

Study on the Solution of a Hybrid Bundle Method Subproblem

Hanyang Li

School of Mathematics, Liaoning Normal University, Dalian Liaoning
Email: 1160629201@qq.com

Received: May 8th, 2019; accepted: May 22nd, 2019; published: May 29th, 2019

Abstract

The solutions of non-smooth optimization problem have been the focus of the optimization theory for a long time. This paper proposes a class of hybrid bundle method subproblem with matrix norm, and then researches this subproblem's Lagrange function by taking advantage of the dual space theorem. Finally, based on relation between the dual problem and the original problem, the solution of hybrid bundle method subproblem is calculated, and the concrete expression is given.

Keywords

Non-Smooth Optimization, Bundle Method, Lagrange Dual Space

一类混合束方法子问题的求解研究

李函阳

辽宁师范大学数学学院, 辽宁 大连
Email: 1160629201@qq.com

收稿日期: 2019年5月8日; 录用日期: 2019年5月22日; 发布日期: 2019年5月29日

摘 要

非光滑优化问题的求解一直以来都是优化理论的重点研究对象, 本文提出了一类带有矩阵范数的混合束方法子问题, 然后利用对偶空间理论研究该问题的Lagrange函数, 最后通过对偶问题与原问题的关系, 计算出该混合束方法子问题的解, 并给出具体的表达式。

关键词

非光滑优化, 束方法, Lagrange对偶空间



1. 引言

对非光滑优化问题来说，束方法是最有效的解决方法之一。很多学者在束方法的理论和应用方面都取得了较多的有益成果，见文献[1] [2] [3] [4] [5]。考虑原始问题 $\min_{x \in R^n} f(x)$ ，其中 f 是 $R^n \rightarrow R$ 的非光滑凸函数。文献[6]中作者提出了一个带矩阵范数的双稳定子问题，文献[7]中作者对目标函数增加迫近项和对可行域增加信赖域约束，灵活地把邻近束方法和信赖域束方法结合起来构造了混合束方法。通过[6]和[7]的启发，本文提出了一类带矩阵范数的混合束方法子问题：

$$\begin{cases} \min \widehat{f}_k(x) + \frac{1}{2\tau_k} \|x - x^k\|_k^2 \\ \text{s.t. } \|x - x^k\|_k^2 \leq \delta_k^2 \end{cases}$$

这里引用了矩阵范数，并通过计算得到该子问题解的表达式。

2. 预备知识

束方法基本原理是在迭代进行到第 k 步时，产生了一组候选点列 $\{y^i\}_{i=1}^k$ 和一组稳定中心点列 $\{x^i\}_{i=1}^k$ ，它利用候选点列的信息来构造目标函数 $f(x)$ 的切平面近似 $\widehat{f}_k(x)$ ，然后求得 d^k ，令 $y^{k+1} = x^k + d^k$ 作为下一个候选点，然后进行下降性检验，以此来确定下一个稳定中心，同时更新束，如此迭代得一点列 $\{x^k\}$ ，满足 $\{x^k\}$ 的任一聚点都是原问题的最优解。

束方法具体来说常见的有信赖域束方法、邻近束方法和水平束方法，其中信赖域束方法和邻近束方法的迭代格式分别为：

$$\begin{aligned} y^{k+1} &= \arg \min \left\{ \widehat{f}_k(x) \mid \|x - x^k\| \leq \delta_k, x \in R^n \right\} \\ y^{k+1} &= \arg \min \left\{ \widehat{f}_k(x) + \frac{1}{2\tau_k} \|x - x^k\|^2 \mid x \in R^n \right\} \end{aligned}$$

其中 $\delta_k > 0$ 为信赖域参数， $\tau_k > 0$ 为邻近参数。

3. 混合束方法子问题的产生

受文献[6]和[7]的启发，本文提出了一类带矩阵范数的混合束方法子问题：

$$\begin{cases} \min \widehat{f}_k(x) + \frac{1}{2\tau_k} \|x - x^k\|_k^2 \\ \text{s.t. } \|x - x^k\|_k^2 \leq \delta_k^2 \end{cases} \quad (2.1)$$

