Least Squares method

Linear least squares is a method for estimating unknown parameters in a linear regression model, with the goal of minimizing the sum of the squares of the difference between observed responses in the given data set and those predicted by a linear function of a set of explanatory variables.

Let $\{(x_i, y_i)\}_{i=1}^{n}$ be a set of given data.

For linear least square, $y = a_0 + a_1 x$

We try to find a_0 and a_1 such that the difference between the linear model and the given data is minimum.

The error between the model and the given data is

$$d_i = y_i - (a_0 + a_1 x_i), \quad i = 1, 2, \cdots, n.$$

Let $D = \sum_{i=1}^{n} d_i^2 = \sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)^2$

To minimize D, from Calculus, we have

$$\frac{\partial D}{\partial a_0} = -\sum_{i=1}^{n} 2(y_i - a_0 - a_1 x_i) = -2\sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)$$

$$\frac{\partial D}{\partial a_1} = -2\sum_{i=1}^{n} (y_i - a_0 - a_1 x_i)(x_i) = -2\left(\sum_{i=1}^{n} x_i y_i - a_0 \sum_{i=1}^{n} x_i - a_1 \sum x_i^2\right)$$

Let $\frac{\partial D}{\partial a_0} = 0$, $\frac{\partial D}{\partial a_1} = 0$, we have

$$\begin{cases} \sum_{i=1}^{n} (a_0 + a_1 x_i - y_i) = 0 \\ \sum_{i=1}^{n} x_i y_i - a_0 \sum_{i=1}^{n} x_i - a_1 \sum_{i=1}^{n} x_i^2 = 0 \end{cases}$$

This implies

$$\begin{cases} n a_0 + \left(\sum_{i=1}^{n} x_i\right) a_1 = \sum_{i=1}^{n} y_i \\ \left(\sum_{i=1}^{n} x_i\right) a_0 + \left(\sum_{i=1}^{n} x_i^2\right) a_1 = \sum_{i=1}^{n} x_i y_i \end{cases}$$

$$3x + 5y = 1$$
$$2x + 3y = 2$$

Cramer's rule :

$$x = \frac{\begin{vmatrix} 1 & 5 \\ 2 & 3 \end{vmatrix}}{\begin{vmatrix} 3 & 5 \\ 2 & 3 \end{vmatrix}} = \frac{3-10}{9-10} = 7$$

$$y = \frac{\begin{vmatrix} 3 & 1 \\ 2 & 2 \end{vmatrix}}{\begin{vmatrix} 3 & 5 \\ 2 & 3 \end{vmatrix}} = \frac{6-2}{-1} = -4$$

By Cramer's rule,

$$a_1 = \frac{\begin{vmatrix} n & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix}}{\begin{vmatrix} n & \sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{vmatrix}} \qquad (1)$$

$$= \frac{n \sum x_i y_i - \sum x_i \cdot \sum y_i}{n \sum x_i^2 - \sum x_i \sum x_i}$$

$$a_0 = \frac{\sum y_i}{n} - \frac{a_1 \left(\sum x_i\right)}{n} \qquad (2)$$

Example 1.

x	-3	-2	1	0	1	2	3
y	0	2	1	-3	-1	-5	-9

For linear least squares, $y = a_0 + a_1 x$,

$$\begin{cases} a_0 + a_1(-3) = 0 \\ a_0 + a_1(-2) = 2 \\ a_0 + a_1(1) = 1 \\ a_0 + a_1(0) = -3 \\ a_0 + a_1(1) = -1 \\ a_0 + a_1(2) = -5 \\ a_0 + a_1(3) = -9 \end{cases}$$

Find a_0 and a_1

In matrix form, we have

$$\underbrace{\begin{bmatrix} 1 & -3 \\ 1 & -2 \\ 1 & 1 \\ 1 & 0 \\ 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}}_{A_{7\times2}} \underbrace{\begin{bmatrix} a_0 \\ a_1 \end{bmatrix}}_{\hat{x}_{2\times1}} = \underbrace{\begin{bmatrix} 0 \\ 2 \\ 1 \\ -3 \\ -1 \\ -5 \\ -9 \end{bmatrix}}_{b_{7\times1}}$$

$$A\hat{x} = b$$

$$\underbrace{(A^T A)}_{2\times2}\hat{x} = \underbrace{A^T b}_{2\times1}$$

$A^T A = 2\times2$ matrix

$$\hat{x} = (A^T A)^{-1}(A^T b)$$

In MATLAB,

```
x = [-3 -2 -1 0 1 2 3]';
y = [0 2 1 -3 -1 -5 -9]';

n = length(x);
A = [ones(n,1) x];
b = y;
A = A'*A;
B = A'*b;
C = A\B;

xt = -3:0.02:3
yt = C(1) + C(2)*xt;
plot(x,y,'*', xt,yt);
```

