

求解非光滑复合约束优化问题的再分配束方法

吴琼, 张宏伟*, 王爽

(大连理工大学数学科学学院, 辽宁大连 116024)

摘要: 针对一类特殊的复合约束优化问题提出了再分配型束方法, 其中目标函数和约束函数为具有 lower- C^2 性质的函数. 利用改善函数, 原约束问题可以被转化为无约束问题, 并且新的目标函数也具有 lower- C^2 性质. 再利用 lower- C^2 函数的性质, 引入了凸化参数来改善子问题目标函数的凸性, 并设计了相应的束方法. 给出了原问题和新问题最优点的关系, 并简单地给出了参数稳定性结论和算法的局部收敛性结论. 通过对 H_2/H_∞ 综合问题的分析, 利用提出的算法计算了最优的 H_2/H_∞ 动态控制器, 表明了算法的有效性.

关键词: 约束优化; 非光滑优化; 改善函数; 束方法; lower- C^2

中图分类号: O221.2

文献标识码: A

doi:10.7511/dllgxb202001014

0 引言

考虑一类带有复合约束的优化问题(1):

$$\begin{aligned} \min f(\mathbf{x}) \\ \text{s. t. } g(\mathbf{x}) = (h \circ c)(\mathbf{x}) \leq 0 \end{aligned}$$

其中 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 和 $g: \mathbf{R}^n \rightarrow \mathbf{R}$ 在开集上是 lower- C^2 函数. 这类问题在很多领域都有重要的应用. 例如, 令 $c: \mathbf{R}^n \rightarrow \mathbf{S}^m$ 为二阶连续可微的 Hermitian 矩阵值函数, $h: \mathbf{S}^m \rightarrow \mathbf{R}$ 为矩阵的最大特征值函数, 则这类特殊问题在很多领域都有很重要的应用. 在鲁棒控制理论中, 一类问题是在最大特征值不超过某个指定的界的条件下设计一类系统. 在问题(1)中, 将最大特征值替换成最小特征值, 将不等号换个方向, 就能将问题应用到结构化控制中, 一类经典的结构化控制问题^[1]是在最小特征值的下界条件下最小化某个函数. 这类模型还适用于很多数学规划, 由于问题(1)中的约束是非凸且非光滑的, 这类问题实际上很难求解. 非凸非光滑无约束问题的求解方法在文献中并不是很多, 而非凸非光滑约束问题的方法则更少. 为了解决这类非凸非光滑的约束优化问题, 经典的手段是引入相应的 Lagrange 函数或惩罚函数, 将问题变成最

小化目标函数和约束函数某种形式的和的无约束问题, 或者使用支撑函数来近似约束使得问题变得易解^[2]. 但是, 在实际计算中, 约束函数的非凸性和非光滑性会使得惩罚函数和 Lagrange 函数变得更加复杂, 而支撑函数只适用于目标函数具有比较简单的形式如目标函数为线性时. 因此, 为了有效利用原问题的复合结构, 本文引用一类特殊的改善函数^[3] $F_x(\mathbf{y}) := \max\{f(\mathbf{y}) - f(\mathbf{x}), g(\mathbf{y})\}$, 则求解问题(1)等价于求解下面的数学规划问题(2):

$$\min_{\mathbf{y} \in \mathbf{R}^n} F_x(\mathbf{y})$$

其中, \mathbf{x} 是某次迭代的逼近中心且随着迭代变化. 这样, 改善函数 F_x 保留了原目标函数 f 和约束函数 g 的性质和原问题的复合结构. 本文针对问题(2)提出一类再分配束方法. 问题(2)也是一类非凸非光滑问题, 而束方法是现今解决非凸非光滑问题最有效的优化方法之一. 之前的很多方法在解决非凸问题时使用了邻近策略^[4-8]. 因为在求解非凸问题时, 线性化误差经常是非正的, 因此之前的算法采取了多种策略来改善线性化误差. 大多数算法都是从重新定义线性化误差出发的, 例

收稿日期: 2019-08-03; 修回日期: 2019-11-29.

基金项目: 山东省自然科学基金资助项目(ZR201807061177).

作者简介: 吴琼(1993-), 女, 博士生, E-mail: wuqionglap@mail.dlut.edu.cn; 张宏伟*(1955-), 男, 博士, 教授, E-mail: hwzhang@dlut.edu.cn.

如取线性化误差的绝对值^[9],取线性化误差的二次惩罚项^[10],或将前两种组合起来^[11-12]. 本文从另一种角度出发,利用目标函数和约束函数的内在性质,使用再分配的策略和加入凸化项来保证线性化误差为正,再应用束方法来求解.

1 假设和最优性条件

首先对问题(2),给出下面的基本假设.

假设 H1 假设存在 $x^0 \in R^n$ 和 $M_0 \geq 0$ 使得 $L_0 := \{x \in R^n : F_x(x) \leq M_0\}$ 为有界集.

这样可以得到下面的引理.

引理 1 令 F_x 为问题(1)的改善函数,则 F_x 在 L_0 上是 lower- C^2 的. 若 F_x 满足上面的假设 H1,对任意 $x \in L_0$,下列结论成立.

