**同 济 大 学**

**实验报告**

利用梯度下降实现线性拟合



**课程名称**   最优化原理

**学生姓名**  姚 天 亮

**学 号**  2150248

**学 院** 电子与信息工程学院

1. **作业要求**

本作业题要求使用线性拟合，利用梯度下降法，求解参数和，令，使得预测值和真实值之间的均方误差（MSE）误差最小。定义误差如下：

其中：学习率设定为0.3，最大迭代次数设定为50次，和的初始值可设定为0到1之间的任意数值，我们可以采用随机数进行生成。

1. **理论推导**

本作业题的目的是搜索LOSS在上的最小点。为此我们使用梯度下降法，使得LOSS沿着负梯度方向下降，从而在规定的迭代次数内搜索到相应的符合要求的点。

我们使用的负梯度下降法的一般公式为：

其中：是一个正实数，称为步长。

1. 的含义是：给定一个搜索点，由此点出发，根据向量指定的方向和幅值运动，可以得到新点；之后不断迭代，到达终止条件即可。

梯度下降法的终止条件通常有以下几种：

1. 达到最大迭代次数：在训练模型时，通常会指定最大的迭代次数，到这个迭代次数时，梯度下降算法就会停止运行；
2. 目标函数的值达到一定精度：如可以指定目标函数的值在两次迭代之间的变化小于某个设定值时，算法停止运行；
3. 达到一定的运行时间：可以通过设定算法运行的最大时间来确定算法是否停止。例如，可以指定算法运行时间在某个时间段之后停止。

