

3.3 线性回归的简洁实现重点摘录与练习解答

问题解答

1、如果将小批量的总损失替换为小批量损失的平均值，需要如何更改学习率？

解：该问有问题，默认MSELoss就是返回小批量损失的平均值，可防止梯度值过大，并方便调节学习率。根据

$$\mathbf{w}_t = \mathbf{w}_{t-1} - \eta \frac{\partial l}{\partial \mathbf{w}_{t-1}}$$

可知，若用小批量的总损失替换小批量损失的平均值，需要将学习率除以 *batch_size*。

2、查看深度学习框架文档，它们提供了哪些损失函数和初始化方法？用Huber损失代替原损失，即

$$l(y, y') = \begin{cases} |y - y'| - \frac{\sigma}{2} & \text{if } |y - y'| > \sigma \\ \frac{1}{2\sigma}(y - y')^2 & \text{其它情况} \end{cases}$$

解：使用 `dir(nn)` 可查到 PyTorch 提供了 BCELoss, BCEWithLogitsLoss 等多种损失函数。

HuberLoss的优点在于：

1. 增强MSE的离群点鲁棒性，减小了对离群点的敏感度问题；
2. 误差较大时使用MAE可降低异常值影响，使得训练更加健壮；
3. Huber Loss下降速度介于MAE与MSE之间，弥补了MAE在Loss下降速度慢的问题，而更接近MSE。

3、如何访问线性回归的梯度

解：直接看代码即可。

```
print(net[0].weight.grad, net[0].bias.grad)
```