2019-04-11

Nankai-Baidu
Joint Laboratory

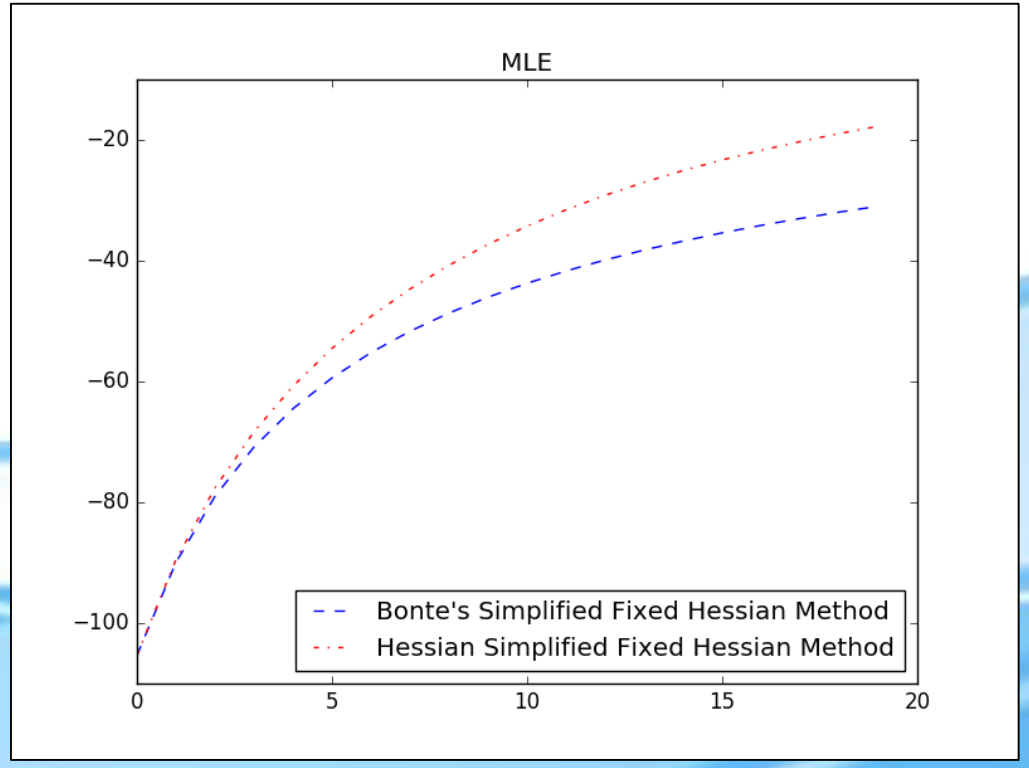
Parallel and Distributed
Software Technology Lab

