e2].'*[e1 e2]; % Matriz de Gram de la base
E=rref([G B]);
s=E(:,end); % Vector de soluciones a,b
syms t
P=s.'*[1 t].'
Expectativa=subs(P,t,2000)
% Para dibujar
plot(e2,Y,'r*')
hold on
tnum=linspace(1920,1990);
Pnum=subs(P,t,tnum);
plot(tnum,Pnum);

```

La gráfica resultante sería:



Ejercicio 7.1 Aproxime por el método de mínimos cuadrados la función $f(x) = \sin(\pi x)$ en $[0, 1]$ mediante un polinomio de segundo grado, y calcule el error cometido en la aproximación.

Ejercicio 7.2 En la fabricación de un cierto producto X , la cantidad de un cierto compuesto β presente es controlada por la cantidad del ingrediente α utilizado en el proceso. Al fabricar un kilogramo de X se registran la cantidad α utilizada y la cantidad β presente. Se obtienen los siguientes datos

Cantidad α utilizada	3	4	5	6	7	8	9	10	11	12
Cantidad β presente	4.5	5.5	5.7	6.6	7	7.7	8.5	8.7	9.5	9.7

Suponiendo que la relación entre la cantidad α y la cantidad β esta dada por una ecuación lineal $\beta = a\alpha + b$, determinar dicha recta utilizando el método de los mínimos cuadrados. Utilizar la ecuación obtenida para predecir la cantidad β presente en un kilogramo de X si se utilizan 30 unidades de α por kilogramo de X .

Ejercicio 7.3 *La tabla siguiente muestra la cantidad de un cierto contaminante, y , con respecto a la cantidad normal encontrada en un punto geográfico, para una cierta cantidad de aire t y en intervalos de media hora:*

t	1	1.5	2	2.5	3	3.5	4	4.5	5
y	-0.15	0.24	0.68	1.04	1.21	1.15	0.86	0.41	-0.08

La gráfica de las medidas sugiere una relación de tipo polinómico. Encontrar el polinomio de grado 2 ($y = at^2 + bt + c$) que produzca un buen modelo a partir de estos datos.