山东大学　　　　计算机科学与技术　　　　　　学院

　机器学习与模式识别　　课程实验报告

学号：202300130183	姓名：　宋浩宇	班级：23 级人工智能班
实验题目：Experiment 1: Linear Regression		
实验学时：2	实验日期：2025/2/25	

实验环境：

软件环境：

系统：Windows 11 家庭中文版 23H2 22631.4317

计算软件：MATLAB 版本: 9.8.0.1323502 (R2020a)

Java 版本: Java 1.8.0_202-b08 with Oracle Corporation Java HotSpot(TM) 64-Bit Server VM mixed mode

硬件环境：

CPU：13th Gen Intel(R) Core(TM) i9-13980HX　　2.20 GHz

内存：32.0 GB (31.6 GB 可用)

磁盘驱动器：NVMe WD_BLACKSN850X2000GB

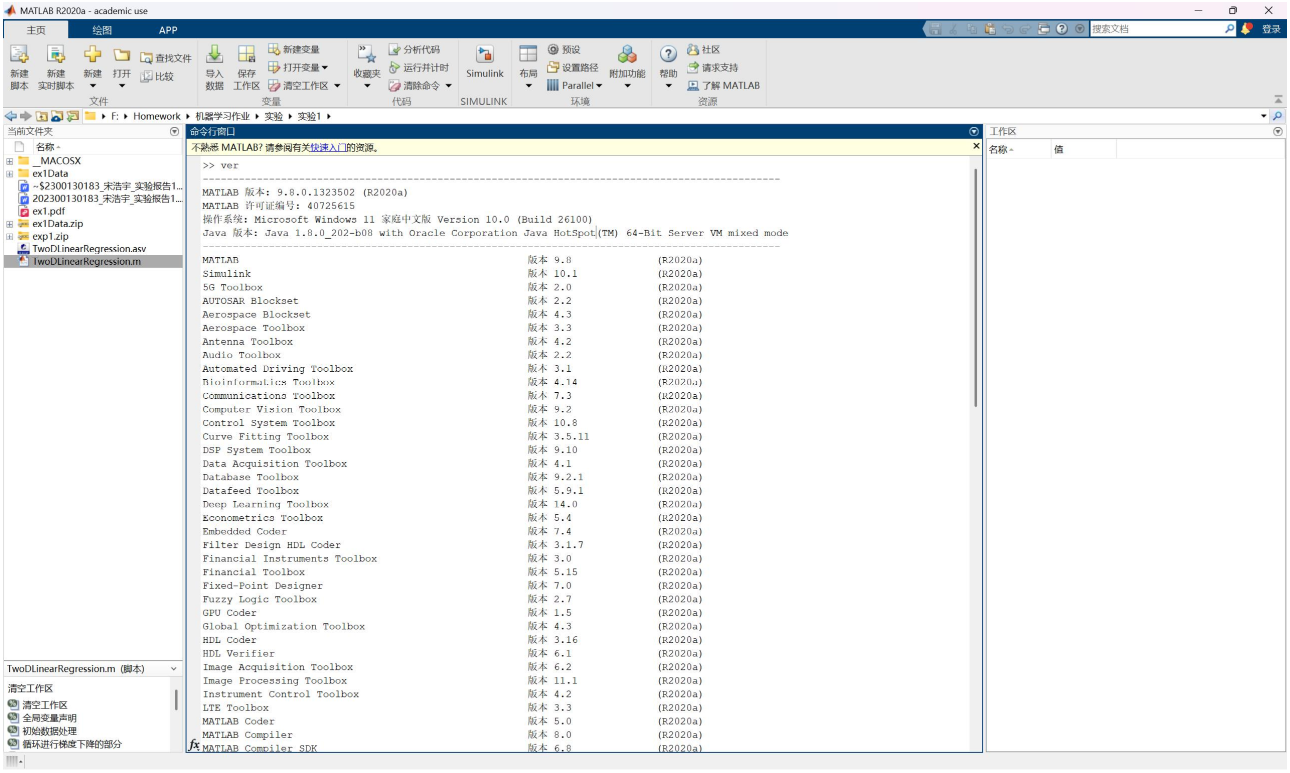
显示适配器：NVIDIA GeForce RTX 4080 Laptop GPU

