实验12-1 牙膏的销售量（验证,参见教材P325-331）

下面给出一组数据，其中：

第 1 列　销售周期；

第 2 列　某公司牙膏销售价格（元）x_4；

第 3 列　其它厂家平均价格（元）x_3；

第 4 列　广告费用（百万元）x_2；

第 5 列　价格差（元）x_1(x_3-x_4)；

第 6 列　销售量（百万支）y。

存放在一个名为mydata.m的（数据）m文件中。

1	3.85	3.80	5.50	-0.05	7.38
2	3.75	4.00	6.75	0.25	8.51
3	3.70	4.30	7.25	0.60	9.52
4	3.70	3.70	5.50	0	7.50
5	3.60	3.85	7.00	0.25	9.33
6	3.60	3.80	6.50	0.20	8.28
7	3.60	3.75	6.75	0.15	8.75
8	3.80	3.85	5.25	0.05	7.87
9	3.80	3.65	5.25	-0.15	7.10
10	3.85	4.00	6.00	0.15	8.00
11	3.90	4.10	6.50	0.20	7.89
12	3.90	4.00	6.25	0.10	8.15
13	3.70	4.10	7.00	0.40	9.10
14	3.75	4.20	6.90	0.45	8.86
15	3.75	4.10	6.80	0.35	8.90
16	3.80	4.10	6.80	0.30	8.87
17	3.70	4.20	7.10	0.50	9.26
18	3.80	4.30	7.00	0.50	9.00
19	3.70	4.10	6.80	0.40	8.75
20	3.80	3.75	6.50	-0.05	7.95
21	3.80	3.75	6.25	-0.05	7.65
22	3.75	3.65	6.00	-0.10	7.27
23	3.70	3.90	6.50	0.20	8.00
24	3.55	3.65	7.00	0.10	8.50
25	3.60	4.10	6.80	0.50	8.75
26	3.65	4.25	6.80	0.60	9.21
27	3.70	3.65	6.50	-0.05	8.27
28	3.75	3.75	5.75	0	7.67
29	3.80	3.85	5.80	0.05	7.93
30	3.70	4.25	6.80	0.55	9.26

实验要求：

1．建立mydata.m;

2．绘制y对x_1的散点图。(验证)

程序如下（运行结果与教材327图1比较）:

```
M=dlmread('mydata.m');%读取ASCII码文件
```

```
x1=M(:,5);  y=M(:,6);
plot(x1,y,'bo');
```

2. 确定 y 对 x_1 的的拟合，绘制散点图与拟合曲线组合图形。(验证)

从 y 对 x_1 的散点图可以发现，可用线性模型（直线）

$$y=\beta_0+\beta_1x_1+\varepsilon$$

来拟合（其中ε 是随机误差)。

程序如下（运行结果与教材 p327 图 1 比较):

```
clc;format short g;
M=dlmread('mydata.m');%读取 ASCII 码文件
x1=M(:,5); y=M(:,6);
plot(x1,y,'bo');
b=regress(y,[ones(size(x1)),x1]); % b=[β0 β1]',列向量
xx=sort(x1); %按升序排序
yy=[ones(size(xx)),xx]*b;
hold on;
plot(xx,yy,'-r');
hold off;
```

3. 绘制 y 对 x_2 的散点图。(验证)

程序如下（运行结果与教材 p327 图 2 比较）:

```
clc;format short g;
M=dlmread('mydata.m');%读取 ASCII 码文件
x2=M(:,4); y=M(:,6);
plot(x2,y,'bo');
```

4. 确定 y 对 x_2 的的拟合，绘制散点图与拟合曲线组合图形。(验证)

