

**《机器学习》课程实验报告**

**学 院 软件学院**

**专 业 软件工程**

**组 员**  颜常霖

**学 号 201530613368**

**邮 箱 693812981@qq.com**

**指导教师**  **吴庆耀**

**提交日期** **2017年 12月 15 日**

## 1. 实验题目: 逻辑回归、线性分类与随机梯度下降

## 2. 实验时间：2017年 12月 2日

## 3. 报告人:颜常霖

## 4. 实验目的:

（1）对比理解梯度下降和随机梯度下降的区别与联系。

（2）对比理解逻辑回归和线性分类的区别与联系。

（3）进一步理解SVM的原理并在较大数据上实践。

## 5. 数据集以及数据分析：

实验使用的是[LIBSVM Data](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/" \t "https://www.zybuluo.com/chenyaofo/note/_blank)的中的[a9a](https://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary.html" \l "a9a" \t "https://www.zybuluo.com/chenyaofo/note/_blank)数据，包含32561 / 16281(testing)个样本，每个样本有123/123 (testing)个属性，标签为+1，-1。

## 实验步骤:

**逻辑回归与随机梯度下降**

（1）读取实验训练集和验证集。

（2）逻辑回归模型参数初始化，可以考虑全零初始化，随机初始化或者正态分布初始化。

（3）选择Loss函数及对其求导。

（4）求得部分样本对Loss函数的梯度G。

（5）使用不同的优化方法更新模型参数（NAG，RMSProp，AdaDelta和Adam）。

（6）选择合适的阈值，将验证集中计算结果大于阈值的标记为正类，反之为负类。在验证集上测试并得到不同优化方法的Loss函数值，，和。

（7）重复步骤4-6若干次，画出，，和随迭代次数的变化图。

**线性分类与随机梯度下降**

（1）读取实验训练集和验证集。

（2）支持向量机模型参数初始化，可以考虑全零初始化，随机初始化或者正态分布初始化。

（3）选择Loss函数及对其求导，过程详见课件ppt。

（4）求得部分样本对Loss函数的梯度。

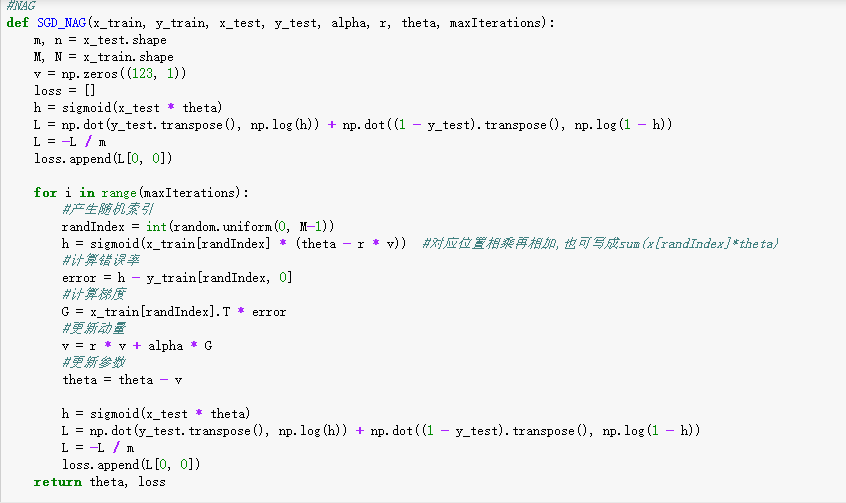
（5）使用不同的优化方法更新模型参数（NAG，RMSProp，AdaDelta和Adam）。

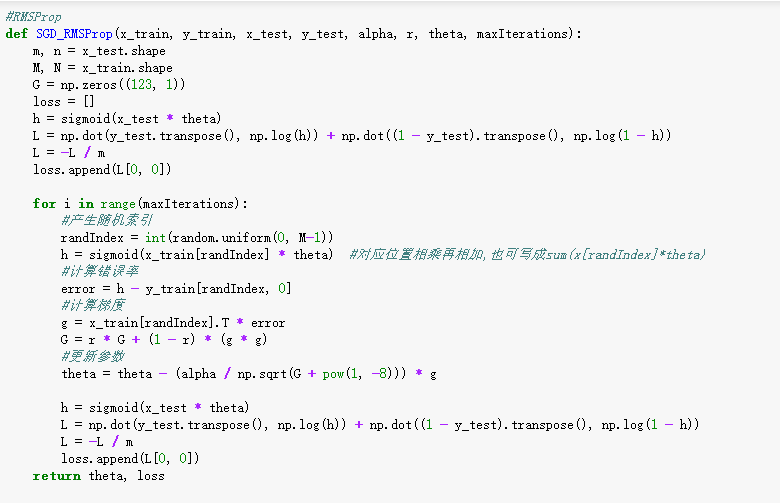
（6）选择合适的阈值，将验证集中计算结果大于阈值的标记为正类，反之为负类。在验证集上测试并得到不同优化方法的Loss函数值，，和。