1. 实验内容（此处为实验指导文件原文）

This first exercise will give you practice with linear regression. These exercises have been extensively tested with Matlab, but they should also work in Octave, which has been called a “free version of Matlab”. If you are using Octave, be sure to install the Image package as well (available for Windows as an option in the installer, and available for Linux from Octave-Forge).

2. 实验步骤

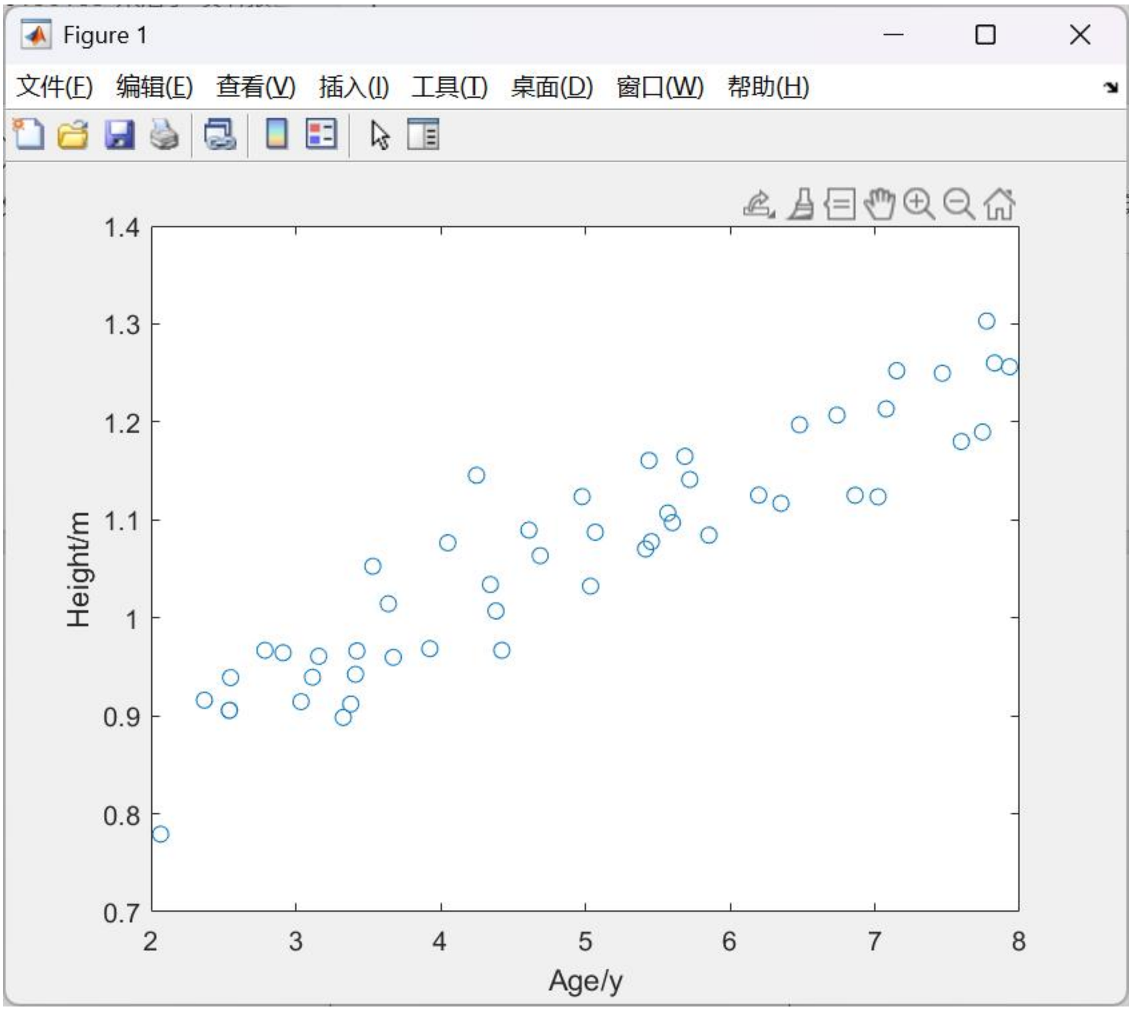
(1) 安装 matlab 并激活。



不熟悉 MATLAB? 请参阅有关快速入门的资源。

```
>> ver
-----------------------------------------------------------------------------------------------
MATLAB 版本: 9.8.0.1323502 (R2020a)
MATLAB 许可证编号: 40725615
操作系统: Microsoft Windows 11 家庭中文版 Version 10.0 (Build 26100)
Java 版本: Java 1.8.0_202-b08 with Oracle Corporation Java HotSpot(TM) 64-Bit Server VM mixed mode
-----------------------------------------------------------------------------------------------
MATLAB                                                版本 9.8                 (R2020a)
Simulink                                              版本 10.1                (R2020a)
5G Toolbox                                            版本 2.0                 (R2020a)
AUTOSAR Blockset                                      版本 2.2                 (R2020a)
Aerospace Blockset                                    版本 4.3                 (R2020a)
Aerospace Toolbox                                     版本 3.3                 (R2020a)
Antenna Toolbox                                       版本 4.2                 (R2020a)
Audio Toolbox                                         版本 2.2                 (R2020a)
Automated Driving Toolbox                             版本 3.1                 (R2020a)
Bioinformatics Toolbox                                版本 4.14                (R2020a)
Communications Toolbox                                版本 7.3                 (R2020a)
Computer Vision Toolbox                               版本 9.2                 (R2020a)
Control System Toolbox                                版本 10.8                (R2020a)
