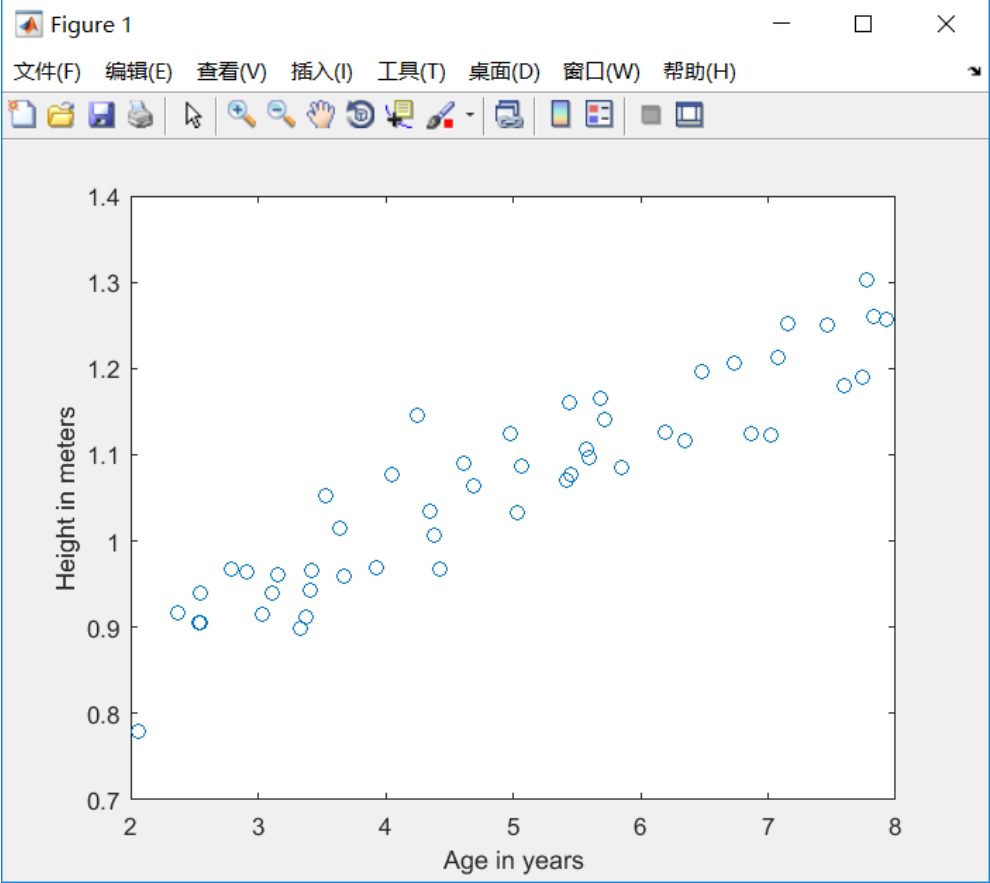


计算机科学与技术 学院

机器学习 课程实验报告

学号：	姓名：	班级：
实验题目：Linear Regression		
实验学时：2.0	实验日期：2018/9/28	
实验目的： 学习线性回归的基本算法应用，熟悉 matla 等软件的使用，锻炼解决实际问题的能力。		
软件环境： Win10 + matlabR2016a		
实验步骤与内容： <h3>1.1 可视化数据集：</h3> <pre>>> load('matlab.mat') >> x=load('ex1x.dat'); >> y=load('ex1y.dat'); >> figure >> plot(x,y,'o'); >> ylabel('Height in meters'); >> xlabel('Age in years');</pre> 		

