应用数学和力学
APPLIED MATHEMATICS AND MECHANICS

存零约束优化问题的序列二次方法

罗美铃，李高西，黄应全，刘丽颖

SQP Methods for Mathematical Programs With Switching Constraints

LUO Meiling, LI Gaoxi, HUANG Yingquan, and LIU Liying

在线阅读 View online: https://doi.org/10.21656/1000-0887.420294 http://www.applmathmech.cn

存零约束优化问题的序列二次方法*

罗美铃，　李高西，　黄应全，　刘丽颖

（重庆工商大学 数学与统计学院，重庆 400067）

摘要： 存零约束优化 (MPSC) 问题是近年来提出的一类新的优化问题, 因存零约束的存在, 使得常用的约束规范不满足, 以至于现有算法的收敛性结果大多不能直接应用于该问题. 应用序列二次规划 (SQP) 方法求解该问题, 并证明在存零约束的线性独立约束规范下, 子问题解序列的聚点为原问题的 Karush-Kuhn-Tucker 点. 同时为了完善各稳定点之间的关系, 证明了强平稳点与 KKT 点的等价性. 最后数值结果表明, 序列二次规划方法处理这类问题是可行的.

关　键　词： 非线性规划；　存零约束优化；　序列二次方法；　全局收敛性
中图分类号： O211　　**文献标志码：** A　　DOI：10.21656/1000-0887.420294

SQP Methods for Mathematical Programs With Switching Constraints

LUO Meiling,　LI Gaoxi,　HUANG Yingquan,　LIU Liying
(*School of Mathematics and Statistics*, *Chongqing Technology and Business University*, *Chongqing* 400067, *P.R.China*)

Abstract: The mathematical program with switching constraint (MPSC) problem makes a new-type optimization issue in recent years. Due to the existence of switching constraints, the common constraint specification is not satisfied, so that the convergence results of existing algorithms can not be directly applied to this problem. The sequential quadratic programming (SQP) method was applied to solve the problem, and to prove that the clustering point of the solution sequence of the subproblem is the Karush-Kuhn-Tucker point of the original problem under the linear independent constraint specification with the switching constraint. At the same time, in order to improve the relationship between stationary points, the equivalence between the strong stationary point and the KKT point was proved. Finally, the numerical results show that, the sequential quadratic programming method is feasible to deal with this type of problems.

Key words: nonlinear programming; mathematical program with switching contraint; sequential quadratic programming method; global convergence

引　言

本文考虑下述非线性规划问题:

---

* **收稿日期：** 2021-09-26; **修订日期：** 2021-11-23
基金项目： 国家自然科学基金 (11901068); 重庆市自然科学基金 (cstc2019jcyj-msxmX0760; cstc2021jcyj-msxmX0499)
作者简介： 罗美铃 (1998—), 女, 硕士生 (E-mail: hana_lml@163.com);
李高西 (1988—), 男, 博士 (通讯作者. E-mail: ligaoxicn@163.com).
引用格式： 罗美铃, 李高西, 黄应全, 刘丽颖. 存零约束优化问题的序列二次方法[J]. 应用数学和力学, 2022, **43**(7): 792-801.

$$\begin{cases} \min \ f(\boldsymbol{x}), \\ \quad \text{s.t.} \ \ g_i(\boldsymbol{x}) \geqslant 0, \quad i = 1,2,\cdots,p, \\ \qquad\quad h_j(\boldsymbol{x}) = 0, \quad j = 1,2,\cdots,q, \\ \qquad\quad G_t(\boldsymbol{x})H_t(\boldsymbol{x}) = 0, \quad t = 1,2,\cdots,l, \end{cases} \tag{1}$$

其中, 函数 $f(\boldsymbol{x})$, $g_1,\cdots,g_p$, $h_1,\cdots,h_q$, $G_1,\cdots,G_l$, $H_1,\cdots,H_l$: $\mathbb{R}^n \to \mathbb{R}$ 均连续可微. 为简单起见, 我们令 $\boldsymbol{g} = (g_1,\cdots,g_p)^{\mathrm{T}}$, $\boldsymbol{h} = (h_1,\cdots,h_q)^{\mathrm{T}}$, $\boldsymbol{G} = (G_1,\cdots,G_l)^{\mathrm{T}}$, $\boldsymbol{H} = (H_1,\cdots,H_l)^{\mathrm{T}}$. 在问题 (1) 任意可行点 $\boldsymbol{x}$ 处, 对任意固定的 t, $G_t(\boldsymbol{x})$和$H_t(\boldsymbol{x})$至少有一个为零, 我们称这样的约束为存零约束, 称问题 (1) 为 “存零约束优化” 问题, 简记为 MPSC 问题. 该模型首先由 Mehlitz[1] 提出并系统研究. 文献 [1] 指出, 最优控制问题的离散化、either-or 约束优化、0-1 规划等问题均可以转化为问题 (1), 因此有必要讨论问题 (1) 的理论与算法. 由于常用的约束规范, 比如线性独立约束规范 (LICQ)、Mangasarian-Fromovitz 约束规范 (MFCQ) 在问题 (1) 的可行点处均不满足, 因此不能将其看作一般的非线性规划处理.