Alternatively, using (1) and (2), we can also obtain a_0 and a_1.

```
a1 = (n * sum(x.* y) - sum(x)*sum(y)) /(n* sum(x.^2) -
                                      (sum(x))^2 )

a0 = sum(y)/n - a1 * sum(x)/n
```

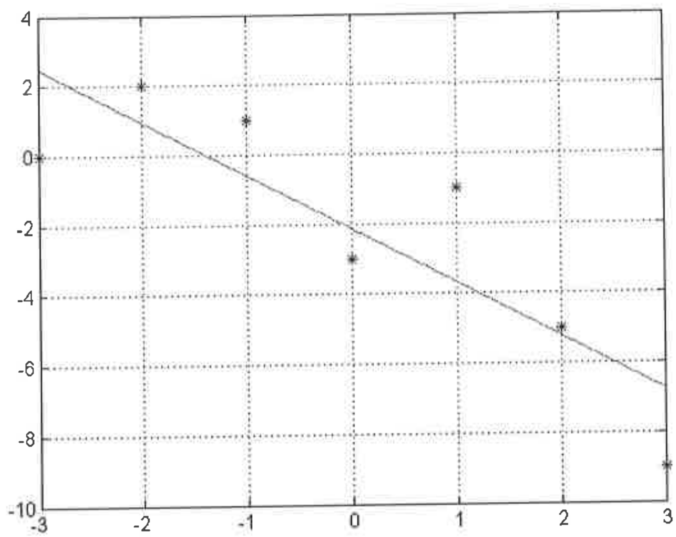
—— Example 2

Let $y = a_0 + a_1 x + a_2 x^2$. (quadratic model)

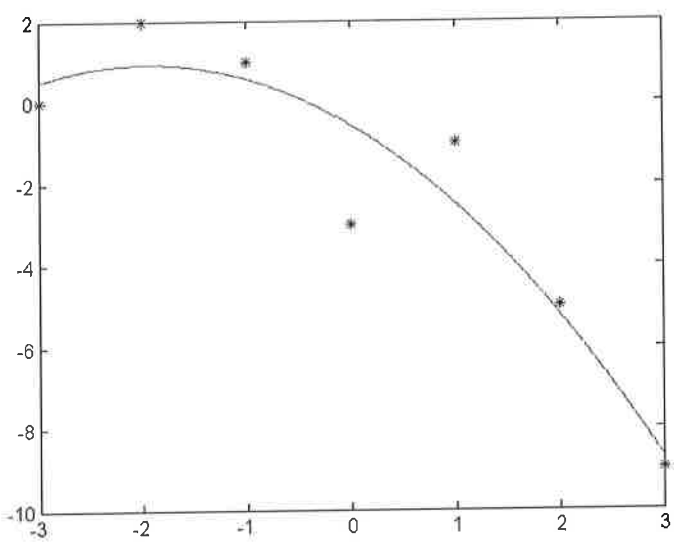
Using the same data $\{(x_i, y_i)\}_{i=1}^{7}$ shown above,

```
n = length(x);
A = [ones(n,1)  x   x.^2];
b = y ;
A = A'* A ;
B = A'* B ;
c = A\B ;          % c = [a0  a1] or c(1) = a0, c(2) = a1 .

xt = -3 : 0.2 : 3 ;
yt = c(1) + c(2)* xt + c(3)* xt.^2 ;
plot(x, y, '*', xt, yt ) ;
```

