



华南理工大学

South China University of Technology

## 《机器学习》课程实验报告

学 院 软件学院

专 业 软件工程

组 员 陈秋伶

学 号 201530611210

邮 箱 1900363950@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 15 日

3. 报告人: 陈秋伶

## 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 的中的 a9a 数据, 包含 32561 / 16281(testing) 个样本, 每个样本有 123/123 (testing) 个属性。

## 6. 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, 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  $L_{NAG}$ ,  $L_{RMSProp}$ ,  $L_{AdaDelta}$  和  $L_{Adam}$ 。

7. 重复步骤 4-6 若干次, 画出  $L_{NAG}$ ,  $L_{RMSProp}$ ,  $L_{AdaDelta}$  和  $L_{Adam}$  随迭代次数的变化图。

---

## 7. 代码内容:

(针对逻辑回归和线性分类分别填写 8-11 内容)

## 8. 模型参数的初始化方法: 全零初始化

## 9. 选择的 loss 函数及其导数:

### 逻辑回归:

```
def loss(X,Y,W):
    n,m=X.shape
    l=0
    for i in range(n):
        l+=math.log((1+exp(-Y[i]*W.T.dot(X[i]))),math.e)
    return l/n
def gradient(X,Y,W):
    n,m=X.shape
    g=0
    for i in range(n):
        g-=Y[i]*X[i]/(1+exp(Y[i]*W.T.dot(X[i])))
    return (g/n).reshape((g.shape[0],1))
```

### 线性分类:

```
def loss(X,Y,W):
    l=0
    C=10
    n,m=np.shape(X)
    for i in range(n):
        l+=max(0,1-Y[i]*W.T.dot(X[i]))
    l/=n
    l*=C
    l+=(W.T.dot(W)/2)[0][0]
    return l
def gradient(x,y,w):
    c=10
    g=np.zeros(np.shape(w))
    num,_=x.shape
    for i in range(num):
        if 1-y[i]*(w.T.dot(x[i]))>=0:
```

---

```
g=w-c*x.T.dot(y)
else:
    g=w
return g
```

## 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

NAG:

eta=0.1

gamma=0.1

batch=100

RMSProp:

eta=0.01

gamma=0.9

epsilon=1e-3

batch=100

AdaDelta

gamma=0.9

epsilon=1e-6

batch=100

Adam:

beta=0.9

gamma=0.999

eta=0.001

epsilon=1e-6

batch=100

线性分类:

NAG:

eta=1e-6

gamma=1e-5

batch=100

v=0

RMSProp:

eta=1e-4

gamma=0.9

epsilon=1e-6

---

batch=100  
AdaDelta  
gamma=0.95  
epsilon=1e-9

Adam  
beta=0.9  
gamma=0.9  
eta=1e-5  
epsilon=1e-9

预测结果（最佳结果）：

逻辑回归：

NAG: 0.32649897391142424

RMSProp: 0.32493229526506023

AdaDelta: 0.32409409130149014

Adam: 0.3260661358741839

线性分类：

NAG: 7.06317033

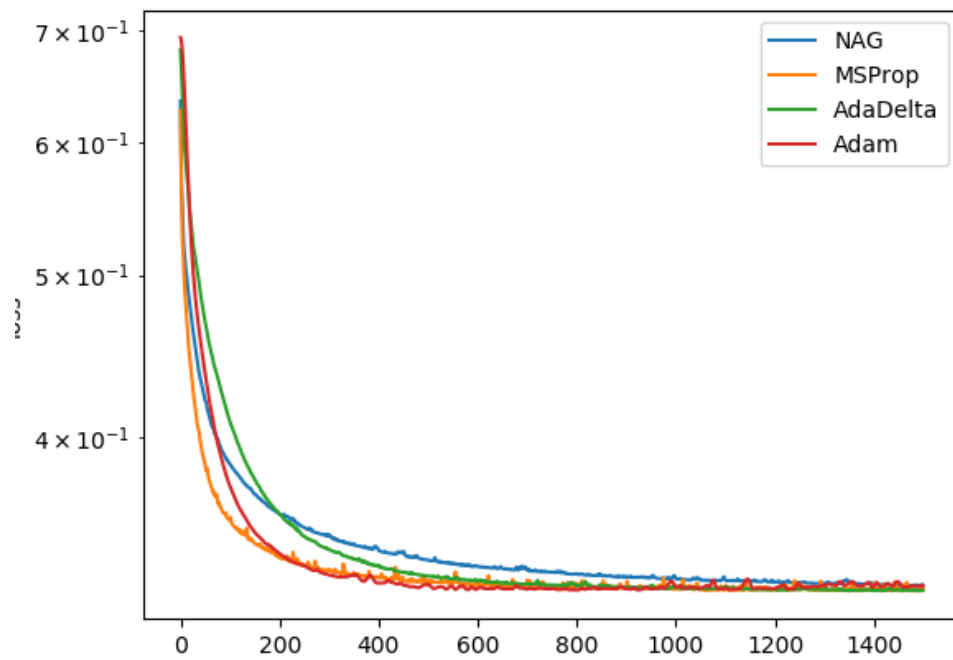
RMSProp: 5.71745986

AdaDelta: 5.87022375

Adam: 5.78553792

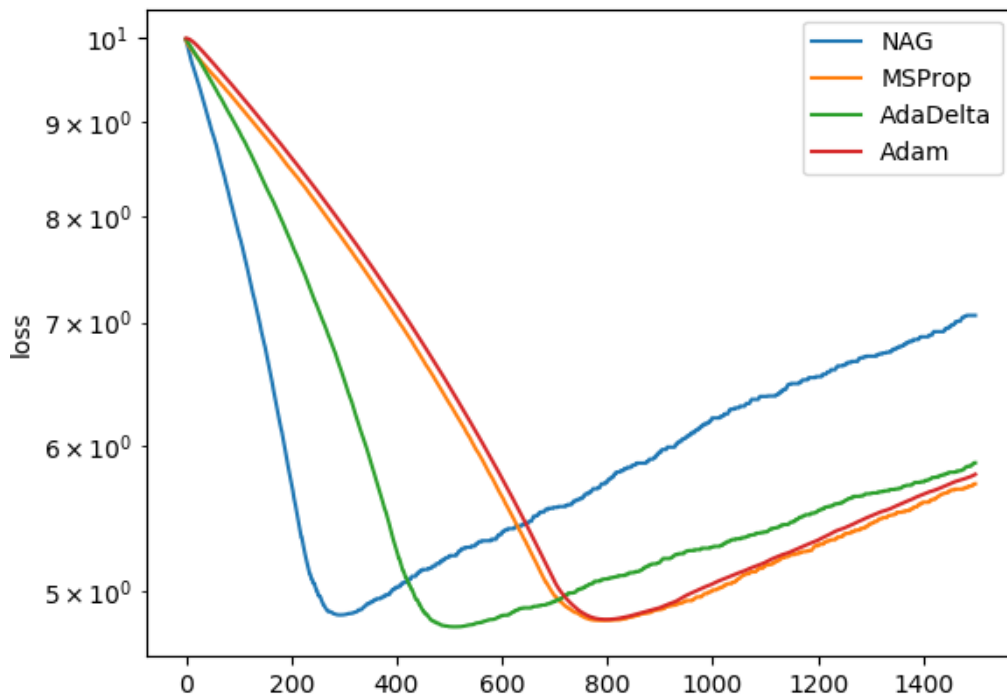
loss 曲线图：

逻辑回归：



x=230.894 y=0.378061

线性分类：



**11.实验结果分析:逻辑回归中，经过合适调参，四种优化方法的相差不大，最终会趋于一致。在线性分类中，由于调参原因无法收敛，但按照实际情况 Adam 的优化效果会更好。**

## **12.对比逻辑回归和线性分类的异同点：**

相同点：

- 1) 逻辑回归和线性分类都是分类算法。
- 2) 逻辑回归和线性分类都是监督学习算法
- 3) 逻辑回归和线性分类都是判别模型

不同点：

- 1) loss function 不同
- 2) 支持向量机只考虑局部的边界线附近的点，而逻辑回归考虑全局（远离的点对边界线的确定也起作用

---

### 13.实验总结：

通过这次实验，对逻辑回归和线性分类的实现原理有了更深的了解，越发觉得机器学习是一门深奥而有趣的学科，各种基础算法模型实现若参照前任的经验实现并不是很困难，重要的是进行调参的过程。