Curve Fitting Toolbox                                 版本 3.5.11              (R2020a)
DSP System Toolbox                                    版本 9.10                (R2020a)
Data Acquisition Toolbox                              版本 4.1                 (R2020a)
Database Toolbox                                      版本 9.2.1               (R2020a)
Datafeed Toolbox                                      版本 5.9.1               (R2020a)
Deep Learning Toolbox                                 版本 14.0                (R2020a)
Econometrics Toolbox                                  版本 5.4                 (R2020a)
Embedded Coder                                        版本 7.4                 (R2020a)
Filter Design HDL Coder                               版本 3.1.7               (R2020a)
Financial Instruments Toolbox                         版本 3.0                 (R2020a)
Financial Toolbox                                     版本 5.15                (R2020a)
Fixed-Point Designer                                  版本 7.0                 (R2020a)
Fuzzy Logic Toolbox                                   版本 2.7                 (R2020a)
GPU Coder                                             版本 1.5                 (R2020a)
Global Optimization Toolbox                           版本 4.3                 (R2020a)
HDL Coder                                             版本 3.16                (R2020a)
HDL Verifier                                          版本 6.1                 (R2020a)
Image Acquisition Toolbox                             版本 6.2                 (R2020a)
Image Processing Toolbox                              版本 11.1                (R2020a)
Instrument Control Toolbox                            版本 4.2                 (R2020a)
LTE Toolbox                                           版本 3.3                 (R2020a)
MATLAB Coder                                          版本 5.0                 (R2020a)
MATLAB Compiler                                       版本 8.0                 (R2020a)
MATLAB Compiler SDK                                   版本 6.8                 (R2020a)
```

