XI'AN JIAOTONG UNIVERSITY

西安交通大学

本科生实验报告

机器学习技术综合训练

实验 1：使用神经网络识别手写数字

姓名： 杨豪

班级： 软件 2101

时间： 2023 年 4 月 5 日

学号： 2206213297

目录

实验 1：使用神经网络识别手写数字

1.1 问题描述

试根据神经网络的授课内容，结合多分类问题的四个反向传播方程，实现一个手写数字识别程序，使其可以对 MNIST 数据集图片中的手写数字进行识别。其中，神经网络包含两个隐层，第一个隐层的神经元个数为 192，第二个隐层的神经元个数为 30

损失函数采用交叉熵损失函数，激活函数使用 Sigmoid 激活函数

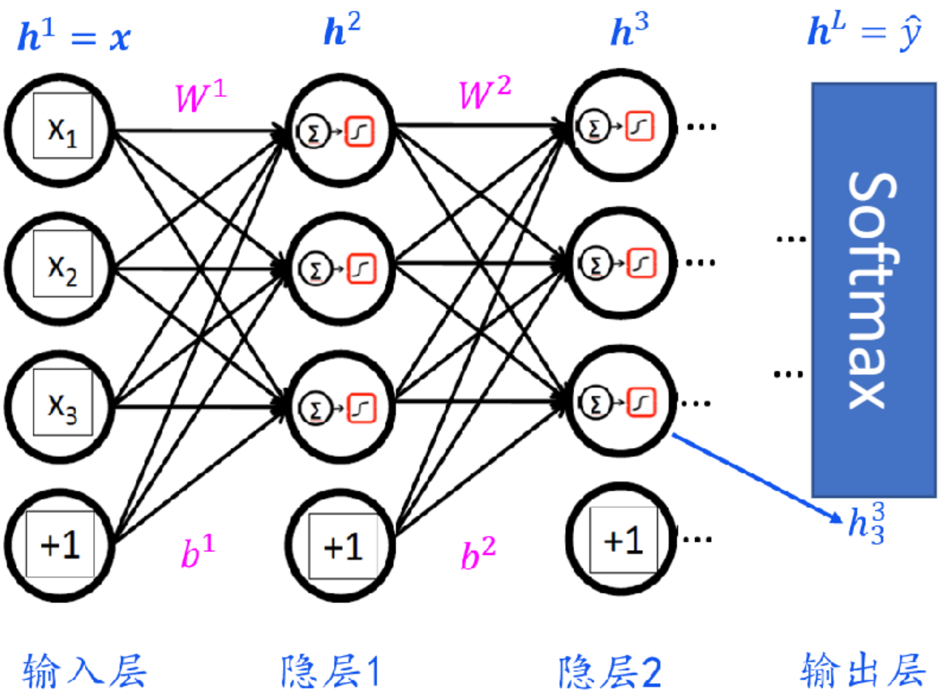
1.2 实验原理

神经网络算法介绍

Algorithm 1: Neutral Network in SGD

输入： 训练集合 x_i，$i = 1, \cdots, n$

for i $in\ 1 : n/m$ **do**
  (1). 向前传播
  for l $in\ 2, \cdots, L$ **do**
    $\mathbf{z^{x,l}} := \left(W^{l-1}\right)^T h^{x,l-1} + b^{x,l-1}$
    $\mathbf{h^{x,l}} := \sigma\left(z^{x,L}\right)$
  end
  (2). 输出误差：$\delta^{x,L}$
  $\boldsymbol{\delta}^{\mathbf{x,L}} = \left(h^{x,L} - y^x\right)$
  (3). 反向传播误差：$\delta^{x,l}$
  for l $in\ L-1 : 2$ **do**
    $\delta^{\mathbf{x,l}} = \sigma'\left(z^{x,l}\right) \odot \left(W^l \delta^{x,l+1}\right)$
  end
  (4). 梯度下降
  for l $in\ 2 : L$ **do**
    $\mathbf{W^{l-1}} := W^{l-1} - \frac{\eta}{m}\sum_x \frac{\partial E}{\partial W^{l-1}} = W^{l-1} - \frac{\eta}{m}\sum_x h^{x,l-1}\left(\delta^{x,l}\right)^T$
    $\mathbf{b^{l-1}} := b^{l-1} - \frac{\eta}{m}\sum_x \frac{\partial E}{\partial b^{l-1}} = b^{l-1} - \frac{\eta}{m}\sum_x \delta^{x,l}$
  end
end