- (1)水平集 L_0 非空且紧致.
- (2)函数 F_x 下有界且邻近有界,其阈值为 0,即 $F_x(\cdot) + \frac{R}{2} |\cdot|^2$ 下有界对所有 $R > 0$ 都成立.
- (3)存在 ρ^c ,使得对任何 $\rho \geq \rho^c$ 和任意的 $y \in L_0$,函数 $F_x(\cdot) + \frac{\rho}{2} |\cdot - y|^2$ 在 L_0 上是凸的.

(4)函数 F_x 在 L_0 上是 Lipschitz 连续的.

命题的证明可参考文献[13]和[14]. 接下来,给出问题(1)和问题(2)局部最优解的关系.

定理 1 若 \bar{x} 是问题(1)的一个最优解,则 \bar{x} 也是问题(2)的一个局部最优解且 $0 \in \partial F_x(\bar{x})$,若 $0 \in \partial F_x(\bar{x})$,则有两种可能性:

- (1) $g(\bar{x}) > 0$,则 \bar{x} 是 g 自己的一个局部极小点,称 \bar{x} 为约束违反的局部极小点.
- (2) $g(\bar{x}) \leq 0$,则 \bar{x} 满足问题(1)的 Fritz John 必要性最优性条件. 若 $g(\bar{x}) < 0$ 或 $0 \notin \partial g(\bar{x})$,则 \bar{x} 满足问题(1)的 KKT 必要性最优性条件.

上面的引理和定理可参考文献[2,15-16].

2 束信息的存储

下面利用引理 1 来定义非凸情况下的束信息和束方法. 首先,定义邻近映射

$$p_R F_x(x) := \operatorname{argmin}_y \left\{ F_x(y) + \frac{1}{2} R |x - y|^2 \right\}$$

当 R 充分大时,上式在 L_0 上是单值且 Lipschitz

的. 引入增广函数

$$F_{\bar{x}^k}^{\eta_n}(\cdot) := F_{\bar{x}^k}(\cdot) + \eta_n |\cdot - \bar{x}^k|^2 / 2 \quad (1)$$

定义函数的线性化误差

$$e_i^k = F_{\bar{x}^k}(\hat{x}^k) - (F_{\bar{x}^k}(x^i) + \langle G_i^k, \hat{x}^k - x^i \rangle)$$

这样就得到了增广束信息

$$\bigcup_{i \in I_n} \{e_i^k, d_i^k, \Delta_i^k, G_i^k\}$$

其中 $d_i^k = \frac{|x^i - \bar{x}^k|^2}{2}$, $\Delta_i^k = x^i - \bar{x}^k$, $G_i^k \in \partial F_{\bar{x}^k}(x^i)$.

因为线性化误差 e_i^k , 偏差范数 d_i^k 和偏差向量 Δ_i^k 都依赖于 \hat{x}^k , 因此,在下降步 \hat{x}^k 改变之后这些变量都需要做对应的更新.

为了定义 QP 问题,把逼近参数 R 分成两个非负项: $R = \eta + \mu$, 其中 η 定义为增广凸化参数, μ 定义为逼近参数. 用之前迭代的信息能够定义 $F_{\bar{x}^k}^{\eta}$ 的割平面模型:

$$\begin{aligned} \varphi^k(y) := \max_i \left\{ \left(F_{\bar{x}^k}(x^i) + \frac{\eta}{2} |x^i - \bar{x}^k|^2 \right) + \right. \\ \left. \langle G_i^k + \eta(x^i - \bar{x}^k), y - x^i \rangle \right\} \end{aligned}$$

上式的一个等价表示为

$$\begin{aligned} \varphi_n^k(y) = F_{\bar{x}^k}(\hat{x}^k) + \max_{i \in I_n} \{ -(e_i^k + \eta_n d_i^k), \\ \langle G_i^k + \eta_n \Delta_i^k, y - \hat{x}^k \rangle \} \quad (2) \end{aligned}$$

假设 $x^{n+1} := p_{\mu_n} \varphi_n^k(\hat{x}^k)$ 是 QP 问题的唯一最优解,则对应的最优性条件为存在乘子 α_n 使得

$$x^{n+1} = \hat{x}^k - \frac{1}{\mu_n} G_{\eta_n}^{-n}(\hat{x}^k) \quad (3)$$

$$G_{\eta_n}^{-n}(\hat{x}^k) = \sum_{i \in I_n} \alpha_i^n (G_i^k + \eta_n \Delta_i^k) \in \partial \varphi_n^k(x^{n+1})$$

每次迭代后,用 I_n^{act} 来表示强有效梯度集,即

$$I_n^{\text{act}} := \{i \in I_n : \alpha_i^n > 0\}$$

在最优性条件(3)中,将 $G_{\eta_n}^{-n}(\hat{x}^k)$ 称为集梯度,对应的集元素为

$$\begin{aligned} (e_{-n}^k, d_{-n}^k, \Delta_{-n}^k, G^{-n}) = \sum_{i \in I_n} \alpha_i^n (e_i^k, d_i^k, \Delta_i^k, G_i^k) = \\ \sum_{i \in I_n^{\text{act}}} \alpha_i^n (e_i^k, d_i^k, \Delta_i^k, G_i^k) \quad (4) \end{aligned}$$

则对所有的 $l \in I_n^{\text{act}}$ 和 $l = -n$ 都有

$$\begin{aligned} \varphi_n^k(x^{n+1}) = F_{\bar{x}^k}(\hat{x}^k) - e_l^k - \eta_n d_l^k + \\ \langle G_l^k + \eta_n \Delta_l^k, x^{n+1} - \hat{x}^k \rangle \end{aligned}$$

