机器学习中的优化问题: 从半监督学习到深度神经网络模型

刘为

摘要

机器学习是研究怎样使用计算机模拟或实现人类学习活动的科学, 是现如今 最前沿的研究领域之一. 本文从最优化角度出发, 系统地研究机器学习中的半监 督学习和深度神经网络模型. 具体地, 我们设计了三类最优化算法, 并验证了这些 算法的有效性和稳定性.

本文首先系统地研究了半监督学习, 并提出了一类新的半监督谱聚类的连续 最优化模型. 新模型不需要类的数目的一个精确值, 取而代之一个上界 d 即可; 新 模型能同时处理常见的四种半监督标签标注方式; 新模型以结果图的邻接矩阵作 为变量, 从而使我们直接得到分类结果. 在此通过对模型结构的分析, 我们提出一 种块坐标下降算法 (CO-BCD), 并证明了 CO-BCD 的全局收玫性. 数据实验显示 我们的方法对上界 d 的取值不敏感, 且相较于其他半监督聚类方法而言能快速有 效地取得更好的聚类效果.

接着, 我们系统地研究了深度神经网络中的最优化方法. 我们从神经网络中 的自编码和稀疏神经网络出发, 介绍如何将深度神经网络模型运用到半监督学习 中. 接着, 我们综述了求解神经网络的随机梯度下降算法以及基于罚方法的算法. 通过综述, 我们发现已有的求解深度神经网络的数学算法缺乏收玫性. 因此, 我们 提出了一类新的 l_1 罚模型. 我们将在后两章里利用该模型设计一些求解神经网络 的收敛算法.

之后, 我们将半 l_1 罚模型应用到自编码中, 并提出新模型 (RP). 同时, 我们构 建了 (RP) 与自编码模型之间在全局极小值点、局部极小值点、各种稳定点之间 的等价性与精确罚性. 进一步, 我们构建了一个有界闭集, 并证明新模型 (RP)在该 有界闭集内存在全局极小值点. 我们设计并提出了一种光滑化临近点算法 (SPG), 并用于在上述有界闭集内求解模型 (RP). 在收玫性分析中, 我们证明 SPG 算法能 稳定找到 (RP) 与自编码模型的一个一阶稳定点. 其次, 在合适的参数选取下,SPG 相对于 SGD 类算法 (例如 Adam, Adadelda, 和 vanilla SGD) 在人工合成数据集和 实际数据集上, 都能得到更低的训练误差和测试误差.

最后, 我们针对带块稀疏正则和 leaky ReLU 激活函数的深度神经网络, 提出 了一个非光滑非凸最优化模型. 我们证明了该最优化模型有一个非空的有界解集, 且其约束函数满足 Mangasarian-Fromovitz 约束规范性条件. 此外, 我们给出了 它的 KKT 点和稳定点的重要性质, 以及其与一个具有互补约束优化问题的稳定 点之间的联系. 接下来, 我们提出了一种不精确的增广拉格朗日算法来求解该最 小化问题, 并证明该算法能稳定找到该问题的一个有界的 KKT 点. 在数值试验 中, 我们主要测试了一类函数拟合问题以及一类半监督分类问题. 数值实验表明, 与已有算法相比, 该算法效率更高, 稳定性更强; 同时, 该算法可被直接用于训练 带 l_{2} 正则的 (leaky) ReLU 网络与自编码问题.

关键词：机器学习, 半监督学习, 自编码, 深度神经网络, 变分分析, 光滑化方法