

PRÁCTICA 4: APROXIMACIÓN POR MÍNIMOS CUADRADOS

1. Aproximación por mínimos cuadrados

1.1. Aproximación por una recta

Supongamos un conjunto de n pares de datos: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Se trata de encontrar una recta, que llamaremos recta de regresión, que mejor se aproxime a los datos en el sentido de los mínimos cuadrados. Esta recta puede obtenerse con Matlab con la orden `polyfit` (véase la Práctica 3).

1.1.1. Ejemplo

Se tienen los siguientes datos relativos al crecimiento de una determinada población:

Tiempo (meses)	Población (miles)
1	5,1
2	7
3	8,9
4	11,1
5	13,2
6	14,8
7	17
8	18,9
9	21,1
10	23,1

Calcular la recta de regresión correspondiente a estos datos y representar en un mismo gráfico la recta y los datos. Calcular el error cuadrático.

[Ejemplo111.m](#)

```
1 % Conjunto de datos
2 x=[1:10];
3 y=[5.1,7,8.9,11.1,13.2,14.8,17,18.9,21.1,23.1];
4 % Grafico de los datos
5 plot(x,y,'r*')
6 hold on
7 % Calculo de la recta de regresion
8 p=polyfit(x,y,1)
9 % Calculo del error
10 errorc=norm(y-polyval(p,x))
11 % Grafica de la recta de regresion
12 xp=[1:0.01:10];
13 yp=polyval(p,xp);
14 plot(xp,yp)
15 xlabel('Tiempo(meses)'),ylabel('Población(miles)')
16 hold off
```

La recta de regresión $y = a + bx$ para el conjunto de datos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ puede obtenerse también resolviendo el sistema de ecuaciones normales:

$$\begin{cases} an + b \sum_{i=1}^n x_i = \sum_{i=1}^n y_i \\ a \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i y_i \end{cases}$$

1.1.2. Ejemplo

Hacer el ejercicio propuesto en el Ejemplo 1.1.1 resolviendo el sistema de ecuaciones normales. En primer lugar creamos un fichero que resuelva el sistema de ecuaciones normales:

```

1 function [a,b]=rectareg(x,y)
2 M=[length(x),sum(x);sum(x),sum(x.^2)];
3 u=[sum(y);sum(x.*y)];
4 sol=M\u;
5 a=sol(1);
6 b=sol(2);

```

Resolvemos el ejercicio:

```

1 % Conjunto de datos
2 x=[1:10];
3 y=[5.1, 7, 8.9, 11.1, 13.2, 14.8, 17, 18.9, 21.1, 23.1];
4 % Grafico de los datos
5 plot(x,y,'r*')
6 hold on
7 % Calculo de la recta de regresion
8 [a,b]=rectareg(x,y)
9 % Calculo del error
10 errorc=norm(y-(a+b*x))
11 % Grafica de la recta de regresion
12 xp=[1:0.01:10];
13 yp=a+b*xp;
14 plot(xp,yp)
15 xlabel('Tiempo(meses)'),ylabel('Población(miles)')
16 hold off

```

1.1.3. Ejercicio

Generar un conjunto de 20 puntos aproximadamente alineados (consúltase en la Ayuda la orden `rand`). Calcular la recta de regresión correspondiente a estos datos y representar en un mismo gráfico la recta y los datos.

```

1 % Conjunto de datos aproximadamente alineados
2 x=[1:20];
3 y=2+3*x+rand(1,20);

```

1.2. Aproximación por una exponencial

Se trata de encontrar la función exponencial $y = ae^{bx}$ que mejor se aproxime a los pares de datos $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$. Para ello se realiza la transformación

$$Y = \ln y,$$

de donde:

$$Y = \ln y = \ln a + bx.$$

Si denotamos $A = \ln a$, se tiene

$$Y = A + bx.$$

De nuevo obtenemos el sistema de ecuaciones normales:

$$\begin{cases} An + b \sum_{i=1}^n x_i = \sum_{i=1}^n Y_i \\ A \sum_{i=1}^n x_i + b \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i Y_i \end{cases}$$

La resolución de este sistema nos proporciona $a = e^A$ y b . Creamos un fichero que resuelva este sistema de ecuaciones:

```
aproxexp.m
1 function [a,b]=aproxexp(x,y)
2 Y=log(y);
3 M=[length(x), sum(x); sum(x), sum(x.^2)];
4 u=[sum(Y); sum(x.*Y)];
5 sol=M\u;
6 a=exp(sol(1));
7 b=sol(2);
```

1.2.1. Ejemplo

Se tienen los siguientes datos relativos a la desintegración de una sustancia radiactiva:

Tiempo (horas)	Cantidad (mg)
0	102,9
1	75,8
2	56,1
3	42,2
4	31,1
5	23,6

Determinar la función exponencial que mejor se ajusta a los datos anteriores y representar en un mismo gráfico la función exponencial y los datos.

```
Ejemplo121.m
1 % Conjunto de datos
2 x=[0:5];
3 y=[102.9, 75.8, 56.1, 42.2, 31.1, 23.6];
4 % Grafico de los datos
5 plot(x,y,'r*')
6 hold on
7 % Calculo de la exponencial
8 [a,b]=aproxexp(x,y)
9 % Calculo del error
10 errorc=norm(y-a*exp(b*x))
11 % Grafica de la exponencial
12 xp=[0:0.01:5];
13 yp=a*exp(b*xp);
14 plot(xp,yp)
15 xlabel('Tiempo(horas)'),ylabel('Cantidad(mg)')
16 hold off
```

1.2.2. Ejercicio

Resolver el ejemplo anterior utilizando la orden `polyfit`.

1.2.3. Ejercicio

Se tienen los siguientes datos:

x	1	2	3	4	5	6	7	8	9	10
y	3,5	3,6	4,1	4,5	5,1	5,5	6,1	6,6	7,4	8,2

Determinar la función exponencial que mejor se ajusta a los datos anteriores y representar en un mismo gráfico la exponencial y los datos.