利用集梯度可以得到

$$G_{\eta_n}^{-n}(\hat{x}^k) = G^{-n} + \eta_n \Delta_{-n}^k = \mu_n (\hat{x}^k - x^{n+1})$$

且割平面 $-e_{-n}^k + \eta_n d_{-n}^k + \langle G_{\eta_n}^{-n}(\hat{x}^k), y - \hat{x}^k \rangle$ 也会

用于生成新的模型函数. 对所有束信息, 当每次下降步变化时, 指标 k 会变成 $k+1$. 相应的束信息会更新为

$$\begin{aligned} e_i^{k+1} &= e_i^k + F_{\hat{x}^k}(\mathbf{x}^i) - F_{\hat{x}^{k+1}}(\mathbf{x}^i) + \\ &\quad \langle \mathbf{G}_i^k, \hat{\mathbf{x}}^k - \mathbf{x}^i \rangle - \langle \mathbf{G}_i^{k+1}, \hat{\mathbf{x}}^{k+1} - \mathbf{x}^i \rangle, \\ d_i^{k+1} &= d_i^k + |\hat{\mathbf{x}}^{k+1} - \hat{\mathbf{x}}^k|^2 / 2 - \langle \mathbf{\Delta}_i^k, \hat{\mathbf{x}}^{k+1} - \hat{\mathbf{x}}^k \rangle, \\ \mathbf{\Delta}_i^{k+1} &= \mathbf{\Delta}_i^k + \hat{\mathbf{x}}^k - \hat{\mathbf{x}}^{k+1} \end{aligned} \quad (5)$$

在下面的算法中, 将 η_n 定义为使增广线性化误差非负的最小值:

$$\eta_n^{\min} := \max_{i \in I_n, d_i^{k(n)} > 0} -\frac{e_i^{k(n)}}{d_i^{k(n)}} \quad (6)$$

显然, 当 $\eta > \eta_n^{\min}$ 时, $e_i^k + \eta d_i^k \geq 0$ 对所有 $i \in I_n$ 都成立. 令 i_k 为使得 $\hat{\mathbf{x}}^k = \mathbf{x}^{i_k}$ 的指标, 则下面的结论成立:

$$i_k \in I_n \Leftrightarrow (e_{i_k}^k, d_{i_k}^k, \mathbf{\Delta}_{i_k}^k, \mathbf{G}_{i_k}^k) = (0, 0, \mathbf{0}, \mathbf{G}_{i_k}^k)$$

在束集合中.

3 算法建立

下面给出求解非光滑复合约束优化问题的再分配束方法.

Step 0 初始化.

选取初始点 $\hat{\mathbf{x}}^0$, 不可接受增长参数 $M_0 > 0$, 参数 $R_0 > 0$, 停止参数 $\epsilon_{\text{stop}} \geq 0$, Armijo 型参数 $m \in (0, 1)$, 凸化增长参数 $\Gamma > 1$. 初始化迭代计数器 $n=0$, 下降步计数器 $k=k(n)=0$ 和 $i_0=0$, 束指标集 $I_0 := \{0\}$ 和初始候选点 $\mathbf{x}^0 := \hat{\mathbf{x}}^0$.

计算 $\mathbf{G}_0^0 \in \partial F_{\hat{\mathbf{x}}^0}(\hat{\mathbf{x}}^0)$ 和其他束信息 $(e_0^0, d_0^0, \mathbf{\Delta}_0^0) = (0, 0, \mathbf{0}) \in \mathbf{R} \times \mathbf{R} \times \mathbf{R}^n$.

选取初始逼近参数对 $(\mu_0, \eta_0) := (R_0, 0)$.

Step 1 停止测试.

若 $0 \in \partial F_{\hat{\mathbf{x}}^k}(\hat{\mathbf{x}}^k)$, 则算法停止.

Step 2 生成局部模型和 QP 子问题.

给定当前下降步逼近点 $\hat{\mathbf{x}}^k$, 束信息集合 $\{(e_i^k, d_i^k, \mathbf{\Delta}_i^k, \mathbf{G}_i^k)\}_{i \in I_n}$ 和逼近参数对 (μ_n, η_n) , 其中 $\eta_n \leq R_n$ 且 $\mu_n = R_n - \eta_n$, 定义局部模型(2), 求解

$$\min_{\mathbf{y} \in \mathbf{R}^n} \varphi_n^k(\mathbf{y}) + \frac{\mu_n}{2} \|\mathbf{y} - \hat{\mathbf{x}}^k\|^2$$

得到最优解 \mathbf{x}^{n+1} 和乘子 α^n , 计算集信息(4).

定义预测下降量

$$\delta_{n+1}^k := F_{\hat{\mathbf{x}}^k}(\hat{\mathbf{x}}^k) + \frac{\eta_n}{2} \|\mathbf{x}^{n+1} - \hat{\mathbf{x}}^k\|^2 - \varphi_n^k(\mathbf{x}^{n+1})$$

Step 3 下降步检测.

检验下降条件

$$F_{\hat{\mathbf{x}}^k}(\mathbf{x}^{n+1}) \leq F_{\hat{\mathbf{x}}^k}(\hat{\mathbf{x}}^k) - m\delta_{n+1}$$

若条件成立, \mathbf{x}^{n+1} 是下降步, 置 $k(n+1) = k+1$, $i_{k+1} = n+1$, $\hat{\mathbf{x}}^{k+1} = \mathbf{x}^{n+1}$, 根据式(5)更新束信息.