(2) 获取实验使用的数据。



(3) 构造模型、损失函数以及设置梯度下降方式。

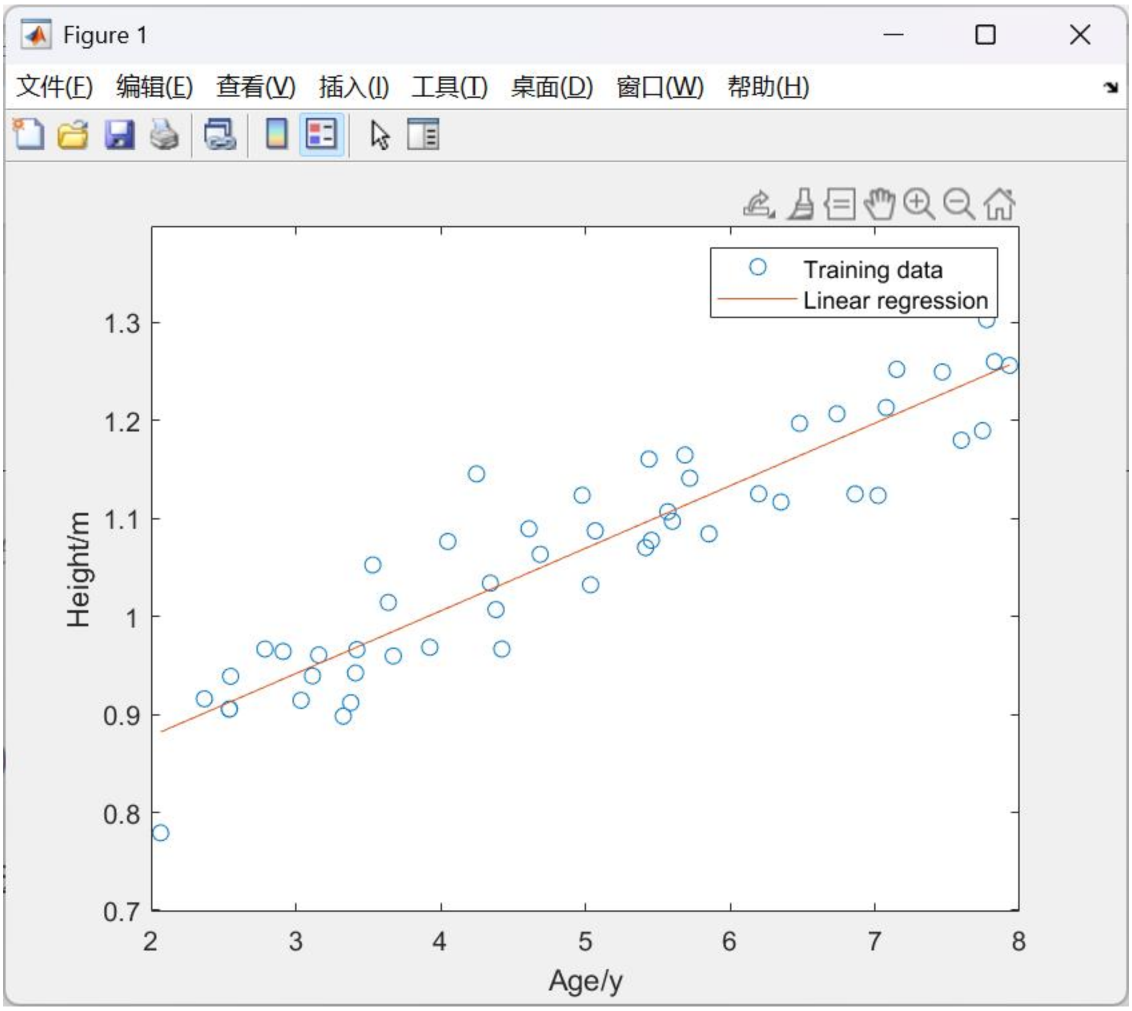
$$h_\theta(x) = \theta^T x = \sum_{i=0}^{n} \theta_i x_i,$$

ıdient descent update rule is

$$\theta_j := \theta_j - \alpha \frac{1}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)}) x_j^{(i)}$$

(4) 用 matlab 代码实现并进行计算。
(5) 进行模型的测试。

3. 测试结果
首先是模型对于原数据集的表现：



然后是模型对于实验指导书提出的问题的的输出：
问题：

Use your model to predict the height for two boys of ages 3.5 and 7.

