

一种有效求解非凸正则化线性支持向量机的并行与分布式方法

关磊¹, 孙涛¹, 乔林波¹, 杨智慧^{2,3}, 李东升¹, 葛可适¹, 卢锡城¹

¹国防科技大学计算机学院, 中国长沙市, 410073

²复旦大学计算机科学技术学院, 中国上海市, 201203

³上海市数据科学重点实验室, 中国上海市, 201203

摘要: 支持向量机 (SVM) 被视为线性分类的有力工具。与可以产生稀疏效果的非凸惩罚项组合时, SVM能同时执行分类和变量选择。然而, 由于其不可微、非凸和非平滑特性, 非凸正则化SVM通常难以有效求得全局最优解。已有针对非凸正则化SVM的求解方案通常以串行方式求解, 因而无法充分利用现代多核机器的并行处理能力。另一方面, 现实世界中数据多以分布式方式存储, 迫切需要一种并行与分布式方法求解非凸正则化SVM问题。为应对这一挑战, 本文提出一种基于交替方向乘子法 (ADMM) 的并行算法高效求解非凸正则化SVM问题。采用有效技术降低并行算法的计算与同步开销。时间复杂度分析证明所提并行算法具有低复杂度。此外, 该并行算法能保证收敛性。在LIBSVM数据集上的实验证明了所提并行算法的有效性。

关键词: 线性分类; 支持向量机; 非凸惩罚项; 交替方向乘子法 (ADMM); 并行算法

<https://doi.org/10.1631/FITEE.1800566>