梯度下降实验

1. Cocoa和Mini-batch SGD

CocoaLocalIteration = 100,00

Mini-batch size = 100

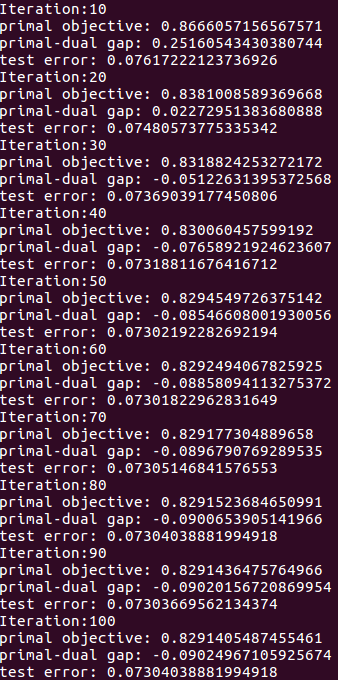
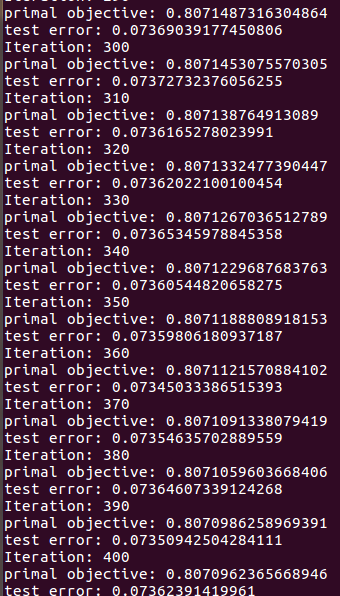
Data set: RCV

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Dataset | Trainging  n | Features  d | 稀疏性 | 精度 | Worker |
| Rcv1 | 677,399 | 47,236 | 0.16% | 1e-6 | 8 |
|  |  |  |  |  |  |

Time

|  |  |  |
| --- | --- | --- |
| Dataset | Iterations(#) | Time(s) |
| cocoa | 100 | 32 |
| Mini-batch SGD | 400 | 81 |

Cocoa Mini-batch SGD



1. SVRG 和 Spark Mini-batch SGD 对比

SVRG 的实现思路如下:

1.计算平均梯度(与Spark mini-batch SGD)一样

2.每个节点计算local w ,

3.全局更新w,我这里只是使用了加权平均

实验发现:在运行时间差不多的情况下,SVRG可以达到更好的LR 的 ROC,