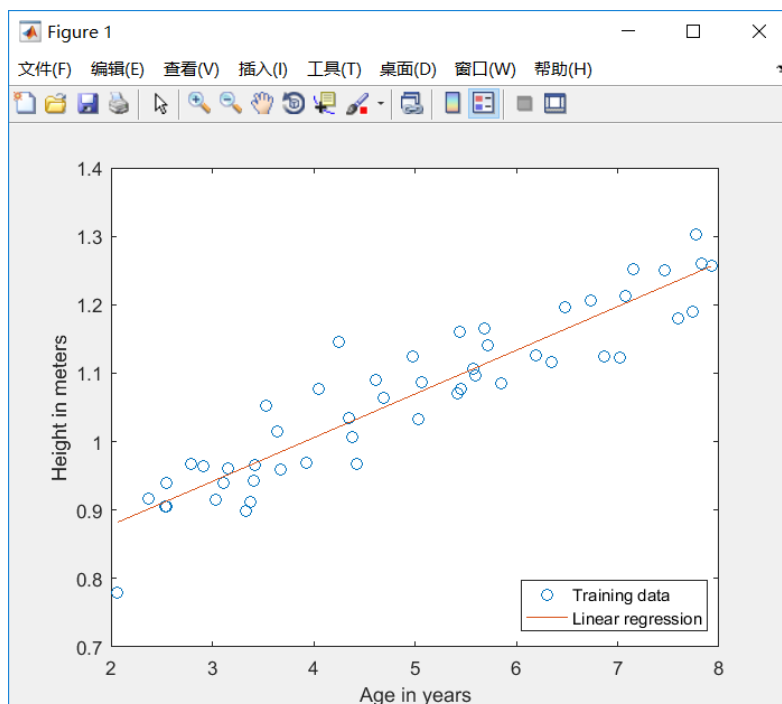
1.2 梯度下降:

```
>> m=length(y);
>> x=[ones(m,1),x];
>> theta=zeros(2,1)
theta =
     0
     0
>> iterations=1500;
>> alpha=0.07;
>> computeCost(x,y,theta);
>> theta=gradientDescent(x,y,theta,alpha,iterations);
```

(其中 `computeCost` 和 `gradientDescent` 函数见文件夹或附件)

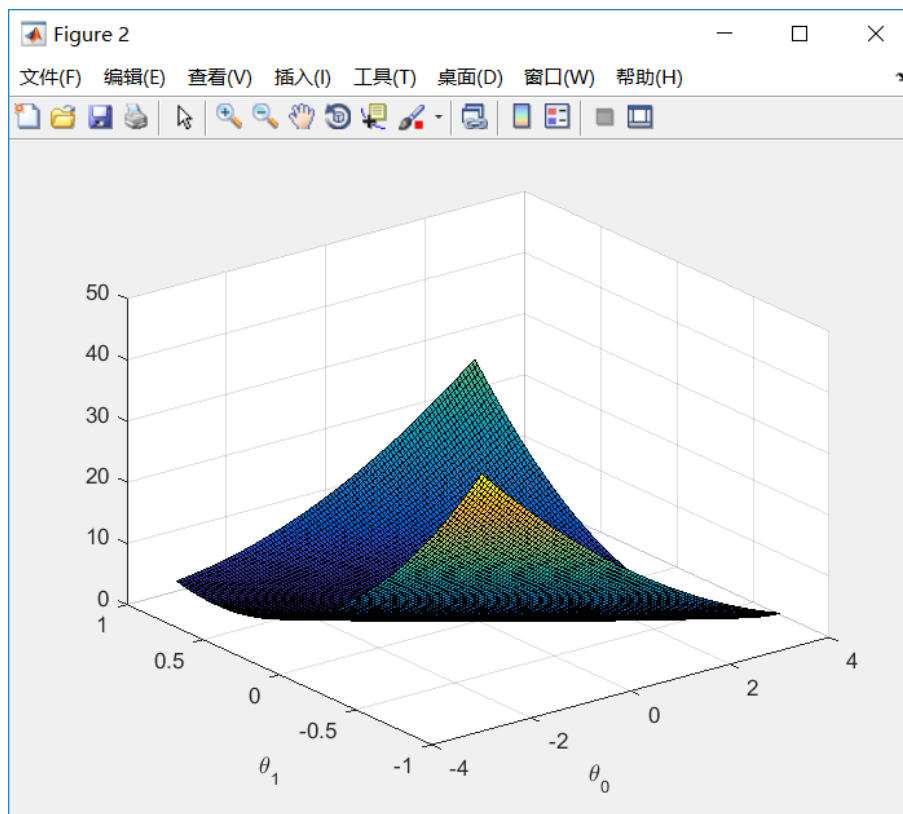
1.3 显示最优 theta 参数和可视化 $J(\theta)$ 进行预测:

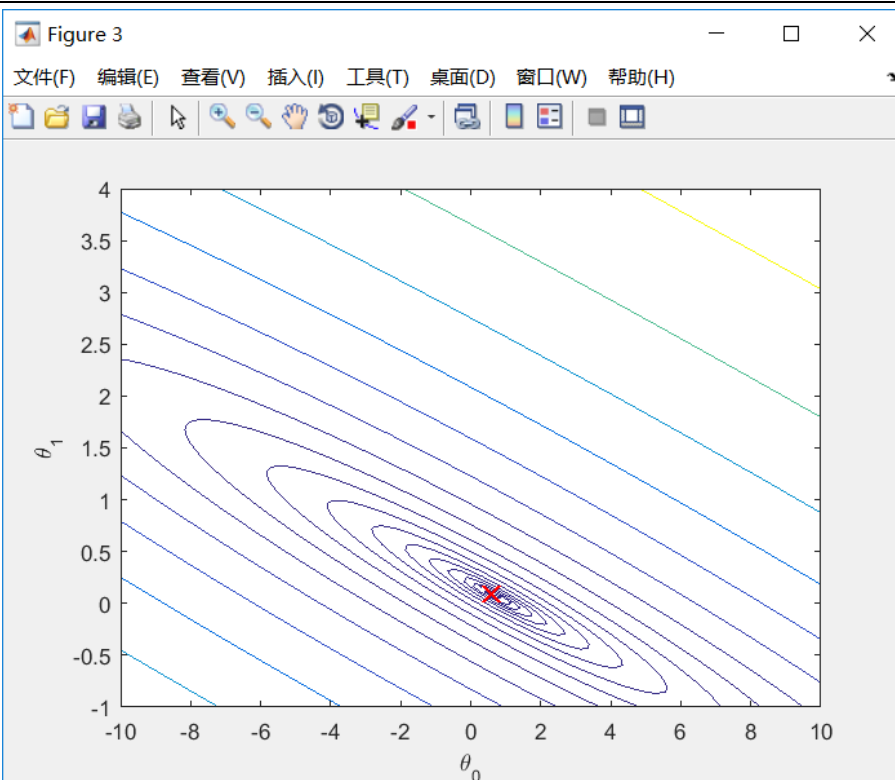
```
>> theta
theta =
    0.7502
    0.0639
>> hold on;
>> plot(x(:,2),x*theta,'-')
>> legend('Training data','Linear regression')
>> hold off
>> pre1=[1,3.5]*theta
pre1 =
    0.9737
>> pre2=[1,7]*theta
pre2 =
    1.1973
```



1.4 可视化 $J(\theta)$:

```
>> J_vals=zeros(100,100);  
>> theta0_vals=linspace(-3,3,100);  
>> theta1_vals=linspace(-1,1,100);  
>> for i=1:length(theta0_vals)  
    for j=1:length(theta1_vals)  
        t=[theta0_vals(i);theta1_vals(j)];  
        J_vals(i,j)=computeCost(x,y,t);  
    end  
end  
>> J_vals=J_vals';  
>> figure;  
>> surf(theta0_vals,theta1_vals,J_vals)  
>> xlabel('\theta_0');  
>> ylabel('\theta_1');
```





结论分析与体会：

比较熟悉掌握了线性回归的用法，有销小收获。

附录：程序源代码

主要算法函数：

1. compute Cost:

```
function J = computeCost(x,y,theta)
%UNTITLED2 此处显示有关此函数的摘要
% 此处显示详细说明
m=length(y);
J = 0;
% 初始化
J = sum((x*theta - y).^2) / (2 * m);
% 计算损失
end
```

2. gradientDescent:

```
function [theta, J_history] = gradientDescent(X, y, theta, alpha, num_iters)
m = length(y);
% 样本数量
J_history = zeros(num_iters, 1);
% 记录 J 最优化的过程, (1500,1), all elements is zero
for iter = 1:num_iters
    H = X * theta;
    % (97,2) * (2*1) = (97,1)
    T = [0 ; 0];
    % (2,1), 记录梯度
```

```
for i = 1 : m,
    T = T + (H(i) - y(i)) * X(i,:)';
    % (1,1)*(1*2)的转置, 结果为 (2, 1)
end
theta = theta - (alpha * T) / m;

J_history(iter) = computeCost(X, y, theta);
% theta 带入, 调用损失函数, 计算损失, 并记录在 J_history 中
end
end
```