否则, \mathbf{x}^{n+1} 是零步, 置 $k(n+1) = k(n)$. 加入束信息

$$\mathbf{\Delta}_{n+1}^k := \mathbf{x}^{n+1} - \hat{\mathbf{x}}^k, d_{n+1}^k := |\mathbf{\Delta}_{n+1}^k|^2 / 2,$$

$$e_{n+1}^k := F_{\hat{\mathbf{x}}^k}(\hat{\mathbf{x}}^k) - (F_{\hat{\mathbf{x}}^k}(\mathbf{x}^{n+1}) + \langle \mathbf{G}_{n+1}^k, \mathbf{\Delta}_{n+1}^k \rangle)$$

选取新的指标集, 满足

$$I_{n+1} \supseteq \{n+1, i_k\}$$

且

$$I_{n+1} \supseteq I_n^{\text{act}} = \{i \in I_n : \alpha_i^n > 0\} \text{ 或 } I_{n+1} \supseteq \{-n\}$$

Step 4 更新 η .

根据下面的规则更新 η . 若 $\eta_n^{\min} \leq \eta_n$, 则 $\eta_{n+1} := \eta_n$. 否则 $\eta_{n+1} := \Gamma \eta_n^{\min}$, $R_n := \mu_n + \eta_{n+1}$. 其中 η_n^{\min} 由式(6)给出.

Step 5 更新 μ .

若 $F_{\hat{\mathbf{x}}^k}(\mathbf{x}^{n+1}) > F_{\hat{\mathbf{x}}^k}(\hat{\mathbf{x}}^k) + M_0$, 则目标函数值的上升是不可接受的. 以下列参数重启算法

$$\eta_0 := \eta_n, \mu_0 := \Gamma \mu_n, R_0 := \eta_0 + \mu_0$$

$$\mathbf{x}^0 := \hat{\mathbf{x}}^0, (e_0^0, d_0^0, \mathbf{\Delta}_0^0) := (0, 0, \mathbf{0})$$

$$k(0) := 0, i_0 := 0, I_0 := \{0\},$$

$$n := 0$$

循环到 Step 1.

不论任何情况, 令 n 加 1, 循环到 Step 1.

在 Step 4 中, 对模型的凸化参数 η_n 的更新保证了 $\eta_n \geq \eta_n^{\min}$ 对所有的迭代都成立, 因此 $e_{-n}^k + \eta_n d_{-n}^k \geq 0$. 此外, 在 Step 2 中定义的预测下降量也满足

$$\delta_{n+1}^k = \frac{R_n + \mu_n}{2} \|\mathbf{x}^{n+1} - \hat{\mathbf{x}}^k\|^2 + e_{-n}^k + \eta_n d_{-n}^k \geq 0 \quad (7)$$

在 Step 5 中的重启能够保证最终所有的迭代点都在集合 L_0 中. 下面的引理能够说明算法的迭代中只有有限次重启.

引理 2(算法是良定义的) 考虑再分配束方法生成的迭代点序列 $\{\mathbf{x}^n\}$, 若 $F_{\hat{\mathbf{x}}^0}$ 满足假设 H1, 且 $i_k \in I_n$, 则在 Step 5 中只有有限次重启. 因此, 最终 $\{\mathbf{x}^n\}$ 完全在 L_0 中, 且逼近参数序列 $\{\mu_n\}$ 会变

成常数.

4 收敛性结论

下面给出关于算法的参数稳定性和渐近收敛性的结论, 这里只给出定理 2 的详细证明.

引理 3(模型性质) 考虑式(2)中给出的函数族 φ_n^k 和再分配算法, 下面的结论成立:

(1) φ_n^k 是凸函数.

(2) 若对式(6)中定义的 η_n^{\min} 有 $\eta_n \geq \eta_n^{\min}$, 则 $\varphi_n^k(\hat{x}^k) \leq F_{\hat{x}^k}(\hat{x}^k) = 0$ 对所有 $w \in \mathbf{R}^n$ 都成立.

(3) 若 $\eta_{n+1} = \eta_n$, $I_{n+1} \supset I_n^{\text{act}}$ 或 $I_{n+1} \supset \{-n\}$, 且 \mathbf{x}^{n+1} 是零步, 则

$$\varphi_{n+1}^k(w) \geq \varphi_n^k(\mathbf{x}^{n+1}) + \mu_n \langle \hat{x}^k - \mathbf{x}^{n+1}, w - \mathbf{x}^{n+1} \rangle$$

(4) 若 $I_n \supset \{n\}$, 则对 $\mathbf{G}_n^k \in \partial F_{\hat{x}^k}(\mathbf{x}^n)$ 有

$$\varphi_n^k(w) \geq F_{\hat{x}^k} + \eta_n d_n^k + \langle \mathbf{G}_n^k + \eta_n \mathbf{A}_n^k, w - \mathbf{x}^n \rangle$$

引理 4(参数的最终稳定性) 考虑函数 φ_n^k , 若 F_{x^0} 满足假设 H1, 则存在 $n' > 0$ 使得所有参数序列满足

$$\eta_n = \bar{\eta}, \mu_n = \bar{\mu}, R_n = \bar{R} = \bar{\mu} + \bar{\eta}, \forall n > n'$$

且引理 3 的结论成立. 此外, 若 $\bar{\eta} > \rho^c$, 则

$$\varphi_n^k(w) \leq F_{\hat{x}^k}(w) + \bar{\eta} |w - \hat{x}^{k(n)}|^2 / 2, \forall w \in L_0, \forall n > n' \quad (8)$$

且上式对 $\{\mathbf{x}^n\}$ 的任一聚点附近的 w 也成立.