从 y 对 x_2 的散点图可以发现，可用二次函数模型

$$y=\beta_0+\beta_1x_2+\beta_2x_2^2+\varepsilon$$

来拟合。程序如下（运行结果与教材 p327 图 2 比较):

```
clc;format short g;
M=dlmread('mydata.m');%读取 ASCII 码文件
x2=M(:,4); y=M(:,6);
plot(x2,y,'bo');
b=regress(y,[ones(size(x2)),x2,x2.^2]); % b=[β0 β1 β2]',列向量
xx=sort(x2);
yy=[ones(size(xx)),xx,xx.^2]*b;
hold on;
plot(xx,yy,'-r');
hold off;
```

5. y 对 x_1,x_2 的回归模型及其求解，销售量预测。(验证)

综上得回归模型

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_2^2+\varepsilon$$

变量 x_1,x_2 为回归变量，参数 $\beta_0,\beta_1,\beta_2,\beta_3$ 为回归系数。

程序如下（运行结果与教材 p328 表 2 比较和 p329 的相应结果比较）：

```
clc;format compact;format short g;
M=dlmread('mydata.m');%读取ASCII码文件
x1=M(:,5);x2=M(:,4); y=M(:,6);
[b,bint,r,rint,stats]=regress(y,[ones(size(x1)),x1,x2,x2.^2],0.05);
disp('β 0,β 1,β 2,β 3估计值    置信区间');
[b,bint]
R2=stats(1)
F=stats(2)
p=stats(3)
disp('销售量预测');
x1=0.2
x2=6.5
y=[1 x1,x2,x2^2]*b
```

[提示]

多元线性回归函数调用格式：

[b,bint,r,rint,stats]=regress(y,x,alpha)

用本例说明，多元回归模型为：

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_2^2+\varepsilon$$

输入：

y 为 n（=30）维列向量数据。

x 为对应于回归系数 $(\beta_0,\beta_1,\beta_2,\beta_3)'$ 的数据矩阵 $[1\ \ x_1\ \ x_2\ \ x_2^2]$ (30×4 矩阵，第 1 列全 1)。

alpha 为置信水平（缺省时为 0.05）。

输出：

b 为 $\beta=(\beta_0,\beta_1,\beta_2,\beta_3)'$ 估计值，4 维列向量。

bint 为 b 的置信区间，4×2 矩阵。

r 为残差 n(=30)维列向量 y-xβ。

rint 为 r 的置信区间，30×2 矩阵。

stats 为回归模型的检验统计量，含 4 个值，第 1 个是回归方程的决定系数 R_2（R 是相关系数），第 2 个是 F 统计值，第 3 个是与 F 统计量对应的概率值 p。

实验报告提交：

1. 实验要求 1 的运行结果。
2. 实验要求 2 的运行结果。
3. 实验要求 3 的运行结果。
4. 实验要求 4 的运行结果。
5. 实验要求 5 的运行结果。

实验12-2 牙膏的销售量——模型改进（验证）

仍使用实验 12-1 的数据。

实验要求：

1．y 对 x_1,x_2 的回归模型的改进和求解，销售量预测。(编程)

改进的模型

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_2^2+\beta_4x_1x_2+\varepsilon$$

参考实验 12-1 实验要求 5 的程序，编写一个类似的程序，运行结果与教材 p329 的表 3 及相关结果相比较。

2．完全二次多项式模型（验证）

$$y=\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_1x_2+\beta_4x_1^2+\beta_5x_2^2+\varepsilon$$