MPSC 的约束结构与具有互补约束的优化问题和具有消失约束的优化问题密切相关 (具体可见文献 [2]). 最近, Mehlitz[1] 给出了一些平稳性概念, 如弱平稳性、Mordukhovich (M-) 平稳性和强 (S) 平稳性来确保这些平稳条件在问题的最优点处成立, 然后提出了一些 MPSC 定制的约束条件, 例如 MPSC Mangasarian-Fromovitz 约束规范 (MPSC-MFCQ)、MPSC 线性独立约束规范 (MPSC-LICQ)、MPSC Abadie 约束规范和 MPSC Guignard 约束规范. 在文献 [1] 的基础上, Luo 等[3] 提出了一种求解 MPSC 的松弛方法; 然后, Achtziger 等[4] 研究了 MPSC 的二阶最优性条件; 最近, Liang 等[5] 讨论了 MPSC 的最优性条件和精确罚问题.

序列二次规划算法, 简称 SQP 算法, 最早由 Wilson 提出. 在数值效果和稳定性方面, 它是目前约束优化问题的一种最有效的方法. 随着 SQP 算法自身理论的不断完善, 许多学者将该算法应用到工程和经济等各种实际优化模型中, 例如均衡问题[6-7]、最优控制问题[8]、极大极小问题[9]、半无限规划问题[10] 等, 并取得了显著的效果. Wright[11] 建议用改进的 SQP 算法处理非线性规划问题. Fletcher 等[6] 将原始的 SQP 算法直接应用于均衡约束优化 (MPEC) 问题, 并得到了相关的收敛性结论. 随后, 朱志斌等[12] 同样以 MPEC 为背景, 引入线搜索和高阶修正方向, 从而获得了更好的结果. 由于求解一般非线性规划问题的标准约束规范 LICQ 和 MFCQ 在 MPSC 中的任一可行点处均失效, 我们便引入 MPSC-LICQ 到问题 (1) 中, 并保证该问题在 SQP 算法下收敛. 本文将 SQP 算法应用于问题 (1) 中, 进而分析该算法的收敛性过程, 这是本文研究的主要内容.

本文主要结构如下: 第 1 节给出本文研究所需的基本定义, 同时讨论了强平稳条件与 KKT 条件的关系; 第 2 节给出了该问题的序列二次算法框架; 第 3 节在 MPSC-LICQ 条件下给出了算法的收敛性; 第 4 节给出数值实验结果; 最后是所得结论.

1　预 备 知 识

在本节中, 我们先建立一些相关的指标集, 同时给出与 MPSC 相关的基本定义, 随之分析问题 (1) 的强平稳点和 KKT 点之间的关系.

为便于后续讨论, 定义指标集:

$$\begin{cases} \mathcal{I}_g^* = \{i \in \{1,2,\cdots,p\} : g_i(\boldsymbol{x}^*) = 0\}, \ \mathcal{I}_h^* = \{1,2,\cdots,q\}, \\ \mathcal{I}_G^* = \{t \in \{1,2,\cdots,l\} : G_t(\boldsymbol{x}^*) = 0 \wedge H_t(\boldsymbol{x}^*) \neq 0\}, \\ \mathcal{I}_H^* = \{t \in \{1,2,\cdots,l\} : G_t(\boldsymbol{x}^*) \neq 0 \wedge H_t(\boldsymbol{x}^*) = 0\}, \\ \mathcal{I}_{GH}^* = \{t \in \{1,2,\cdots,l\} : G_t(\boldsymbol{x}^*) = 0 \wedge H_t(\boldsymbol{x}^*) = 0\}, \end{cases} \tag{2}$$

其中, $\{\mathcal{I}_G^*, \mathcal{I}_H^*, \mathcal{I}_{GH}^*\}$为$\{1,2,\cdots,l\}$上不相交的划分.

定义 1[1]　设$\boldsymbol{x}^* \in \mathbb{R}^n$为 MPSC 的可行点, 如果存在乘子满足

$$\boldsymbol{0} = \nabla f(\boldsymbol{x}^*) - \sum_{i \in \mathcal{I}_g^*} \lambda_i \nabla g_i(\boldsymbol{x}^*) + \sum_{j \in \mathcal{I}_h^*} \lambda_j \nabla h_j(\boldsymbol{x}^*) + \sum_{t=1}^{l} [\mu_t \nabla G_t(\boldsymbol{x}^*) + \nu_t \nabla H_t(\boldsymbol{x}^*)],$$
$$\forall i \in \mathcal{I}_g^* : \lambda_i \geqslant 0; \ \forall t \in \mathcal{I}_G^* : \nu_t = 0; \ \forall t \in \mathcal{I}_H^* : \mu_t = 0; \ \forall t \in \mathcal{I}_{GH}^* : \mu_t = 0 \wedge \nu_t = 0, \tag{3}$$

则我们称$\boldsymbol{x}^*$为强平稳点, 简记为 S-平稳点.