定理 2(算法的渐近收敛性) 考虑 F_{x^0} 满足假设 H1, 假设算法不在有限步终止. 令 $\bar{\eta}$ 为凸化参数序列的稳定值, 下列情况只有一种成立.

(1) $\bar{\eta} > \rho^c$ 且

(a) \hat{x} 为最后一个下降步, 则 $\mathbf{x}^{n+1} \rightarrow \hat{x}$, 且 $0 \in \partial F_{\hat{x}}(\hat{x})$.

(b) 存在无穷多个下降步, 则 $\{\hat{x}^n\}$ 的任一聚点 \bar{x} 都满足 $0 \in \partial F_{\bar{x}}(\bar{x})$.

(2) $\bar{\eta} \leq \rho^c$.

证明 (i) 若 $\hat{x} := \hat{x}^k$ 后只有零步, 由文献 [17] 可以得到序列 $\{\mathbf{x}^{n+1}\} \rightarrow \mathbf{p} := \mathbf{p}_{\bar{R}} F_{\hat{x}}(\hat{x})$,

$\varphi_n^k(\mathbf{x}^{n+1}) \rightarrow F_{\hat{x}}(\mathbf{p}) - \frac{\bar{\eta}}{2} |\mathbf{p} - \hat{x}|^2$. 这样就有

$$\begin{aligned} \delta_{n+1}^k &= \frac{\bar{\eta}}{2} |\mathbf{x}^{n+1} - \hat{x}|^2 - \varphi_n^k(\mathbf{x}^{n+1}) \rightarrow \\ &\frac{\bar{\eta}}{2} |\mathbf{p} - \hat{x}|^2 - F_{\hat{x}}(\mathbf{p}) - \frac{\bar{\eta}}{2} |\mathbf{p} - \hat{x}|^2 = \\ &-F_{\hat{x}}(\mathbf{p}) \end{aligned}$$

由于 Step 3 中的下降步测试不成立, 则有 $F_{\hat{x}}(\mathbf{x}^{n+1}) > -m\delta_{n+1}^k$. 令 $n \rightarrow \infty$ 有 $F_{\hat{x}}(\mathbf{p}) > mF_{\hat{x}}(\mathbf{p})$, 故 $F_{\hat{x}}(\mathbf{p}) \geq 0$. 但 $\mathbf{p} = \mathbf{p}_{\bar{R}} F_{\hat{x}}(\hat{x})$ 意味着 $F_{\hat{x}}(\mathbf{p}) + R \cdot \frac{1}{2} |\mathbf{p} - \hat{x}|^2 \leq F_{\hat{x}}(\hat{x})$, 则有 $\hat{x} = \mathbf{p}$. 因此, $\hat{x} = \mathbf{p}_{\bar{R}} F_{\hat{x}}(\hat{x})$, 所以 $0 \in \partial F_{\hat{x}}(\hat{x})$.

(ii) 首先, 注意 $\{\hat{x}^k\} \subset L_0$ 且 L_0 是紧集. 因此, 存在 $K \subset \mathbf{N}$, 对任意 $k \in K$ 有 $\hat{x}^k \rightarrow \mathbf{x}^{\text{inf}} \in L_0$, 因为 $\hat{x}^{k+1} = \mathbf{x}^{i_{k+1}}$, 令 $j_k = i_{k+1} - 1$, 故 $\hat{x}^{k+1} = \mathbf{p}_{\bar{\eta}} \varphi_{j_k}^k(\hat{x}^k)$. 又因为下降步测试成立, 则有

$$F_{\hat{x}^k}(\hat{x}^{k+1}) \leq F_{\hat{x}^k}(\hat{x}^k) - m\delta_{k+1}^k \leq 0$$

即

$$\max\{f(\hat{x}^{k+1}) - f(\hat{x}^k), g(\hat{x}^{k+1}) - g(\hat{x}^k)_+\} \leq 0$$

则 $f(\hat{x}^k)$ 单调递减下有界, 且 $f(\hat{x}^k) \rightarrow f(\mathbf{x}^{\text{inf}})$. 因此

$$\liminf_{k \rightarrow \infty} F_{\hat{x}^k}(\hat{x}^{k+1}) \geq \lim_{k \rightarrow \infty} f(\hat{x}^{k+1}) - f(\hat{x}^k) = 0$$

又因为 $F_{\hat{x}^k}(\hat{x}^{k+1}) \leq 0$, 则 $F_{\hat{x}^k}(\hat{x}^{k+1}) \rightarrow 0$, $\delta_{i_k}^k \rightarrow 0$, 由式(7)可以得到 $|\hat{x}^{k+1} - \hat{x}^k|^2 \rightarrow 0$ 且 $e_{-j_k} + \bar{\eta}^k e_{-j_k} \rightarrow 0$. 这样, 由式(2)可以得到 $\varphi_{j_k}(\hat{x}^{k+1}) - F_{\hat{x}^k}(\hat{x}^k) \rightarrow 0$, 且 $\varphi_{j_k}(\hat{x}^{k+1}) \rightarrow F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}})$, $\forall k \in K$. 但 $\hat{x}^{k+1} = \mathbf{p}_{\bar{\eta}} \varphi_{j_k}(\hat{x}^k)$ 且 $\bar{\eta} > \rho^{\text{id}}$, 则对任意 $w \in L_0$, 由式(8)可得

$$\varphi_{j_k}^k(\hat{x}^{k+1}) + \frac{\bar{\mu}}{2} |\hat{x}^{k+1} - \hat{x}^k| \leq F_{\hat{x}^k}(w) + \frac{\bar{R}}{2} |w - \hat{x}^k|$$