输出：

```
height of boy who has age 3.5: 0.973742 m
height of boy who has age 7  : 1.197334 m
```

然后是损失函数关于参数取值变化的可视化：

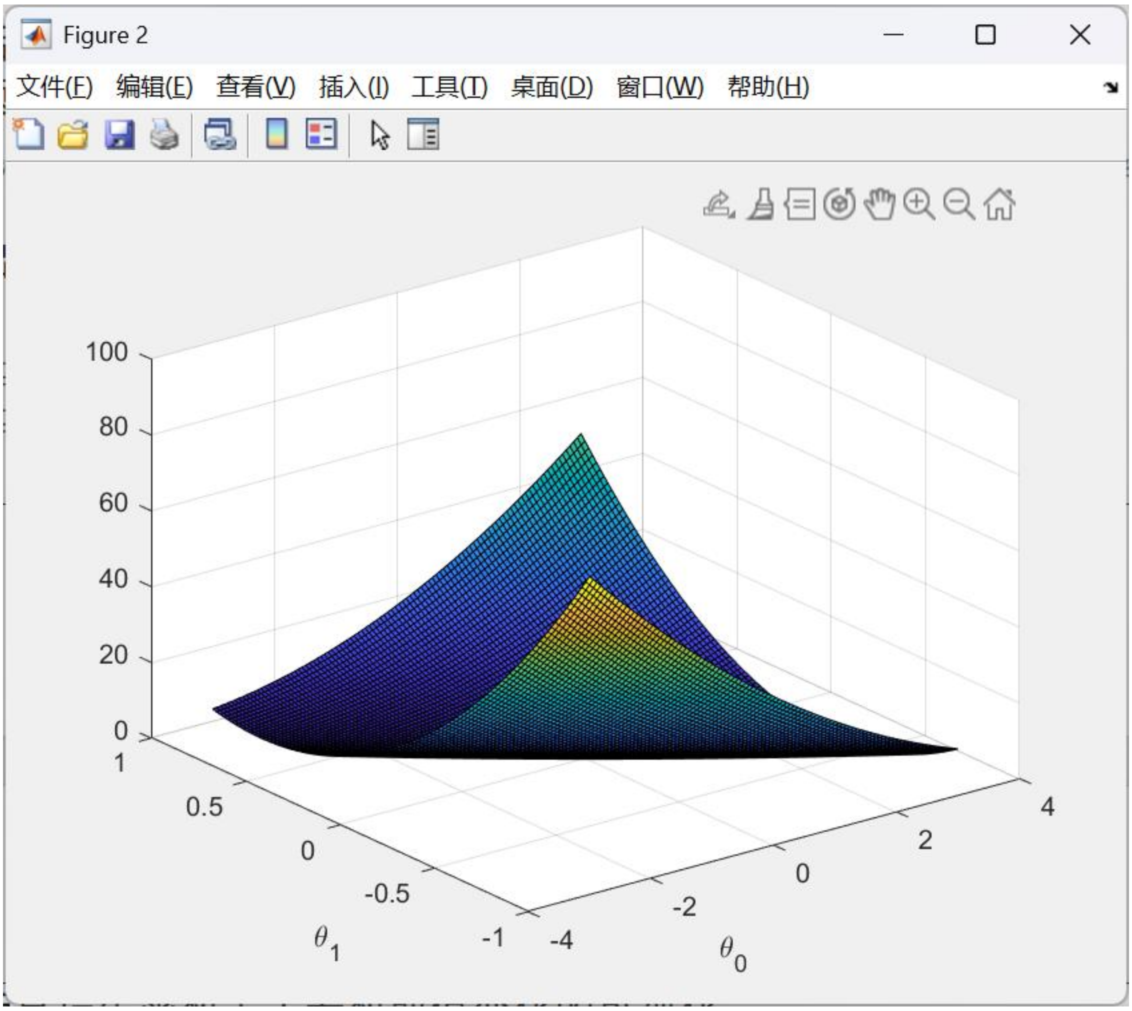


Figure 2
文件(F) 编辑(E) 查看(V) 插入(I) 工具(T) 桌面(D) 窗口(W) 帮助(H)
100
80
60
40
20
0
1
0.5
0
-0.5
-1
θ_1
-4
-2
0
2
4
θ_0

