正交约束优化：理论、算法与应用

高 斌

摘 要

正交约束优化问题是指变量为矩阵且满足正交性约束的最优化问题, 其可行集也被称为 Stiefel 流形. 这类优化模型在科学计算、材料科学及数据科学等领域有着广泛的应用. 由于正交化过程的高计算代价和低可扩展性, 正交约束优化算法的研究遇到了新的瓶颈和挑战. 本文系统地研究了正交约束优化问题, 设计了三类优化算法, 并将新算法应用于实际问题中.

首先我们系统地研究了正交约束优化问题. 我们从黎曼流形优化和欧式空间约束优化两个角度, 分别推导了问题的最优性条件. 其中, 通过分析一大类黎曼度量下的黎曼梯度, 我们得到了问题各种描述下一阶最优性条件的对应关系. 此外, 我们还证明了在任意一阶稳定点处, 正交约束优化问题的 Lagrange 乘子具有显式表达式. 这些性质的刻画将对本文的算法设计起到至关重要的作用.

接着我们针对一大类正交约束优化问题, 提出了乘子校正算法框架. 其主要包含两个步骤, 分别是函数值下降步和乘子校正步. 不同于以往的经典算法, 在我们的算法框架中, 函数值下降步采用标准的欧式负梯度方向, 而不是 Stiefel 流形的切空间方向. 另一方面, 我们构造的乘子校正步进一步使函数值下降, 同时也保证了乘子的对称性. 基于此算法框架, 我们提出了两大类算法. 第一类是梯度下降方法, 其中包括梯度反射法和梯度投影法. 第二类采用以列为块的块坐标下降方法. 进一步, 我们证明了算法的全局收敛性. 数值实验表明我们的算法框架具有很大的潜力.

然后我们将乘子校正步推广到一般的 Stiefel 流形收缩类算法, 得到了子空间加速的收缩类算法. 通过利用乘子校正步的子空间最优性质, 我们设计了一类两阶段的子空间加速算法. 第一阶段是函数值下降法, 第二阶段是子空间加速步. 将 Stiefel 流形的收缩类线搜索算法应用于第一阶段, 我们证明了加速算法的全局收敛性和局部线性收敛速度. 数值实验展示了加速技术的有效性.

之后针对一般的正交约束优化问题, 我们提出了基于增广 Lagrange 函数的并行算法. 由于正交化过程的可扩展性较低, 我们考虑不可行方法并采用增广 Lagrange 罚函数. 不同于经典的增广 Lagrange 函数法, 在我们的算法中, 原始变量的更新通过极小化增广 Lagrange 函数的邻近点线性化逼近得到. 同时, Lagrange 乘子由其一阶稳定点处的显式表达式更新得到. 由此, 算法的主要步骤都可以很自然地进行矩阵计算层面的并行化. 进一步, 我们建立了算法的全局收敛性, 分析了算法的最坏情况复杂度以及局部收敛速度. 此外, 为了减弱算法对罚参数的敏感性, 我们提出了改进的可并行列极小化算法. 串行的数值实验说明乘子的新更新方式显著加速了算法的收敛速度, 并且数值表现与已有的可行

方法不相上下. 并行环境下的数值实验验证了我们的算法具有较高的可扩展性.

最后我们将乘子校正算法和基于增广 Lagrange 函数的并行算法应用于电子结构计算. 我们考虑了电子结构计算中的 Kohn-Sham 密度泛函理论 , 其离散模型通常表述为一个带有正交约束的优化问题. 在串行环境下, 我们测试了 18 个不同的分子结构, 数值实验显示了我们的新算法优于已有的经典算法. 在并行环境下, 我们测试了简化的Kohn-Sham总能量极小化问题, 数值结果显示我们提出的并行算法具有较高的可扩展性.

关键词：正交约束, Stiefel 流形, 块坐标下降法, 收缩类方法, 增广 Lagrange 函数, 并行计算, 电子结构计算