Linear least squares



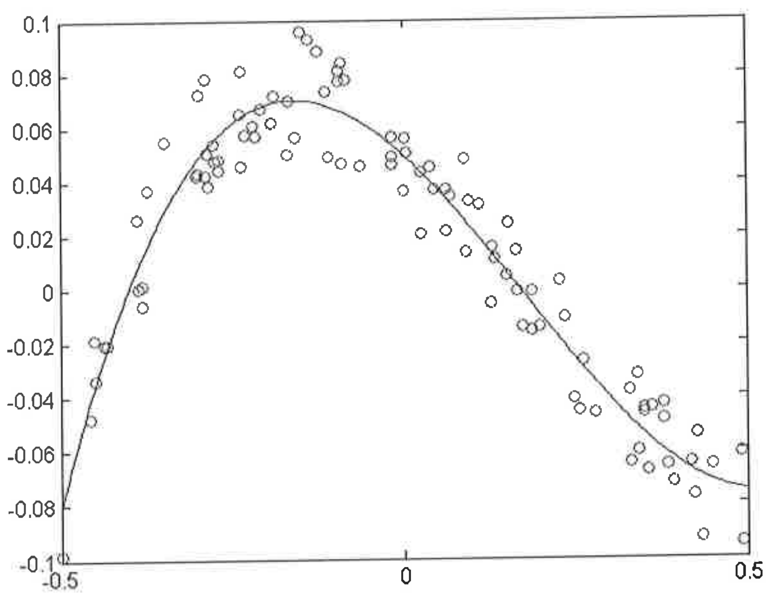
Quadratic least squares



Example 3

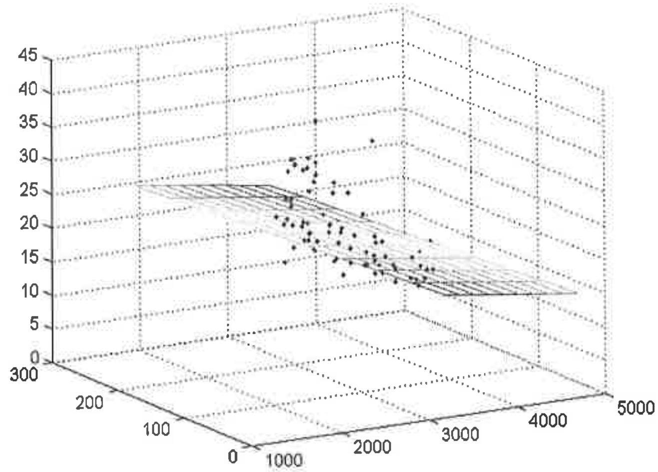
$$y = a_0 + a_1 x + a_2 x^2 + a_3 x^3$$

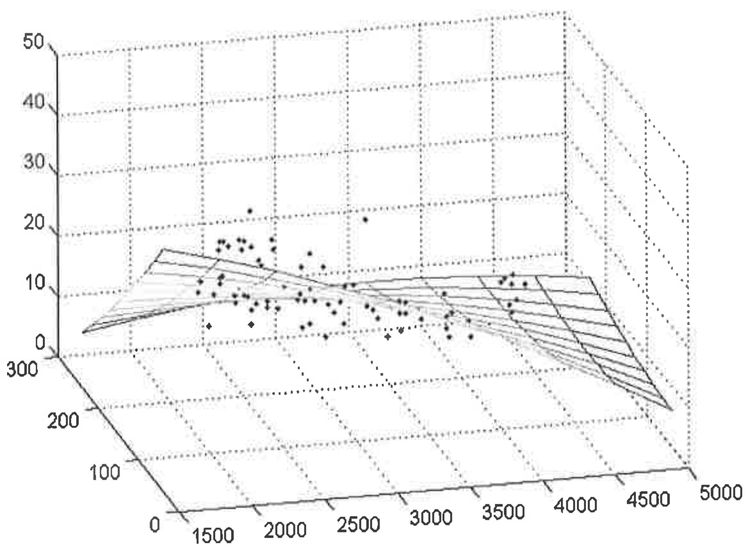
```
load data1

A = [t.^3 t.^2 t.^1 t.^0];

x=A\y;   %% is theoreticaly equivalent to x=(A'*A)\(A'*y)

xt=linspace(-.5,.5,200);

yt=x(1)*xt.^3+x(2)*xt.^2+x(3)*xt+x(4);

plot(t,y,'or',xt,yt,'b');
```



Example 4.

```
load Carsmall
x = Weight ;  y = Horsepower;  z = MPG ;
plot3 ( x,y,z, 'r.');
hold on ;
c = ones ( length (x) , 1);
b = regress ( z, [x  y  c]);  % linear regression  z = b(1) * x + b(2) * y + b(3)
[X, Y] = meshgrid(linspace(1500,5000, 10) , linspace (40,240, 10))
C = ones (10) ;
mesh ( X, Y, b(1) * X + b(2)* Y + b(3) * C]);
b = regress ( z, [x.^2, y.^2 , x.*y, x, y, c]);
figure
plot (x,y,z, 'r.') ;
hold on
mesh (X,Y, b(1) * X.^2 + b(2) * Y.^2 + b(3) * X.*Y + b(4) * X +
                b(5) * Y + b(6) * C) ;
```


45
40
35
30
25
20
15
10
5
0
300
200
100
0
1000
2000
3000
4000
5000


50
40
30
20
10
0
300
200
100
0
1500
2000
2500
3000
3500
4000
4500
5000

Example

```
load data2.mat
n=length(t);
A=[ones(n,1),zeros(n,1),cos(2*pi*t/T);
zeros(n,1),ones(n,1),sin(2*pi*t/T)];
z=[x;y];
X=A\z;
fprintf('M(t)=%f+%fcos(\theta)\n',X(1),X(3))
fprintf('    %f+%fsin(\theta)\n',X(2),X(3))
theta=linspace(0,2*pi,200);
xt=X(1)+X(3)*cos(theta);
yt=X(2)+X(3)*sin(theta);
plot(x,y,'+r',xt,yt);
axis equal
```

Consider $\{(x_i, y_i)\}_{i=1}^{n}$

$$x(t) = a + r\cos\left(2\pi\frac{t}{T}\right) \qquad t \geq 0$$
$$y(t) = b + r\sin\left(2\pi\frac{t}{T}\right)$$

Minimize the error

$$E(a,b,r) = \sum_{i=1}^{n} \left\| \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} - \begin{pmatrix} x_i \\ y_i \end{pmatrix} \right\|^2$$

Find a, b, r.