最后是 J（θ）的等高线图

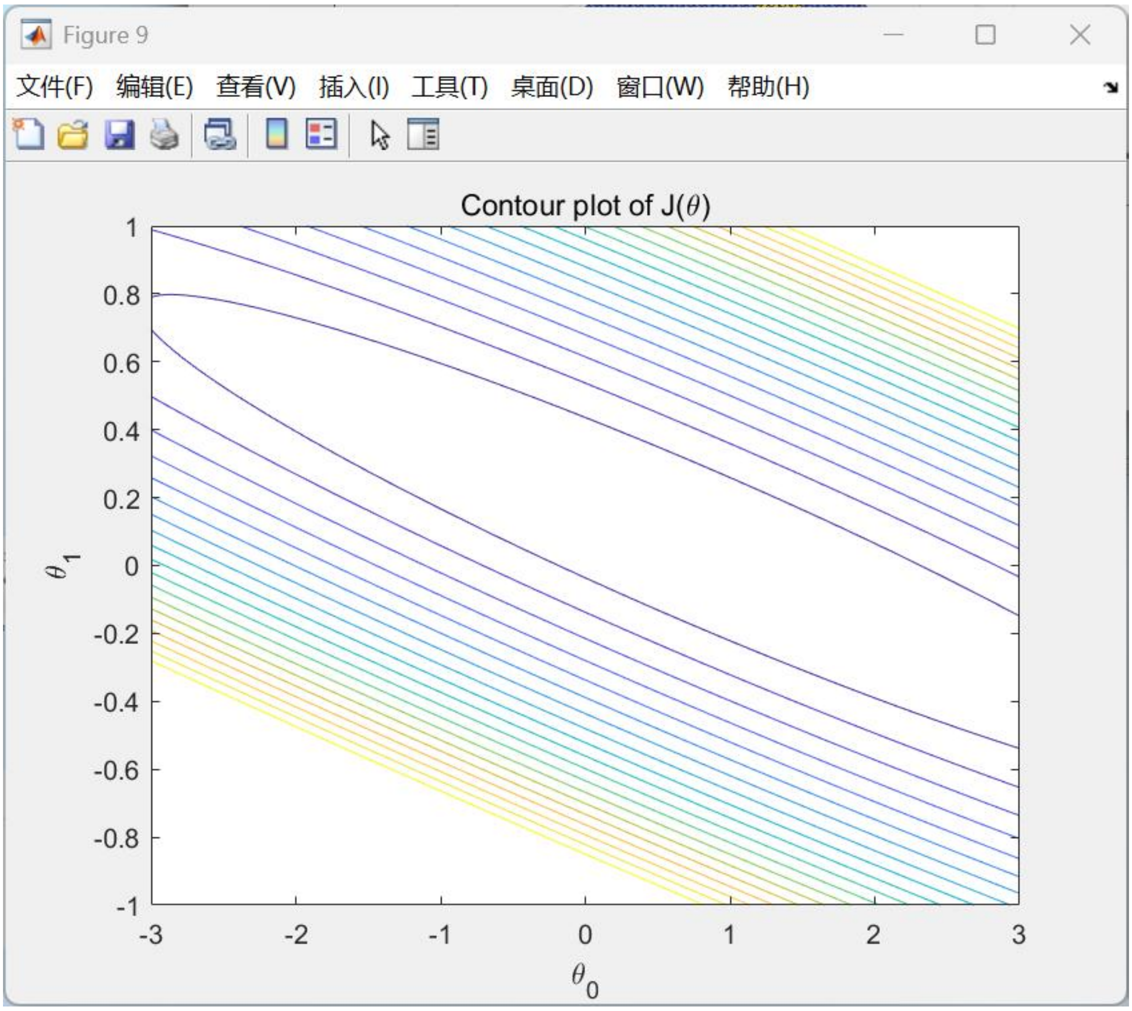


Figure 9
文件(F) 编辑(E) 查看(V) 插入(I) 工具(T) 桌面(D) 窗口(W) 帮助(H)
Contour plot of J(θ)
θ_1
θ_0
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-3
-2
-1
0
1
2
3