在本实验中：我们的目标函数LOSS定义为：

在使用梯度下降法时，我们需要分别计算LOSS关于w和b的偏导数的值，并更新它们的值，相关定义如下：

由此，我们可以实现梯度下降的迭代计算，达到迭代50次的题目要求后终止即可。

1. **实验结果**

图1，图2分别展示了线性拟合结果与迭代过程中LOSS值的变化情况与趋势。

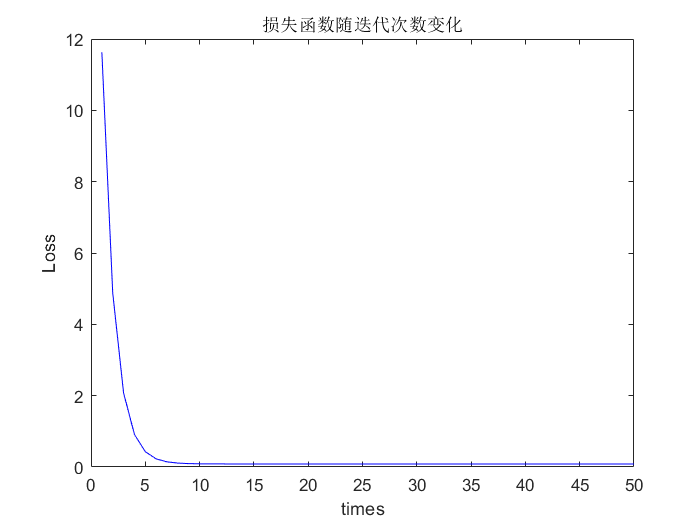


图2 LOSS变化曲线

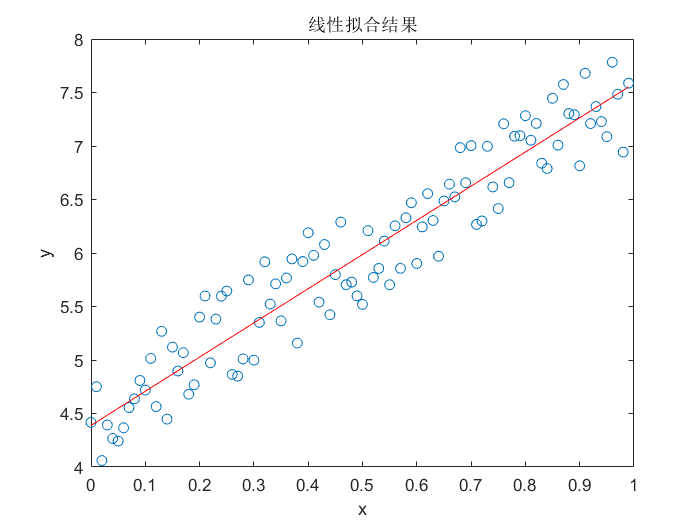


图1 线性拟合结果示意

表1记录了前30次迭代中w、b和 Loss 的值。

表1 前30次迭代中w、b与Loss的值

|  |  |  |  |
| --- | --- | --- | --- |
| 次数 | w | b | Loss |
| 1 | 1.0122 | 1.8359 | 13.674 |
| 2 | 1.6054 | 2.8380 | 5.7524 |
| 3 | 1.9914 | 3.4821 | 2.4568 |
| 4 | 2.2438 | 3.8956 | 1.0847 |
| 5 | 2.4098 | 4.1603 | 0.5127 |
| 6 | 2.5202 | 4.3292 | 0.2738 |
| 7 | 2.5947 | 4.4364 | 0.1735 |
| 8 | 2.6459 | 4.5038 | 0.1311 |
| 9 | 2.6820 | 4.5457 | 0.1129 |
| 10 | 2.7084 | 4.5710 | 0.1049 |
| 11 | 2.7284 | 4.5858 | 0.1011 |
| 12 | 2.7442 | 4.5938 | 0.0992 |
| 13 | 2.7573 | 4.5975 | 0.0980 |
| 14 | 2.7686 | 4.5984 | 0.0971 |
| 15 | 2.7786 | 4.5975 | 0.0966 |
| 16 | 2.7877 | 4.5955 | 0.0961 |
| 17 | 2.7963 | 4.5929 | 0.0955 |
| 18 | 2.8044 | 4.5898 | 0.0946 |
| 19 | 2.8121 | 4.5865 | 0.0945 |
| 20 | 2.8196 | 4.5831 | 0.0941 |
| 21 | 2.8269 | 4.5797 | 0.0937 |
| 22 | 2.8339 | 4.5762 | 0.0933 |
| 23 | 2.8408 | 4.5727 | 0.0930 |
| 24 | 2.8475 | 4.5693 | 0.0926 |
| 25 | 2.8540 | 4.5660 | 0.9232 |
| 26 | 2.8604 | 4.5627 | 0.9200 |
| 27 | 2.8667 | 4.5594 | 0.0916 |
| 28 | 2.8728 | 4.5562 | 0.0913 |
| 29 | 2.8788 | 4.5531 | 0.9111 |
| 30 | 2.8847 | 4.5501 | 0.0904 |

**附录**

**A.代码**

% 2150248 姚天亮 实验一

% 读取数据

x = load('C:\Users\YTL\OneDrive\Desktop\data\_x.txt');

y = load('C:\Users\YTL\OneDrive\Desktop\data\_y.txt');

% 初始化参数

w = rand(1); % 随机生成一个0到1之间的数

b = rand(1); % 随机生成一个0到1之间的数

eta = 0.3; % 学习率

iter = 50; % 最大迭代次数

% 定义损失函数

n = length(x); % 数据的个数

L = @(w,b) (1/(n))\*sum((x\*w+b-y).^2); % 均方误差

% 定义梯度下降法的更新规则

dw = @(w,b) (1/n)\*sum((x\*w+b-y).\*x); % w的偏导数

db = @(w,b) (1/n)\*sum(x\*w+b-y); % b的偏导数

% 初始化损失函数历史记录

costHistory = zeros(iter,1);

% 初始化参数历史记录

wHistory = zeros(iter,1);

bHistory = zeros(iter,1);

% 进行梯度下降法

for i = 1:iter

% 更新参数

w = w - eta\*dw(w,b);

b = b - eta\*db(w,b);

% 计算并记录损失函数值

costHistory(i) = L(w,b);

% 判断是否需要记录参数值

if i <= 30

% 记录参数值

wHistory(i) = w;

bHistory(i) = b;

end

end

% 绘制数据点和拟合直线的图像

figure(1)

plot(x,y,'o') % 绘制数据点

hold on

plot(x,x\*w+b,'r') % 绘制拟合直线

hold off

title('线性拟合结果')

xlabel('x')

ylabel('y')

% 绘制损失函数随迭代次数变化的图像

figure(2)

plot(1:iter,costHistory,'b')

title('损失函数随迭代次数变化')

xlabel('times')

ylabel('Loss')

% 创建一个表格，将wHistory、bHistory和costHistory的前30个值放入其中

T = table((1:30)',wHistory(1:30),bHistory(1:30),costHistory(1:30),'VariableNames',{'times','w','b','Loss'});

% 显示表格

disp(T)