在这里令 $\widehat{f}_k(x) = \max_{j \in J^k} \{f(y^j) + \langle g^j, x - y^j \rangle\} = \max_{j \in J^k} \{f(x^k) + \langle g^j, x - x^k \rangle - \alpha_j^k\}$ ，其中， f 在 x^k 点的线性化误差为 $\alpha_j^k = f(x^k) - f(y^j) + \langle g^j, x - y^j \rangle$ ，由此知 $g^j \in \partial_{\alpha_j} f(y^j)$ ， J^k 为指标集， x^k 为当前稳定中心，当 $k \rightarrow \infty$ 时， $\delta_k \rightarrow 0$ ，记问题(2.1)的最优解为 x^{k+1} 。我们定义原空间范数 $\|\cdot\|_k^2 = \langle M_k \cdot, \cdot \rangle$ ，其对偶空间范数为 $\|\cdot\|_k^2 = \langle \cdot, M_k^{-1} \cdot \rangle$ ， M_k 是对称正定矩阵。

当迭代次数的增加，束中的元素原来越多，对之后的计算和储存会带来巨大的麻烦，所以采用现已

有的聚合技术, 将束进行压缩至只保留 np_k 个元素, np_k 甚至有可能小于 k , 这时 $j \in \{1, 2, \dots, np_k\}$, 接着我们对其对偶问题展开研究。

4. 混合束方法子问题解的表达式

由于问题(2.1)是非光滑凸优化问题, 对非光滑凸优化问题来说在满足 Slater 约束规格的条件下, 对问题的求解可以转化为其极小化 Lagrange 函数问题的求解, 也就是说必然存在一个乘子 $0 \leq \lambda_k \in R$, 能使 $x^{k+1} = \arg \min_{x \in R^n} L(x, \lambda_k)$, 在这里 $L(x, \lambda_k) := \widehat{f}_k(x) + \frac{1}{2\tau_k} |x - x^k|_k^2 + \frac{1}{2} \lambda_k (|x - x^k|_k^2 - \delta_k^2)$, 故问题(2.1)等价于下述问题:

$$\begin{cases} \min \widehat{f}_k(x) + \frac{1}{2\tau_k} |x - x^k|_k^2 + \frac{1}{2} \lambda_k (|x - x^k|_k^2 - \delta_k^2) \\ x \in R^n \end{cases} \quad (3.1)$$

问题(3.1)还可以等价的写成:

$$\begin{cases} \min \widehat{f}_k(x) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} \\ x \in R^n \end{cases} \quad (3.2)$$

我们定义相应的额定下降:

$$\xi_{k+1} := f(x^k) - \left(\widehat{f}_k(x^{k+1}) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x^{k+1} - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} \right)$$

定理 3.1: 令问题(2.1)的最优解为 x^{k+1} , 假设参数 $\delta_k, \tau_k > 0$, 则有

$$x^{k+1} = x^k - \frac{\tau_k}{1 + \lambda_k \tau_k} M_k^{-1} \widehat{g}^k, \quad \text{其中 } \widehat{g}^k = \sum_{j=1}^{np_k} \beta_j g^j$$

并且 $\bar{\beta} = (\bar{\beta}_1, \bar{\beta}_2, \dots, \bar{\beta}_{np_k})$ 是下述问题的最优解

$$\begin{cases} \min \frac{\tau_k}{2(1 + \lambda_k \tau_k)} \left\| \sum_{j=1}^{np_k} \beta_j g^j \right\|_k^2 + \sum_{j=1}^{np_k} \beta_j \alpha_j^k \\ \text{s.t. } \beta \in \Delta_k := \left\{ z \in [0, 1]^{np_k} : \sum_{j=1}^{np_k} z_j = 1 \right\} \end{cases} \quad (3.3)$$

证明: 因为问题(3.2)和问题(2.1)是等价的, 所以可以通过引入变量 r 得到问题(3.2)的等价形式如下:

$$\begin{cases} \min_{(x, r) \in R^n \times R} r + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} \\ \text{s.t. } r \geq f(x^k) + \langle g^j, x - x^k \rangle - \alpha_j^k, j \in \{1, 2, \dots, np_k\} \end{cases} \quad (3.4)$$