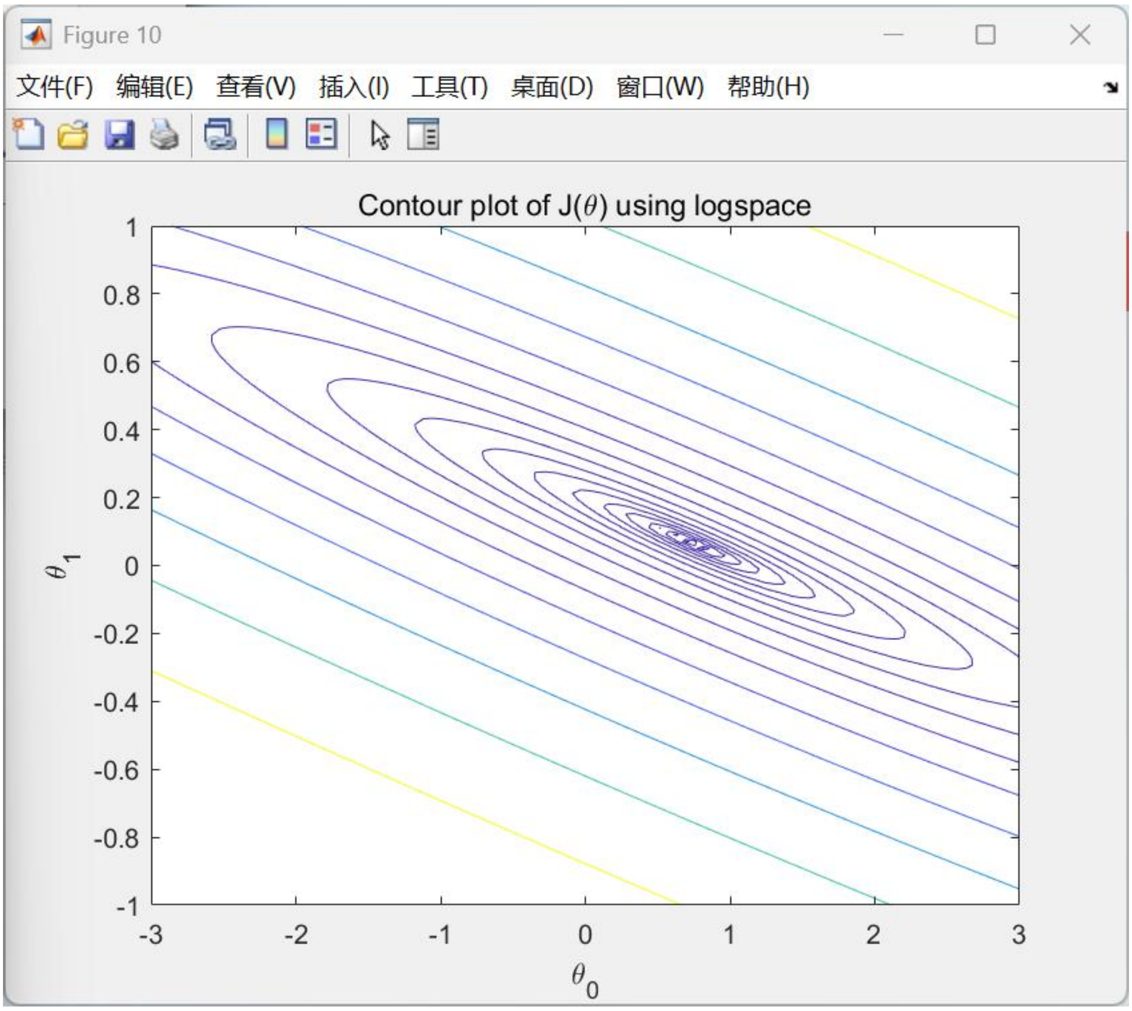


Figure 10
文件(F) 编辑(E) 查看(V) 插入(I) 工具(T) 桌面(D) 窗口(W) 帮助(H)
Contour plot of J(θ) using logspace
1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
θ_1
-3
-2
-1
0
1
2
3
θ_0

4. 附录：实现源代码

```
%% 清空工作区

clear;
clc;


%% 全局变量声明

global alpha;
global theta;

alpha = 0.07;        %学习率设置

theta = [0,0];       %参数初始值

sum_times = 1500;    %梯度下降次数


%% 初始数据处理


x = load("ex1Data/ex1x.dat");
```

```
y = load("ex1Data/ex1y.dat");

% disp(x);
% disp(y);

figure;
plot(x,y,'o');
ylabel("Height/m");
xlabel("Age/y");

m = length(y);

x = [ones(m,1),x];


%% 循环进行梯度下降的部分
for times_cnt = 1:sum_times

h_ans = h(x(1,:));
disp(h_ans);

sum_w = 0;
sum_b = 0;
for i = 1:size(x,1)
sum_w = sum_w + (h(x(i,:)) - y(i)) * x(i,2);
sum_b = sum_b + (h(x(i,:)) - y(i)) * x(i,1);
end
disp(sum_w);
disp(sum_b);


theta(1) = theta(1) - alpha * sum_b / size(x,1);
theta(2) = theta(2) - alpha * sum_w / size(x,1);
end
% 以上为一次更新


%% 打印参数结果


disp(theta);
fprintf("y = %fx + %f\n",theta(2),theta(1));


%% 检验模型效果
```

```
hold on;
plot(x(:,2),h(x),'-');
legend('Training data','Linear regression');


%% 测试预测效果

fprintf("height of boy who has age 3.5: %f m\n",h([1,3.5]));
