The Experiment Report of Machine Learning



Supervisor:

 Qingyao Wu

**SUBJECT:**SOFTWARE ENGINEERING

**SCHOOL:** SCHOOL OF SOFTWARE ENGINEERING

Author:

林庆晓

Grade:

Undergraduate

Student ID：

201730293209

November 21, 2019

[[1]](#footnote-1)线性回归的闭式解

Abstract—求线性回归的闭式解，并输出loss值

# INTRODUCTION

在统计学中，线性回归(Linear Regression)是利用称为线性回归方程的最小平方函数对一个或多个自变量和因变量之间关系进行建模的一种回归分析。线性回归通过已有的X和Y预测出一个新的x所对应的y值。

线性回归广泛应用于金融、经济学领域，可用于股票、证券价格的预测。线性模型形式简单、易于建模，应用广泛，且许多功能更为强大的非线性模型可在线性回归的基础上引入层级结构或高维映射得到。

本实验采用L2范式损失函数，通过最小二乘法推导出线性回归的闭式解，然后输出在训练集和验证集上的损失函数值。

# METHODS AND THEORY

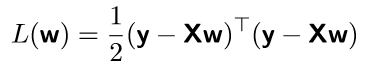
1.线性回归直线方程：



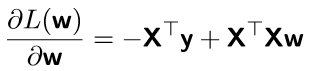
其中w0可以通过增添一列全1到w种来等效

2.确定损失函数：L2范式损失函数（L1范式损失函数也可，不过效果没有L2范式损失函数好）

**矩阵形式**



3.求解闭式解：通过损失函数L对权重w求导出闭式解



令上式为0可求得



4.计算损失函数值：通过上面的损失函数计算闭式解的权重在训练集和验证集上的损失函数值并输出

# Experiment

1.使用sklearn库的load\_svmlight\_file函数读取LIBSVM中的Housing\_scaled数据集，并将其切分为训练集，验证集

2.选取L2范式损失函数用于计算训练集上的损失函数值Loss，并输出

3.使用闭式解公式更新权重参数w

4.在训练集和验证集上用更新后的w计算损失函数值loss\_train和loss\_val，并输出

# conclusion

得到的实验结果：

Loss：



loss\_train:



loss\_val:



总结：通过数学公式的推导，非常简单且方便地就获得了闭式解，理论的支撑对于软件工程也非常重要。不过由于闭式解太过简单，所以也并没有可优化的空间。

1. [↑](#footnote-ref-1)