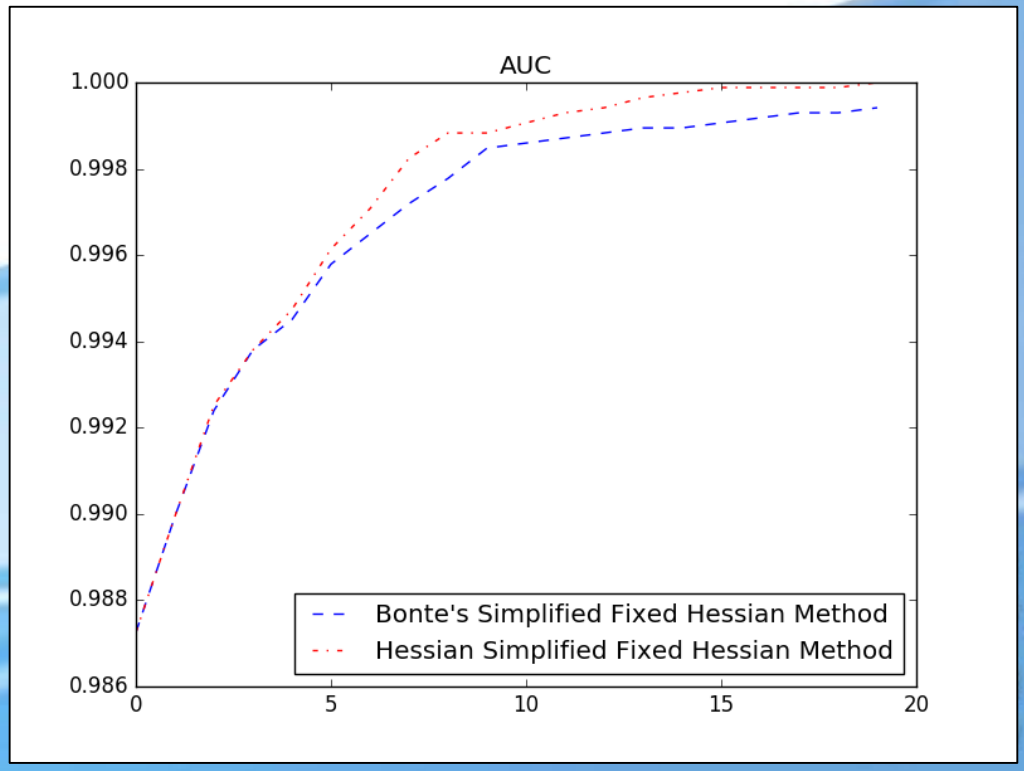
MNIST

- 使用手写数字识别数据集MNIST的前1000个样本数据
  - 数字3（当作正类）和数字8（当作负类）构建二分类问题
  - 使用的是直接从海森矩阵计算出来的替代矩阵（对角矩阵）
  - 发现鲁汶大学提出的SFH方法中的一个问题：B_{ii}可能为0.
  - 不进行归一化预处理训练数据，可能导致Sigmoid函数中指数函数上溢
- 对鲁汶大学的方法中改进学习率的想法：设置学习率
- IDASH 2019
- 实验设计问题

MNIST

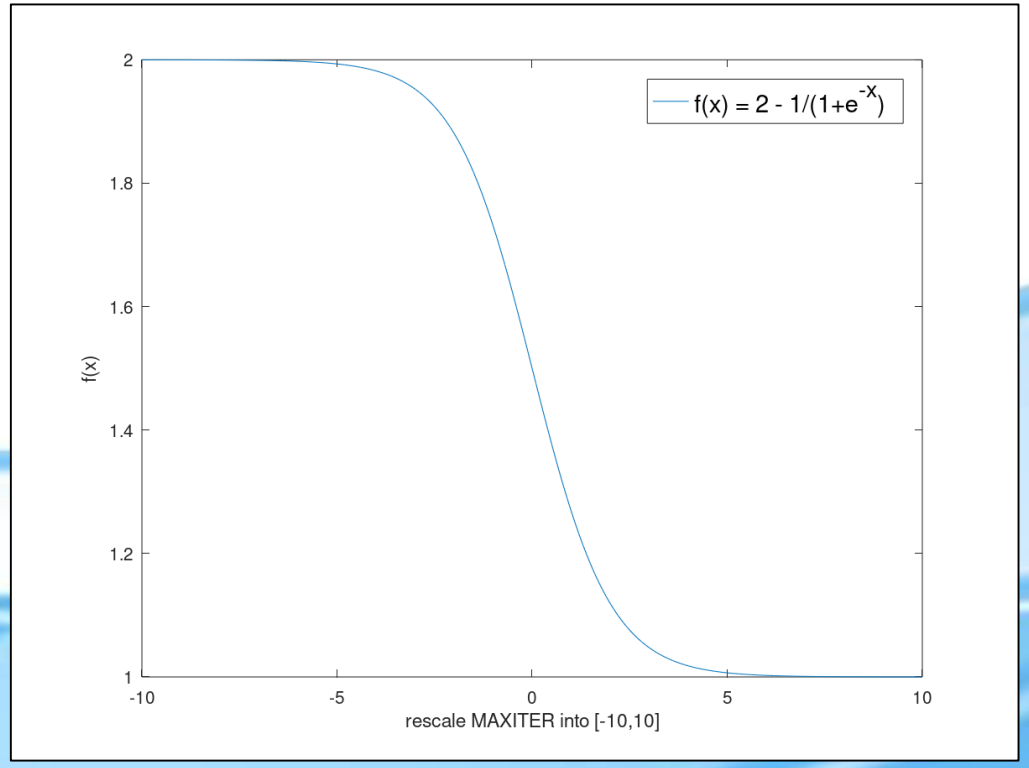
- 使用手写数字识别数据集MNIST的前1000个样本数据
  - 没有设置额外的学习率，修改了鲁汶大学B_{ii} =0的问题
  - 使用的是直接从海森矩阵计算出来的替代矩阵（对角矩阵）时间相对慢