上式两边同时取极限可以得到

$$F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}}) \leq F_{\mathbf{x}^{\text{inf}}}(w) + \frac{\bar{R}}{2} |w - \hat{x}^k|^2, \forall w \in L_0$$

因为 $\mathbf{x}^{\text{inf}} \in L_0$, 对 $w \notin L_0$,

$$F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}}) \leq M_0 \leq F_{x^0}(w) \leq F_{\mathbf{x}^{\text{inf}}}(w) + \frac{\bar{R}}{2} |w - \hat{x}^k|^2$$

因此, $F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}}) \leq F_{\mathbf{x}^{\text{inf}}}(w) + \frac{\bar{R}}{2} |w - \hat{x}^k|^2$ 对所有的 $w \in \mathbf{R}^n$ 都成立, 即 $\mathbf{x}^{\text{inf}} = \mathbf{p}_{\bar{R}} F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}})$, $\bar{R} > \rho^{\text{id}}$, 因为 f 和 g 都是 lower- C^2 的且 $\bar{R} > \rho^{\text{id}}$, 故有 $0 \in \partial F_{\mathbf{x}^{\text{inf}}}(\mathbf{x}^{\text{inf}})$. □

5 数值实验

本章将给出一个数值算例来验证上面提出的非光滑再分配束方法在实际问题中的有效性. 这里引用了一个典型的实际控制问题: 车辆悬挂系

统的 H_2/H_∞ 控制器综合问题^[18]. 车辆悬挂的数学模型展示了车辆的主要动态特征, 例如支撑车辆的重量, 进行不同操作时保持稳定, 提供一定的舒适度, 最小化马路扰动的力量造成的影响等. 文献^[18]中应用的悬挂模型是图 1 所示的“1/4”车辆模型, 这是早期提出的能够分析相关动态性能最简单的模型. 动态方程是通过简单的力平衡建立的. 状态变量 x_1, x_2 分别为从平衡状态的簧下质量 ($M_{us}=28.58 \text{ kg}$) 位移和从平衡状态的簧上质量 ($M_{sp}=288.9 \text{ kg}$) 位移, x_3, x_4 分别为 x_1, x_2 的导数. 如图所示, 根据由车辆悬挂系统的动力学特征得到的数学模型, 系统的状态空间表示为

$$\begin{aligned} \dot{\boldsymbol{x}} &= \boldsymbol{A}\boldsymbol{x} + \boldsymbol{B}_1\boldsymbol{w} + \boldsymbol{B}_2\boldsymbol{u} \\ \boldsymbol{y} &= \boldsymbol{C}_2\boldsymbol{x} \\ \boldsymbol{z} &= \boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{u} \end{aligned} \quad (9)$$

这里 \boldsymbol{x} 是状态向量, \boldsymbol{u} 是控制输入向量, \boldsymbol{y} 是量测输出向量, \boldsymbol{w} 是外部扰动向量, \boldsymbol{z} 是受控输出向量, 剩余的系数矩阵为

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{K_{us}}{M_{us}} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\boldsymbol{B}_1 = \begin{pmatrix} 0 & 0 & \frac{K_{us}}{M_{us}} & 0 \end{pmatrix}^T,$$

$$\boldsymbol{B}_2 = \begin{pmatrix} 0 & 0 & -\frac{1}{M_{us}} & -\frac{1}{M_{sp}} \end{pmatrix}^T,$$

$$\boldsymbol{C}_2 = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

$$\boldsymbol{C} = \begin{pmatrix} 3.3084 & -0.2336 & 0 & 0 \\ -0.2336 & 0.9723 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\boldsymbol{D} = (0 \ 0 \ 0 \ 0 \ 0.2828 \times 10^{-4})$$

其中 $K_{us}=1.559 \times 10^5 \text{ N/m}$, 初始条件为 $\boldsymbol{x}(0)=0$.

在本例中, 考虑扰动信号为阶跃信号, 要设计输出反馈控制器 \boldsymbol{K} , 使得闭环系统满足下列性质:

(1) 内部稳定性. 在闭环系统中 $\boldsymbol{K}(s)$ 渐近稳定 P , 即任意初始状态引起的零输入响应趋于 0.

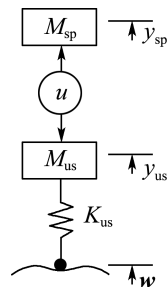


图 1 1/4 车辆模型

Fig. 1 Quarter-car model

(2) 指定的 H_∞ 性能. 从 \boldsymbol{w} 到 \boldsymbol{z} 的闭环传递函数 $\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}$ 的 H_∞ 范数不超过给定的上界 γ_∞ .

(3) 最优的 H_2 性能. 在所有满足(1)和(2)的 \boldsymbol{K} 中, 最小化从 \boldsymbol{w} 到 \boldsymbol{z} 的闭环传递函数 $\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}$ 的 H_2 范数.