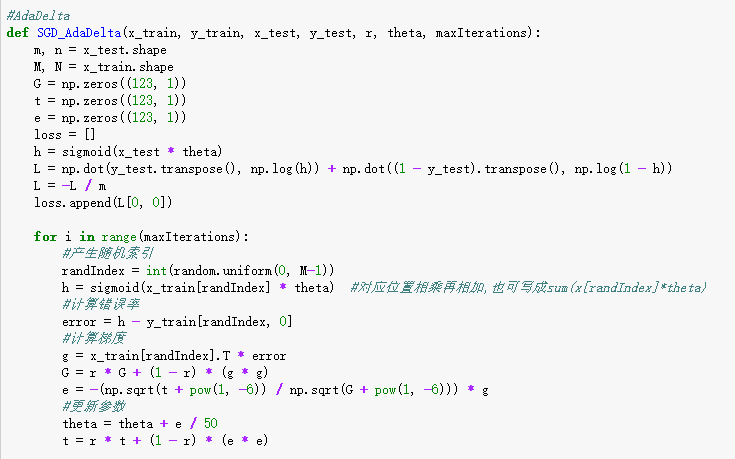
（7）重复步骤4-6若干次，画出，，和随迭代次数的变化图。

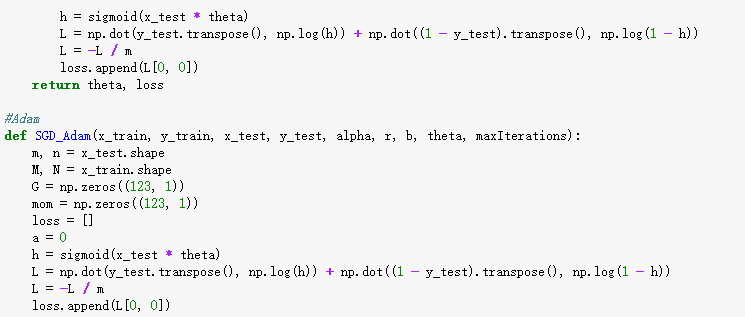
## 代码内容:

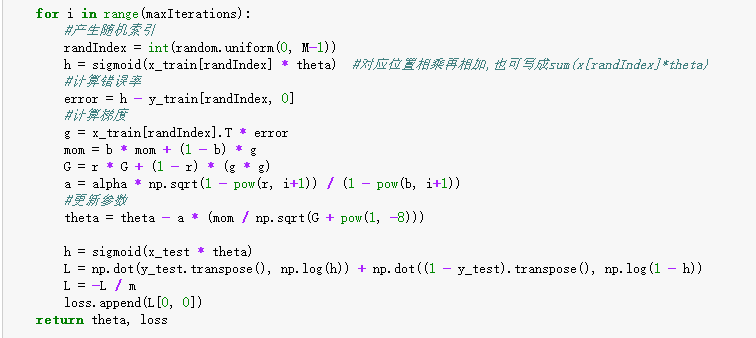
**逻辑回归与随机梯度下降**



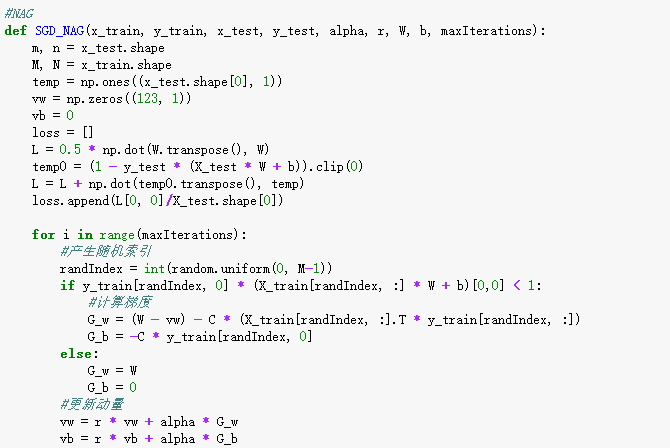


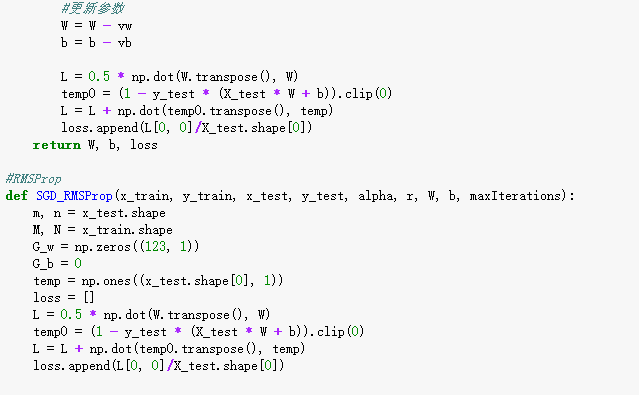


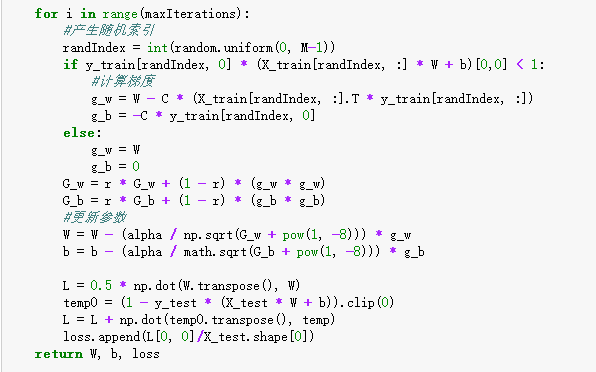


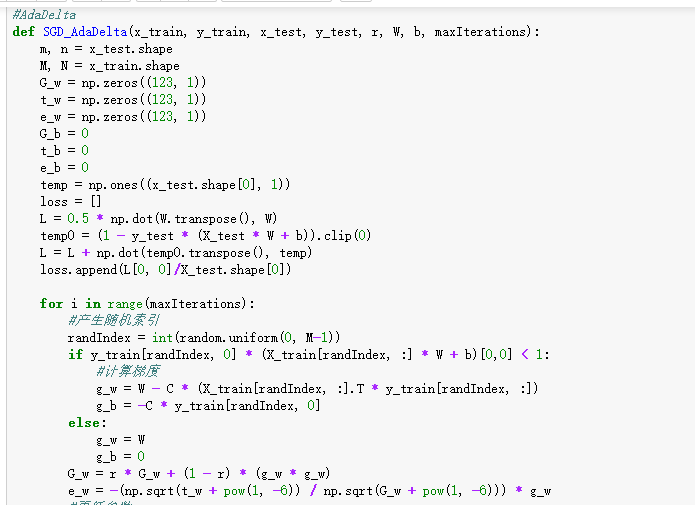


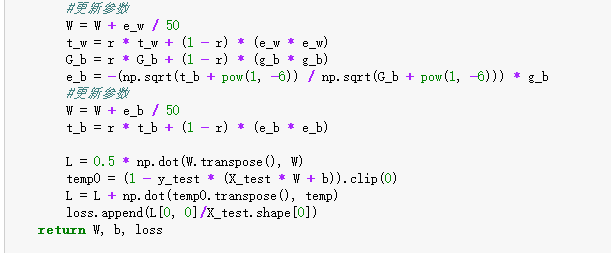
**线性分类与随机梯度下降**

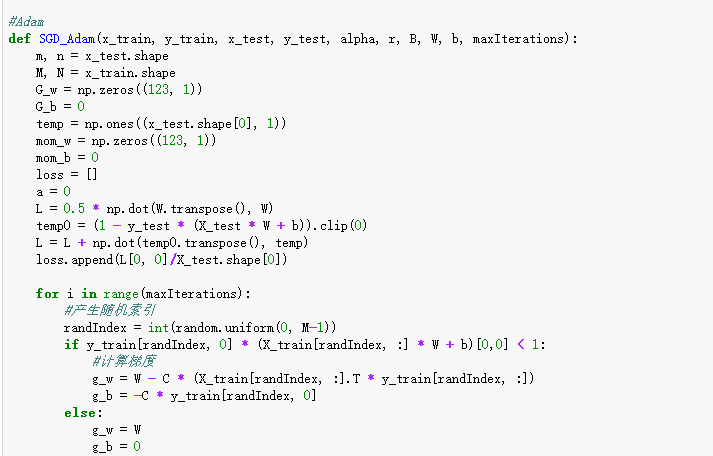


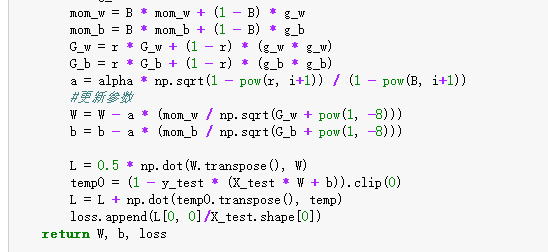












（针对逻辑回归和线性分类分别填写8-11内容）

## 模型参数的初始化方法:

**逻辑回归与随机梯度下降**

逻辑回归采取的初始化方法为随机初始化，参数为0-1之间的浮点数。

**线性分类与随机梯度下降**

线性分类采取的初始化方法为：w采用随机初始化，参数为0-1之间的浮点数。b初始化为0， C初始化为0.1。

## 选择的loss函数及其导数:

**逻辑回归与随机梯度下降**

loss = -1/m∑(y \* log(h(x)) + (1 - y) \* log(1 - h(x)))

G = (h(x) - y) \* x

**线性分类与随机梯度下降**

loss函数为：L=1/n(1/2||w||^2 + C∑max(0, 1 - y(wx+b)))，其中C为大于0的常数,n为样本数目。

对w偏导：G\_w = C∑(w - yx)，当1 - y(wx + b)>0时。

对b偏导：G\_b = w - C∑y

## 10.实验结果和曲线图:（各种梯度下降方式分别填写此项）

## 超参数选择

## 逻辑回归与随机梯度下降：

NAG：超参数有：alpha,r。 alpha为学习率，r为指数衰减。

RMSProp:超参数有：alpha,r。 alpha为学习率，r为指数衰减。

AdaDelta:超参数有：r。 r为指数衰减。