当 $\beta \in R_+^{np_k}$ 时, 问题(3.4)的 Lagrange 函数为:

$$L(x, r, \beta) = r + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} + \sum_{j=1}^{np_k} \beta_j (f(x^k) + \langle g^j, x - x^k \rangle - \alpha_j^k - r)$$

整理得:

$$L(x, r, \beta) = \left(1 - \sum_{j=1}^{np_k} \beta_j \right) r + \sum_{j=1}^{np_k} \beta_j (f(x^k) + \langle g^j, x - x^k \rangle - \alpha_j^k) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2}$$

由于目标函数是强凸的, 所以问题(2.1)有且仅有一个解。记最优乘数为 $\bar{\beta}$, $(x^{k+1}, \bar{\beta})$ 可以从问题(3.4)和它的对偶问题中获得。

$$\min_{(x,r) \in R^n \times R} \max_{\beta \in R_+^{npk}} L(x,r,\beta) = \max_{\beta \in R_+^{npk}} \min_{(x,r) \in R^n \times R} L(x,r,\beta)$$

该问题与问题(3.4)的有限最优值是一样的, 不一样的是对偶问题的内层是无约束最小化优化问题, 所以必有 $1 - \sum_{j=1}^{npk} \beta_j = 0$ 。基于以上结论, 问题(3.2)的最优解 x^{k+1} 和最优乘数 $\bar{\beta}$ 分别可以从原问题和对偶问题中得到。

$$\min_{x \in R^n} \max_{\beta \in R_+^{npk}} L(x,\beta) = \max_{\beta \in R_+^{npk}} \min_{x \in R^n} L(x,\beta)$$

其中

$$\begin{aligned} L(x,\beta) &= \sum_{j=1}^{npk} \beta_j \left(f(x^k) + \langle g^j, x - x^k \rangle - \alpha_j^k \right) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} \\ &= f(x^k) + \sum_{j=1}^{npk} \left(\langle \beta_j g^j, x - x^k \rangle - \beta_j \alpha_j^k \right) + \frac{1 + \lambda_k \tau_k}{2\tau_k} (x - x^k)^T M_k (x - x^k) - \frac{\lambda_k \delta_k^2}{2} \end{aligned}$$

考虑对偶问题, 对于 $\forall \beta \in \Delta_k$, 应用 $x(\beta) = \arg \min_x L(x,\beta)$ 的最优性条件 $\nabla_x L(x(\beta), \beta) = 0$ 。即

$$\nabla_x L(x(\beta), \beta) = \sum_{j=1}^{npk} \beta_j g^j + \frac{1 + \lambda_k \tau_k}{\tau_k} M_k (x(\beta) - x^k) = 0 \quad (3.5)$$

记 $\hat{g}^k = \sum_{j=1}^{npk} \bar{\beta}_j g^j$, 当 $\beta = \bar{\beta}$, $x(\bar{\beta}) = x^{k+1}$ 时, 有 $x^{k+1} = x^k - \frac{\tau_k}{1 + \lambda_k \tau_k} M_k^{-1} \hat{g}^k$ 。

接下证明 $\bar{\beta}$ 是问题(3.3)的解。

首先, 在式子(3.5)的两边同时乘 $x(\beta) - x^k$ 得

$$\sum_{j=1}^{npk} \beta_j g^j (x(\beta) - x^k) + \frac{1 + \lambda_k \tau_k}{\tau_k} (x(\beta) - x^k)^T M_k (x(\beta) - x^k) = 0$$

即

$$\sum_{j=1}^{npk} \beta_j \langle g^j, x(\beta) - x^k \rangle + \frac{1 + \lambda_k \tau_k}{\tau_k} |x(\beta) - x^k|_k^2 = 0 \quad (3.6)$$

同理在式(3.5)两边同时乘 M_k^{-1} 得

$$M_k^{-1} \sum_{j=1}^{npk} \beta_j g^j + \frac{1 + \lambda_k \tau_k}{\tau_k} (x(\beta) - x^k) = 0。$$