1.3 框架代码解读

题目已给定的代码有三个部分

- main.py 主函数，限定了神经网络的层数和每层的节点数，并修改工作路径以方便数据读入和输出
- mnist_loader.py 由相对路径加载 mnist 数据集并读取为方便 network.py 处理的形式
- network.py 神经网络的主干，上面提到的算法由 update_mini_batch 实现，剩余部分主要负责提供 loss 函数的求值、求导和绘图输出
  - update_mini_batch 函数以最小批的方法更新参数，调用 backprop（对应算法描述中的前三步）得到反向传播误差，并应用到梯度下降上（对应算法描述中的第四步）

1.4 对于 TODO 的解答

一共有 4 个 TODO，都是参照算法描述和代码采用的数据类型调用对应的 numpy 实现，没有什么难度

1.4.1 计算 BP1

该部分对应算法描述中的第 2 步，已给出实现

```
# TODO: 此处为BP1的计算，已经给出参考代码如下三行
delta = (self.cost).delta(zs[-1], activations[-1], y)
nabla_b[-1] = delta
nabla_w[-1] = np.dot(activations[-2], delta.transpose())
```

1.4.2 计算 δ

该部分对应算法描述中的第 3 步，注意 np.dot 和 * 的区别和递推关系中等号两边矩阵的上标

```
# TODO: 完成 delta 的计算，对应PPT中的BP2
delta = sp * np.dot(self.weights[-l + 1], delta)
```

1.4.3 计算 BP3

该部分对应算法描述中的第 4 步中计算 $\dfrac{\partial E}{\partial W^{l-1}} = h^{x,l-1}\left(\delta^{x,l}\right)^T$ 和 $\dfrac{\partial E}{\partial b^{l-1}} = \delta^{x,l}$，注意递推关系中的先后次序和等号两边矩阵的上标

```
# TODO: 完成 Partial E / Partial b 和 Partial E / Partial w, 对应PPT中的BP3
    和BP4
nabla_b[-l] = delta
nabla_w[-l] = np.dot(activations[-l - 1], delta.transpose())
```

1.4.4 计算 mini_batch 的 nabla_b 和 nabla_w

该部分对应算法描述中的第 4 步，注意 nabla_b 和 nabla_w 的数据类型为列表，不可以直接相加

```
# TODO: 完成一个mini_batch中 nabla_b 和 nabla_w 的计算
nabla_b = [n_b + d_n_b for n_b, d_n_b in  zip(nabla_b, delta_nabla_b)]
nabla_w = [n_w + d_n_w for n_w, d_n_w in  zip(nabla_w, delta_nabla_w)]
```

1.5 结果展示

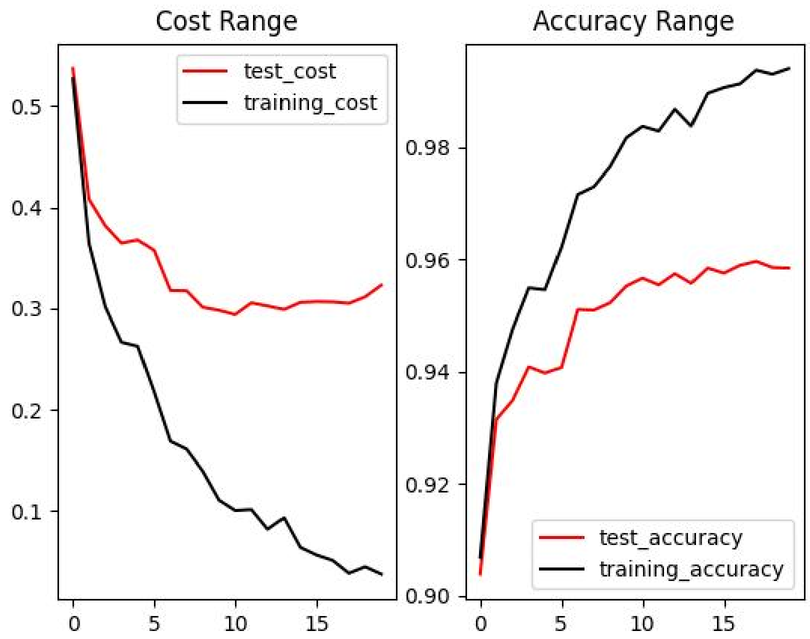


图 1: result

Listing 1: 程序输出

```
Epoch 0 training complete
Cost on training data: 0.4864990500995526
Accuracy on training data: 45648.0 / 50000
Cost on evaluation data: 0.5171438222818319
Accuracy on evaluation data: 9082.0 / 10000

Epoch 1 training complete
Cost on training data: 0.32762718180961936
Accuracy on training data: 47155.0 / 50000
Cost on evaluation data: 0.38299592178724307
Accuracy on evaluation data: 9341.0 / 10000

Epoch 2 training complete
Cost on training data: 0.3086852753678556
Accuracy on training data: 47275.0 / 50000
Cost on evaluation data: 0.399539299143793
Accuracy on evaluation data: 9345.0 / 10000

