

Exercise1-Softmax Regression

在这个练习中，需要训练一个分类器来完成对MNIST数据集中 0-9 10个手写数字的分类。

初始代码可以在课堂提供的压缩包的 `ex1/` 中找到。

MNIST数据集可在[数据集下载地址](#)中下载，共有4个.gz文件，解压后放在初始代码的 `ex1/` 目录下。

初始代码中已经包含以下内容：

- 在 `main.py`文件中加载训练和测试数据集
- 在 `main.py`文件中调用训练和测试相关函数
- 在 `train.py`文件中调用 `softmax_regression()` 函数

当训练完成，分类器会输出10个类别的测试准确率。

你的任务是在 `softmax_regression.py` 文件中实现 `softmax_regression()` 函数，计算每一次迭代的损失值 $J(\theta, x, y)$ ，将它存储在变量 f 中，并计算梯度 $\nabla_{\theta} J(\theta, x, y)$ ，将它存储在变量 g 中。初始代码会将 θ 的形状定义为一个 $k \times n$ 的矩阵 ($K=10$ 个类别)。

此外，你需要在 `evaluate.py`文件中实现`cal_accuracy()`函数，输出分类器在测试集上的准确率。

一开始你可以使用 `for循环` 来得到正确的梯度（最好尽早使用调试策略来进行梯度检查）。之后你会发现用 `for循环` 实现的代码计算速度太慢了。当你得到正确的梯度之后，尽可能尝试对你的代码进行优化，再用完整的数据集来跑实验。

程序中涉及到的超参可以自行设置，尽可能得到较高的分类准确率。

实验环境：硬件环境：CPU。软件环境：Python 3.7, Numpy。

作业以压缩包方式提交，压缩包命名以本次作业为例，命名为 **“第一次作业-学号-姓名”**。