再在两边同乘 $\frac{\tau_k}{1 + \lambda_k \tau_k} \sum_{j=1}^{npk} \beta_j g^j$ 得

$$\frac{\tau_k}{1 + \lambda_k \tau_k} \sum_{j=1}^{npk} \beta_j g^j M_k^{-1} \sum_{j=1}^{npk} \beta_j g^j + \sum_{j=1}^{npk} \beta_j \langle g^j, x(\beta) - x^k \rangle = 0。$$

即

$$\frac{\tau_k}{1 + \lambda_k \tau_k} \left\| \sum_{j=1}^{npk} \beta_j g^j \right\|_k^2 + \sum_{j=1}^{npk} \beta_j \langle g^j, x(\beta) - x^k \rangle = 0 \quad (3.7)$$

将式子(3.6)和式子(3.7)联立, 得

$$\frac{\tau_k}{1 + \lambda_k \tau_k} \left\| \sum_{j=1}^{np_k} \beta_j g^j \right\|_k^2 = \frac{1 + \lambda_k \tau_k}{\tau_k} |x(\beta) - x^k|_k^2$$

故有

$$\begin{aligned} L(x(\beta), \beta) &= f(x^k) + \sum_{j=1}^{np_k} (\langle \beta_j g^j, x - x^k \rangle - \beta_j \alpha_j^k) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} \\ &= f(x^k) + \frac{1 + \lambda_k \tau_k}{2\tau_k} |x - x^k|_k^2 - \frac{\lambda_k \delta_k^2}{2} + \left\langle \sum_{j=1}^{np_k} \beta_j g^j, x - x^k \right\rangle - \sum_{j=1}^{np_k} \beta_j \alpha_j^k \\ &= f(x^k) + \frac{\tau_k}{2(1 + \lambda_k \tau_k)} \left\| \sum_{j=1}^{np_k} \beta_j g^j \right\|_k^2 - \frac{\lambda_k \delta_k^2}{2} + \left\langle \hat{g}^k, -\frac{\tau_k}{1 + \lambda_k \tau_k} M_k^{-1} \hat{g}^k \right\rangle - \sum_{j=1}^{np_k} \beta_j \alpha_j^k \\ &= f(x^k) + \frac{\tau_k}{2(1 + \lambda_k \tau_k)} \|\hat{g}^k\|_k^2 - \frac{\lambda_k \delta_k^2}{2} - \frac{\tau_k}{1 + \lambda_k \tau_k} \|\hat{g}^k\|_k^2 - \sum_{j=1}^{np_k} \beta_j \alpha_j^k \\ &= f(x^k) - \left(\frac{\tau_k}{2(1 + \lambda_k \tau_k)} \|\hat{g}^k\|_k^2 + \frac{\lambda_k \delta_k^2}{2} + \sum_{j=1}^{np_k} \beta_j \alpha_j^k \right) \end{aligned}$$

综上所述 $\bar{\beta}$ 是问题(3.3)的解。

5. 结论

本文从 Lagrange 对偶空间理论出发, 研究了一类带有矩阵范数的混合束方法子问题, 并通过计算得到该子问题的解, 最后给出解的具体表达式, 为以后束方法问题研究工作的展开提供了新的理论基础。

参考文献

- [1] 高岩. 非光滑优化[M]. 北京: 科学出版社, 2008.
- [2] 王宜举, 修乃华. 非线性最优化理论与方法[M]. 北京: 科学出版社, 2012.
- [3] 沈洁, 曹天水, 李娜, 等. 关于复合逼近束方法对偶问题的研究[J]. 吉林师范大学学报(自然科学版), 2013, 34(4): 1-4.
- [4] 沈洁, 赵睿, 高亚丽. 水平束方法子问题的求解研究[J]. 吉林师范大学学报(自然科学版), 2017, 38(2): 54-57.
- [5] 沈洁, 李轩, 李娜. 关于双稳定束方法对偶问题的研究[J]. 吉林师范大学学报(自然科学版), 2014, 35(3): 64-67.
- [6] 沈洁, 李娜, 田佳茜. 双稳定束方法以及收敛性分析[J]. 沈阳师范大学学报(自然科学版), 2015, 33(2): 177-180.
- [7] 张清叶, 高岩. 求解非光滑凸规划的一种混合束方法[J]. 运筹学学报, 2016, 20(2): 113-120.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2163-1476, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: orf@hanspub.org