fprintf("height of boy who has age 7  : %f m\n",h([1,7  ]));

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) = h_j(x,y,t);
end
end

J_vals = J_vals';
figure;
surf(theta0_vals,theta1_vals,J_vals);
xlabel('\theta_0');
ylabel('\theta_1');


%% 模型使用的函数


function result = h(x)
global theta;
% sum = 0;
%
% for i = 1:size(x,1)
%     sum = sum + theta .* x(i,:);
% end


result = theta * x';

end

%% 构造 J 平面使用的函数


function result = h_j(x,y,tem_theta)
sum_j = 0;
% sum_b = 0;
% disp(tem_theta);
% disp(x(1,:));
% disp(x(1,:)*tem_theta);
% disp([1.0,2.0]*[-3;-1]);
for i = 1:size(x,1)
sum_j = sum_j + ((x(i,:) * tem_theta) - y(i)) * ((x(i,:) * tem_theta)
- y(i));
end
% disp(sum_w);
% disp(sum_b);
% sum_w = sum_w / size(x,1);
% sum_b = sum_b / size(x,1);
result = sum_j / size(x,1);
End

%% 使用 linspace 绘制等高线图

figure;
contour(theta0_vals, theta1_vals, J_vals, linspace(0, 30, 20));
xlabel('\theta_0');
ylabel('\theta_1');
title('Contour plot of J(\theta)');

%% 使用 logspace 绘制等高线图


% 使用 logspace 生成等高线的值范围

figure;
contour(theta0_vals, theta1_vals, J_vals, logspace(-3, 1.5, 20));
xlabel('\theta_0');
ylabel('\theta_1');
title('Contour plot of J(\theta) using logspace');
```