```

```
Epoch 3 training complete
Cost on training data: 0.24972857496684978
Accuracy on training data: 47834.0 / 50000
Cost on evaluation data: 0.3585633140924723
Accuracy on evaluation data: 9397.0 / 10000

Epoch 4 training complete
Cost on training data: 0.22723462892513618
Accuracy on training data: 47965.0 / 50000
Cost on evaluation data: 0.35218769857094817
Accuracy on evaluation data: 9446.0 / 10000

Epoch 5 training complete
Cost on training data: 0.16733928319261923
Accuracy on training data: 48490.0 / 50000
Cost on evaluation data: 0.2993975071402639
Accuracy on evaluation data: 9539.0 / 10000

Epoch 6 training complete
Cost on training data: 0.13263494238211188
Accuracy on training data: 48820.0 / 50000
Cost on evaluation data: 0.2848000556186426
Accuracy on evaluation data: 9534.0 / 10000

Epoch 7 training complete
Cost on training data: 0.12413167242826531
Accuracy on training data: 48926.0 / 50000
Cost on evaluation data: 0.280473272887883
Accuracy on evaluation data: 9561.0 / 10000

Epoch 8 training complete
Cost on training data: 0.10126083659446701
Accuracy on training data: 49146.0 / 50000
Cost on evaluation data: 0.2909174340393559
Accuracy on evaluation data: 9552.0 / 10000

```

```
Epoch 9 training complete
Cost on training data: 0.07950659594425181
Accuracy on training data: 49331.0 / 50000
Cost on evaluation data: 0.28240616318534334
Accuracy on evaluation data: 9589.0 / 10000

Epoch 10 training complete
Cost on training data: 0.07638307314555179
Accuracy on training data: 49361.0 / 50000
Cost on evaluation data: 0.2893474599986748
Accuracy on evaluation data: 9598.0 / 10000

Epoch 11 training complete
Cost on training data: 0.07294671256650084
Accuracy on training data: 49367.0 / 50000
Cost on evaluation data: 0.30407742756999756
Accuracy on evaluation data: 9567.0 / 10000

Epoch 12 training complete
Cost on training data: 0.07402528046767626
Accuracy on training data: 49372.0 / 50000
Cost on evaluation data: 0.3076031010374318
Accuracy on evaluation data: 9555.0 / 10000

Epoch 13 training complete
Cost on training data: 0.047381324957852036
Accuracy on training data: 49610.0 / 50000
Cost on evaluation data: 0.2991721910718146
Accuracy on evaluation data: 9613.0 / 10000

Epoch 14 training complete
Cost on training data: 0.039001067374808
Accuracy on training data: 49705.0 / 50000
Cost on evaluation data: 0.28855569143429394
Accuracy on evaluation data: 9589.0 / 10000

```

```
Epoch 15 training complete
Cost on training data: 0.02830440419954917
Accuracy on training data: 49800.0 / 50000
Cost on evaluation data: 0.28475461901351434
Accuracy on evaluation data: 9619.0 / 10000

Epoch 16 training complete
Cost on training data: 0.02440925100559722
Accuracy on training data: 49846.0 / 50000
Cost on evaluation data: 0.28767409965738916
Accuracy on evaluation data: 9626.0 / 10000

Epoch 17 training complete
Cost on training data: 0.01862185593626093
Accuracy on training data: 49883.0 / 50000
Cost on evaluation data: 0.29037706748223857
Accuracy on evaluation data: 9645.0 / 10000

Epoch 18 training complete
Cost on training data: 0.013902421587521778
Accuracy on training data: 49932.0 / 50000
Cost on evaluation data: 0.28878484204108124
Accuracy on evaluation data: 9639.0 / 10000

Epoch 19 training complete
Cost on training data: 0.014437471805492598
Accuracy on training data: 49916.0 / 50000
Cost on evaluation data: 0.29415119177326415
Accuracy on evaluation data: 9635.0 / 10000
```