

# Exercise1-Softmax Regression

## 问题描述

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中的 `ex1/softmax_regression.py` 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：

- 加载训练和测试数据集
- 训练和测试主体代码
- 在 `train()` 函数中调用 `softmax_regression()` 函数和 `cost()` 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` 目标函数，计算损失值  $J(\theta, x, y)$ ，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  $\theta$  的形状定义为一个 `kxn` 的矩阵 ( $K=10$ 个类别)。此外，你需要完成函数 `cal_accuracy()`，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。