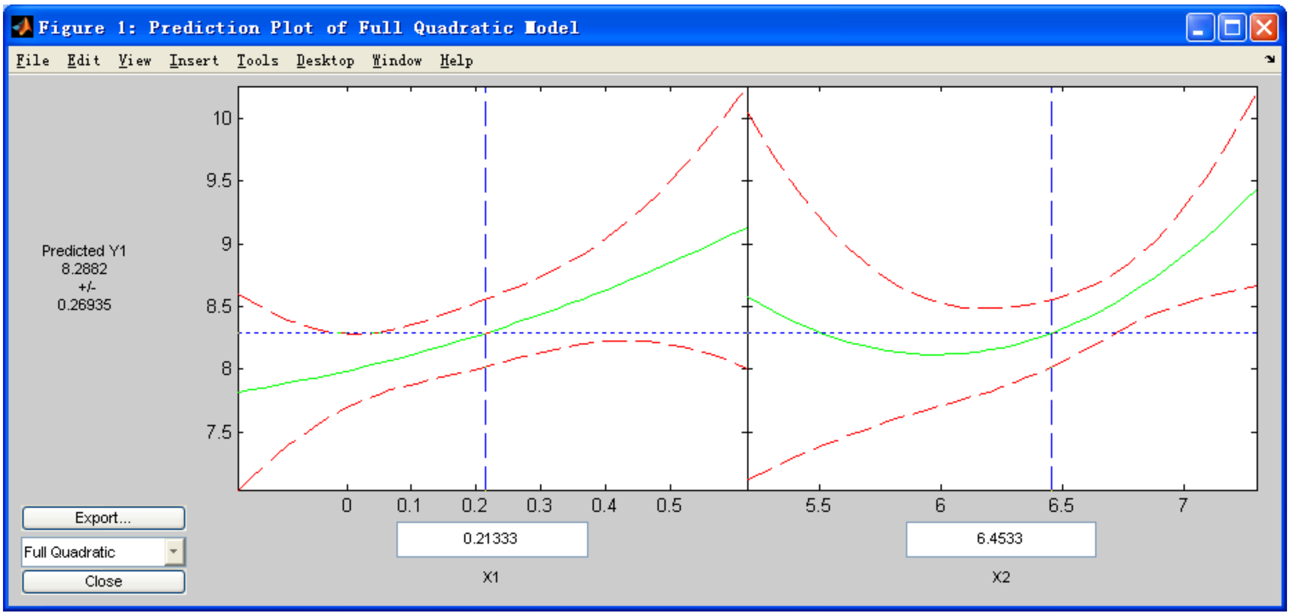
运行以下程序（参考教材 p332)：

```
clear;clc;format compact;format short g;
M=dlmread('mydata.m');%读取 ASCII 码文件
x1=M(:,5);x2=M(:,4); y=M(:,6);
rstool([x1,x2],y,'quadratic')
```

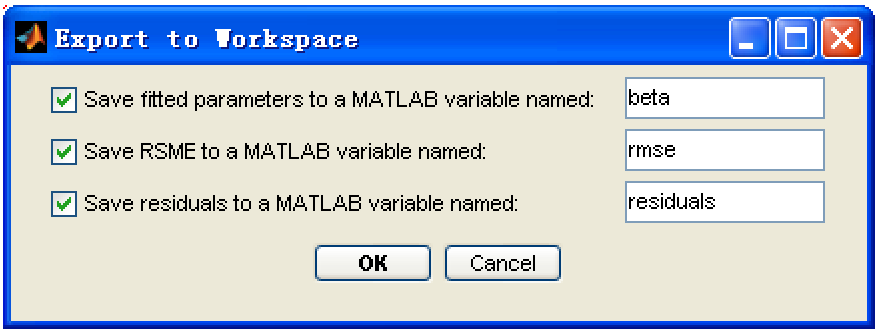
得以下的交互画面。画面中的两个座标系给出 y 的估计值和预测区间。

用鼠标移动交互式画面中的十字线，或在图下方的窗口内输入，可改变 x_1 和 x_2 的数值。

改变 x_1=0.2，x_2=6.5，观察窗口左边的 y 估计值和预测区间。



点击所得交互画面左下方的输出按钮“Export”，所得画面（导出到工作空间）第 1 个复选框是“将拟合参数存到一个名为 beta 的 MATLAB 变量中”，点击 OK。



在命令窗口提示符键入变量名 beta 将得到参数 $(\beta_0,\beta_1,\beta_2,\beta_3,\beta_4,\beta_5)'$的值。

实验报告提交：

1. 实验要求 1 的程序和运行结果。
2. 实验要求 2 的运行结果：参数 $(\beta_0,\beta_1,\beta_2,\beta_3,\beta_4,\beta_5)'$的值。