标准的动态控制器为

$$\boldsymbol{K}(s): \begin{pmatrix} \dot{\boldsymbol{x}}_K \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_K & \boldsymbol{B}_K \\ \boldsymbol{C}_K & \boldsymbol{D}_K \end{pmatrix} \begin{pmatrix} \boldsymbol{x}_K \\ \boldsymbol{y} \end{pmatrix}$$

其中 \boldsymbol{x}_K 是控制器的状态. 将 $\boldsymbol{K}(s)$ 代入对象(9)得到闭环系统, 引入闭环状态矩阵

$$\boldsymbol{A}(\boldsymbol{K}) = \begin{pmatrix} \boldsymbol{A} + \boldsymbol{B}_2 \boldsymbol{D}_K \boldsymbol{C}_2 & \boldsymbol{B}_2 \boldsymbol{C}_K \\ \boldsymbol{B}_K \boldsymbol{C}_2 & \boldsymbol{A}_K \end{pmatrix},$$

$$\boldsymbol{B}(\boldsymbol{K}) = \begin{pmatrix} \boldsymbol{B}_1 \\ \mathbf{0} \end{pmatrix},$$

$$\boldsymbol{C}(\boldsymbol{K}) = (\boldsymbol{C}_1 \ \mathbf{0}), \quad \boldsymbol{D}(\boldsymbol{K}) = 0$$

则可以得到传递函数

$$\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K}, \omega) = \boldsymbol{C}(\boldsymbol{K})(\omega \boldsymbol{I} - \boldsymbol{A}(\boldsymbol{K}))^{-1} \boldsymbol{B}(\boldsymbol{K}) + \boldsymbol{D}(\boldsymbol{K})$$

因此要解下面的数学规划:

$$\begin{aligned} \min \quad & \|\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K})\|_2^2 \\ \text{s. t.} \quad & \|\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K})\|_\infty^2 \leq \gamma_\infty^2 \end{aligned}$$

即

$$\begin{aligned} \min \quad & f(\boldsymbol{K}) := \|\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K})\|_2^2 \\ \text{s. t.} \quad & g(\boldsymbol{K}) \leq 0 \end{aligned}$$

其中

$$g(\boldsymbol{K}) := \max_{\omega \in [0, \infty]} \lambda_1(\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}^H(\boldsymbol{K}, \omega) \boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K}, \omega)) - \gamma_\infty^2$$

λ_1 是矩阵的最大特征值函数. 注意 $f(\boldsymbol{K})$ 是 C^2 函数, 当 $\boldsymbol{T}_{\boldsymbol{w} \rightarrow \boldsymbol{z}}(\boldsymbol{K}, \omega)$ 对 \boldsymbol{x} 和 ω 是二阶连续可微时, 显然 $g(\boldsymbol{K})$ 是 lower- C^2 的, 满足算法需要的条件. 取 $\gamma_\infty = 5.225$, 经过 43 次迭代得到了二阶的最优控制器

$$\mathbf{K}_1(s) = 10^3 \times$$

$$\begin{pmatrix} 0.0896 & 0.3308 & 0.7272 & 0.0640 \\ -0.0662 & -0.1544 & 0.4871 & -0.0109 \\ 0.7986 & 0.2764 & 0.1617 & 1.7366 \end{pmatrix}$$

控制器 $\mathbf{K}_1(s)$ 的 H_2 范数为 33.2432, H_∞ 范

$$\mathbf{K}_2(s) = 10^3 \times \begin{pmatrix} -0.2158 & -0.5613 & -0.0012 & 0.0006 & -0.0927 & 0.0817 \\ 0.0724 & 0.1512 & -0.0042 & 0.0019 & -0.5677 & -0.0374 \\ 0.0011 & -0.0003 & 0.0011 & 0.0044 & 0.0004 & 0.0005 \\ 0.0044 & -0.0017 & -0.0006 & -0.0018 & 0.0006 & 0.0005 \\ 0.4742 & -0.2764 & 0.0004 & 0.0004 & 0.0922 & 1.7370 \end{pmatrix}$$

控制器 $\mathbf{K}_2(s)$ 的 H_2 范数为 33.2549, H_∞ 范数为 5.2116. 将得到的结果与一般的非光滑束方法^[15]做对比, 文献[15]中得到的控制器的 H_2 范数为 33.3110, H_∞ 范数为 5.2236, 并且需要 157 次迭代. 可以看出提出的再分配束方法在计算所需的时间及求得解的可行性和目标函数值上都有所提高. 图 2 给出了得到的控制器 $\mathbf{K}_1(s)$ 与式(9)构成的闭环系统的阶跃响应.

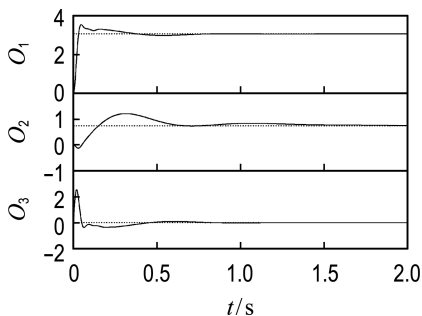


图 2 由受控对象和控制器 $\mathbf{K}_1(s)$ 构成闭环系统的阶跃响应

Fig. 2 Step response of closed-loop system with controlled object and the controller $\mathbf{K}_1(s)$

6 结 语

本文研究了一类特殊的复合约束优化问题, 这类问题在很多重要领域都有广泛的应用. 利用改善函数, 将原问题转换成无约束问题来求解. 利用 lower- C^2 函数的性质, 引入凸化参数来改善新目标函数的凸性, 使得问题容易求解, 本文还使用了集技术来避免迭代过程中束集合过大的缺陷.