Adam:超参数有：alpha,r，b。alpha为学习率，r为指数衰减。

**线性分类与随机梯度下降**

NAG：超参数有：alpha,r。 alpha为学习率，r为指数衰减。

RMSProp:超参数有：alpha,r。 alpha为学习率，r为指数衰减。

AdaDelta:超参数有：r。 r为指数衰减。

Adam:超参数有：alpha,r，b。alpha为学习率，r为指数衰减。

## 预测结果（最佳结果）：

## 逻辑回归与随机梯度下降：

:0.37037 :0.39916

:0.36433 :0.42358

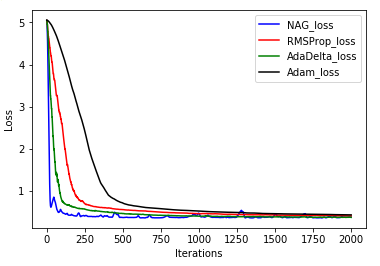
**线性分类与随机梯度下降**

## :0.47745 :0.47754

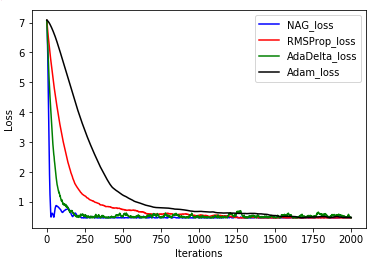
## : 0.48235 :0.48951

## loss曲线图：

## 逻辑回归与随机梯度下降：



**线性分类与随机梯度下降**

****

## 实验结果分析:

## 逻辑回归与随机梯度下降：

从实验结果可知使用NAG优化方法的收敛速度最快，loss值最终都趋向于0.5左右，使用NAG和RMSProp优化方法的结果比其他好。

**线性分类与随机梯度下降**

从实验结果可知使用NAG优化方法的收敛速度最快，loss值最终都趋向于0.4左右，使用AdaDelta优化方法的结果比其他好。

## 12.对比逻辑回归和线性分类的异同点：

两种方法都是常见的分类算法，从目标函数来看，区别在于逻辑回归采用的是logistical loss，SVM采用的是hinge loss。这两个损失函数的目的都是增加对分类影响较大的数据点的权重，减少与分类关系较小的数据点的权重。SVM的处理方法是只考虑support vectors，也就是和分类最相关的少数点，去学习分类器.而逻辑回归通过非线性映射，大大减小了离分类平面较远的点的权重，相对提升了与分类最相关的数据点的权重。

## 实验总结：

通过这次实验，我更清晰理解了梯度下降和随机梯度下降的区别与联系。

让我理解了逻辑回归和线性分类的区别与联系。

让我进一步理解SVM的原理并在较大数据上实践

让我懂得了使用不同优化方法去更新模型参数