**Training Classifiers via SGD**

陈晓宇 MF16333007 [traumcxy@163.com](mailto:traumcxy@163.com)

1. 实现细节

本实验使用逻辑回归来实现二类分类，即利用来判断数据属于哪一个类别，其中，一旦，则数据属于第一类，反之则属于第二类。本次使用利用SGD算法分别求出log-likelihood和ridge-regression中的最小值时的theta值，来带入概率计算来算得数据的归属。

***首先利用log-likelihood来实现二类分类：***

其中，而

载入数据集data\_matrix和数据集的原来标签label\_matrix（m\*n）

输入默认参数lamda = 1，步长alpha = 0.1, 

对于每一次迭代：

随机选取，利用的偏导数去代替的偏导数，其中，由于norm的不连续性，只能利用来代替norm的微分。

然后更新

最后返回

利用得出的，配合概率计算公式，就可以得到每一个数据的就算出来的标签值了，与原有标签进行对比，就可以得到准确率了。

***下面是利用ridge-regression来实现二类分类***：

其中，而

载入数据集data\_matrix和数据集的原来标签label\_matrix（m\*n）

输入默认参数lamda = 1，步长alpha = 0.1, 

对于每一次迭代：

随机选取，利用的偏导数去代替的偏导数，其中

然后更新

最后返回

利用得出的，配合概率计算公式，就可以得到每一个数据的就算出来的标签值了，与原有标签进行对比，就可以得到准确率了。

2． 结果

见附录的result.xlsx文件

（很相近）

（很相近）