

# 深度学习小作业 1 报告

刘翰文 522030910109

## 1. 线性回归问题

线性回归问题可以表述为：给定一批数据  $x_1, x_2, \dots, x_n \in R^m$  以及其对应真实标签  $y_1, y_2, \dots, y_n \in R$ ，希望可以找到一个线性关系使得可以用给定数据预测它的标签。即假设  $x_i = (a_1, a_2, \dots, a_m)$  以及对应标签  $y_i$ ，希望可以找到一组权重参数  $w_1, w_2, \dots, w_m, b$ ，使得  $y_i$  尽可能接近数据与权重参数的线性组合  $w_1 a_1 + w_2 a_2 + \dots + w_m a_m + b$ ，其中  $b$  是偏置参数。预测标签  $\hat{y}_i$  与真实标签  $y_i$  的差距可以用一个函数表示，即损失函数。在线性回归中常用平方损失函数  $L(w, b) = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} (\hat{y}_i - y_i)^2$  表示。为了最小化损失函数，常使用随机梯度下降的优化算法减小损失  $(w, b) \leftarrow (w, b) - \eta \nabla_{(w, b)} L(w, b)$ 。

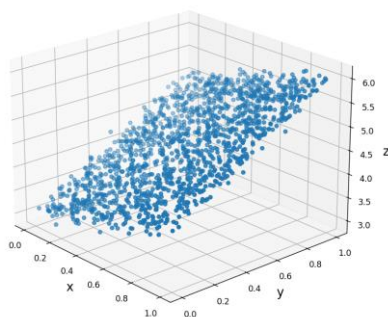
## 2. 算法实现

### 2.1 Jittor 框架配置

将 Jittor 配置到本地参考了 Jittor 官方文档，通过与本机适配的 Jittor 的官方命令行安装代码，完成了 Jittor 在本地的部署。通过对官方文档的学习与自己线性回归的理解，将线性回归的代码实现分为数据生成、模型搭建、模型训练、生成结果四个部分。

### 2.2 数据生成

本次作业使用了简单的二维数据  $x_i = (a_1, a_2)$ ，利用 numpy 的 random 函数生成一系列数据点  $x_1, x_2, \dots, x_m$ ，并定义权重参数  $(w_1, w_2, b) = (1, 2, 3)$ 。利用数据和权重参数生成数据的真实标签  $y_i = w_1 a_1 + w_2 a_2 + b = a_1 + 2a_2 + 3$ ，同时加上一个在  $(-0.1, 0.1)$  之间按高斯分布的随机噪声，模拟真实的数据，生成的数据如下图所示



### 2.3 模型搭建

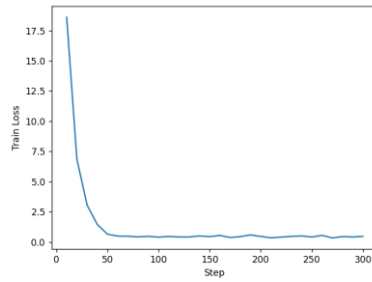
本次使用的模型仅包含了一个简单的线性层 `nn.Linear(3,1)`，包含了三个权重参数分别用来表示  $w_1, w_2, b$ 。并且拓展了输入数据  $x_i = (a_1, a_2)$  到  $x_i = (a_1, a_2, 1)$  用来对齐线性层的输入维度，1 表示和偏置参数相乘使得得到的标签都由一倍的偏置参数加得。

### 2.4 模型训练

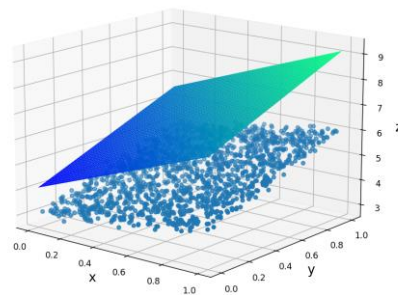
定义损失函数为平方损失函数  $L(w, b) = \frac{1}{m} \sum_{i=1}^m \frac{1}{2} (\hat{y}_i - y_i)^2$ ，定义优化器为随机梯度下降 SGD，用 0.01 的学习率将所有数据送入模型进行训练，记录训练过程中的损失下降情况。

### 2.5 生成结果

通过模型训练中得到的一系列训练损失数据进行画图，得到结果如图

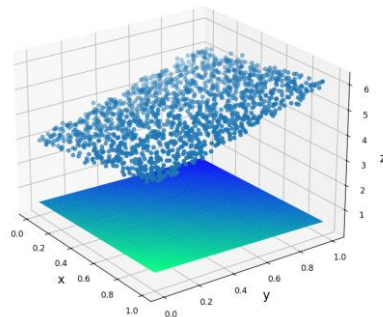


将训练结束后的模型参数输出得到权重参数 $w_1, w_2, b$ ，并在三维图中进行画图，得到三维空间的一个超平面，同时添加真实数据的三维散点，得到如下结果图

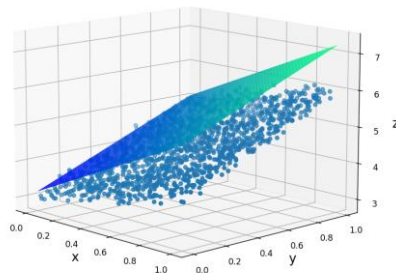


### 3. 结果分析

从得到的训练损失下降结果图来看，训练在 Step=50 时基本收敛，Loss 趋于 0 且基本稳定。从结果的超平面和散点分布情况图来看，得到的超平面和理想的超平面仍有一定误差，打印出训练得到的参数 $w_1 = 2.6, w_2 = 2.4, b = 4.1$ ，与训练之前的超平面比较



可以看得出，训练之后的超平面更趋于理想超平面。同时推测误差出现可能是由于数据范围太小，精度要求高，随机梯度下降和学习率无法准确的优化，于是减小了学习率，增加了训练数据，得到结果



相比第一次得到的结果，新的超平面更接近理想的超平面，可见通过优化，模型的线性回归精度得到了一定的提升。