数为 5.2181. 将得到的结果与一般的非光滑束方法^[15]做对比, 文献[15]中得到的控制器的 H_2 范数为 33.3120, H_∞ 范数为 5.2250, 并且需要 496 次迭代. 还计算了四阶的最优控制器, 经过 25 次迭代得到了

算法中的参数经过迭代能达到稳定且迭代点列能够收敛到问题的局部最优解. 在数值实验部分, 算法被应用到具有实际意义的 H_2/H_∞ 控制问题, 说明了算法的有效性.

参 考 文 献:

- [1] ACHTZIGER W, KOCVARA M. Structural topology optimization with eigenvalues [J]. *SIAM Journal on Optimization*, 2007, 18(4): 1129-1164.
- [2] MENGI E. A support function based algorithm for optimization with eigenvalue constraints [J]. *SIAM Journal on Optimization*, 2017, 27(1): 246-268.
- [3] POLAK E. *Optimization Algorithms and Consistent Approximates* [M]. New York: Springer, 1997.
- [4] KIWIEL K C. A linearization algorithm for nonsmooth minimization [J]. *Mathematics of Operations Research*, 1985, 10(2): 185-194.
- [5] MAKELA M M, NEITTAANMAKI P. *Nonsmooth Optimization: Analysis and Algorithms with Applications to Optimal Control* [M]. Singapore: World Scientific Publishing Co. Pte. Ltd., 1992.
- [6] VLCEK J, LUKSAN L. Globally convergent variable metric method for nonconvex nondifferentiable unconstrained minimization [J]. *Journal of Optimization Theory and Applications*, 2001, 111(2): 407-430.
- [7] FUDULI A, GAUDIOSO M, GIALLOMBARDO G. Minimizing nonconvex nonsmooth functions via cutting planes and proximity control [J]. *SIAM Journal on Optimization*, 2004, 14(3): 743-756.
- [8] HAARALA N, MIETTINEN K, MAKELA M M.

- Globally convergent limited memory bundle method for large-scale nonsmooth optimization [J]. **Mathematical Programming**, 2007, **109**(1): 181-205.
- [9] KIWIEL K C. Restricted step and Levenberg Marquardt techniques in proximal bundle methods for nonconvex nondifferentiable optimization [J]. **SIAM Journal on Optimization**, 1996, **6**(1): 227-249.
- [10] LUKSAN L, VLCEK J. A bundle-Newton method for nonsmooth unconstrained minimization [J]. **Mathematical Programming**, 1998, **83**(3): 373-391.
- [11] MIFFLIN R. A modification and an extension of Lemarechal's algorithm for nonsmooth minimization [C] // SORENSEN D C, WETS R J B, Eds. **Mathematical Programming Study**. Amsterdam: North-Holland, 1982.
- [12] SCHRAMM H, ZOWE J. A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results [J]. **SIAM Journal on Optimization**, 1992, **2**(1): 121-152.
- [13] HARE W, SAGASTIZABAL C. A redistributed proximal bundle method for nonconvex optimization [J]. **SIAM Journal on Optimization**, 2010, **20**(5): 2442-2473.
- [14] ROCKAFELLAR R T, WETS R. **Variational Analysis** [M]. Berlin: Springer, 1998.
- [15] APKARIAN P, NOLL D, RONDEPIERRE A. Mixed $H_2 = H_1$ control via nonsmooth optimization [J]. **SIAM Journal on Control and Optimization**, 2008, **47**(3): 1516-1546.
- [16] JEYAKUMAR V. On optimality conditions in nonsmooth inequality constrained minimization [J]. **Numerical Functional Analysis and Optimization**, 1987, **9**(5/6): 535-546.
- [17] HARE W, SAGASTIZABAL C. Computing proximal points of nonconvex functions [J]. **Mathematical Programming**, 2009, **116**(1/2): 221-258.
- [18] CAMINO J F, ZAMPIERI D E, PERES P L D. Design of a vehicular suspension controller by static output feedback [C] // **Proceedings of the American Control Conference**. San Diego: IEEE, 1999: 3168-3172.

Redistributed bundle method for solving nonsmooth composite constrained optimization problem

WU Qiong, ZHANG Hongwei*, WANG Shuang

(School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, China)

Abstract: A redistributed-type bundle method is proposed for solving a specific kind of composite optimization problem whose objective function and constraint function are functions with the property of lower- C^2 . By using improvement function, the original constrained problem is transformed into an unconstrained problem, in which the new objective function remains lower- C^2 . By applying the properties of lower- C^2 function, a convexification parameter is introduced to improve the convexity of the objective function in the subproblem, and a bundle method is designed. The relationship between the optimum points of original problem and the new problem is given, together with the results about the stability of parameters and local convergence. Via analyzing H_2/H_∞ control problem, an optimal H_2/H_∞ dynamic controller is computed by the proposed method and the effectiveness of the method is demonstrated.

Key words: constrained optimization; nonsmooth optimization; improvement function; bundle method; lower- C^2