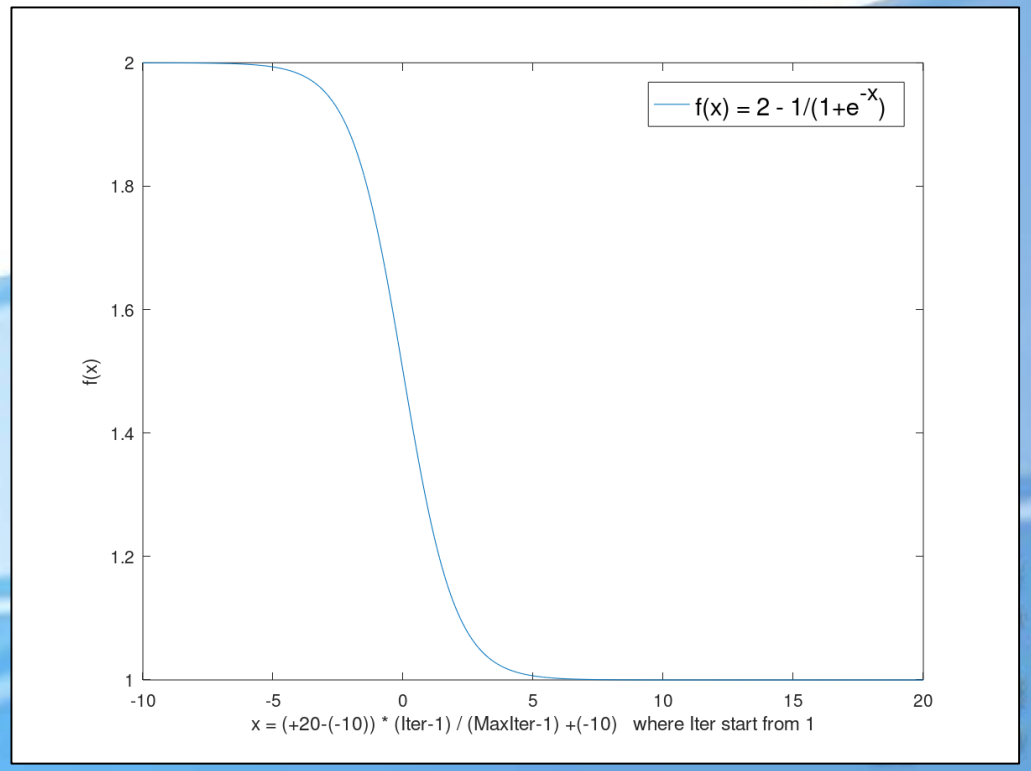




MNIST

- 对鲁汶大学的方法中改进学习率的想法：设置学习率
  - 希望前半部分接近2.0，后期接近1.0 （保证算法收敛）
  - 将区间 [1, MAX_ITER] 缩放到 [-10, +10] 或 [-10, +20]





MNIST

- IDASH 2019
- 2019 IDASH时间表:

  注册时间: 2019-03-25 - 2019-04-15

  提交结果: 2019-04-15 - 2019-08-16

  评选冠军: 2019-08-17 - 2019-09-30

  宣布冠军: 2019-10-01

  研讨会: 2019-10-26 （一天）

  论文提交: 2019-11-30 （截止日期）

MNIST

- IDASH 2019
- 四道题目： 具体的题目内容目前还没有发布
  - Track 1: 基于区块链和智能合约的分布式基因 - 药物相互作用数据共享
    (Distributed Gene-Drug Interaction Data Sharing based on Blockchain and Smart Contracts)
  - Track 2: 使用同态加密的安全基因型插补
    (Secure Genotype Imputation using Homomorphic Encryption)
  - Track 3: 通过TEE提供的服务外包隐私保护下的机器学习任务
    (Outsourcing Privacy-preserving Machine Learning as a Service through TEE)
  - Track 4: 通过安全多方计算建立决策树模型的安全协作
    (Secure Collaboration on Building a decision tree model using Secure Multiparty)

实验设计问题 LR on HE

- HE library : HEAAN
- Data Set : IDASH2017(103x1579), MNIST(3和8)
- Sigmoid (Polynomial) : SNU + Annealing

  Q : 公式化多项式拟合更大区间
- Experiment : Simplified Fixed Hessian Newton Method;

  The Presented Method 1，2

  Nesterov Accelerated Gradient (SNU)
- Criterion : ROC, AUC
- 需要分析：乘法电路深度>>参数的设置