定义 2[1] 设$\boldsymbol{x}^* \in \mathbb{R}^n$为 MPSC 的可行点, 如果在$\boldsymbol{x}^*$点处约束函数的梯度

$$\begin{cases} \nabla g_i(\boldsymbol{x}^*), & i \in \mathcal{I}_g^*, \\ \nabla h_j(\boldsymbol{x}^*), & j \in \mathcal{I}_h^*, \\ \nabla G_t(\boldsymbol{x}^*), & t \in \mathcal{I}_G^* \cup \mathcal{I}_{GH}^*, \\ \nabla H_t(\boldsymbol{x}^*), & t \in \mathcal{I}_H^* \cup \mathcal{I}_{GH}^* \end{cases} \tag{4}$$

满足线性独立, 我们称 MPSC-LICQ 在$\boldsymbol{x}^*$点处成立.

在文献 [1] 中, Mehlitz 通过构造 MPSC 特定的约束规格, 从而来研究原问题 (1) 的局部最优解与各平稳点之间的关系. 下面我们将直接分析原问题的 KKT 点与 S-平稳点的联系.

定理 1 设$\boldsymbol{x}^*$为问题 (1) 的 KKT 点, 当且仅当$\boldsymbol{x}^*$为该问题的 S-平稳点.

证明 必要性

由题设, 存在乘子$\lambda_i(i \in \mathcal{I}_g^*)$, $\lambda_j(j \in \mathcal{I}_h^*)$, $\lambda_t(t = 1,2,\cdots,l)$是 MPSC 的 KKT 系统的解, 即

$$\mathbf{0} = \nabla f(\boldsymbol{x}^*) - \sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*) + \sum_{t=1}^{l}\lambda_t[G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*) + H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*)], \qquad \forall i \in \mathcal{I}_g^* : \lambda_i \geqslant 0.$$

先定义函数$K_t : \mathbb{R}^n \to \mathbb{R}$, $K_t(\boldsymbol{x}) = G_t(\boldsymbol{x})H_t(\boldsymbol{x})$, $\forall \boldsymbol{x} \in \mathbb{R}$. 对于任意$t \in \{1,2,\cdots,l\}$, 函数$K_t$在$\boldsymbol{x}^*$的梯度为

$$\nabla K_t(\boldsymbol{x}^*) = H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*) + G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*) = \begin{cases} H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*), & t \in \mathcal{I}_G^*, \\ G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*), & t \in \mathcal{I}_H^*, \\ \mathbf{0}, & t \in \mathcal{I}_{GH}^*. \end{cases}$$

则有

$$\mathbf{0} = \nabla f(\boldsymbol{x}^*) - \sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_H^*}\lambda_t G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_G^*}\lambda_t H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*), \qquad \forall i \in \mathcal{I}_g^* : \lambda_i \geqslant 0.$$

进一步, 定义$\mu_t := \lambda_t H_t(\boldsymbol{x}^*)(t \in \mathcal{I}_G^*)$, $\nu_t := \lambda_t G_t(\boldsymbol{x}^*)(t \in \mathcal{I}_H^*)$, 从而有

$$\mathbf{0} = \nabla f(\boldsymbol{x}^*) - \sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_H^*}\nu_t\nabla H_t(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_G^*}\mu_t\nabla G_t(\boldsymbol{x}^*), \qquad \forall i \in \mathcal{I}_g^* : \lambda_i \geqslant 0.$$

即得证$\boldsymbol{x}^*$为原问题的 S-平稳点.

充分性

令$\boldsymbol{x}^*$是原问题的 S-平稳点, 由定义 1, 设对任意$t \in \{1,2,\cdots,l\}$, 有

$$\lambda_t := \begin{cases} \dfrac{\mu_t}{H_t(\boldsymbol{x}^*)}, & t \in \mathcal{I}_G^*, \\ \dfrac{\nu_t}{G_t(\boldsymbol{x}^*)}, & t \in \mathcal{I}_H^*, \\ 0, & t \in \mathcal{I}_{GH}^*. \end{cases}$$

则有

$$\nabla f(\boldsymbol{x}^*) - \sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*) + \sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_G^*}\lambda_t H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*) + \sum_{t\in\mathcal{I}_H^*}\lambda_t G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*) = \mathbf{0}, \qquad \forall i \in \mathcal{I}_g^* : \lambda_i \geqslant 0.$$

从而等价于

$$
\begin{aligned}
&\nabla f(\boldsymbol{x}^*)-\sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*)+\sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*)+\sum_{t\in\mathcal{I}_G^*}\lambda_t[H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*)+G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*)]+\\
&\quad\sum_{t\in\mathcal{I}_H^*}\lambda_t[G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*)+H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*)]+\\
&\quad\sum_{t\in\mathcal{I}_{GH}^*}\lambda_t[H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*)+G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*)]=\boldsymbol{0},\qquad \forall i\in\mathcal{I}_g^*:\lambda_i\geqslant 0.
\end{aligned}
$$

又由于$\{1,2,\cdots,l\}=I_G^*\cup I_H^*\cup I_{GH}^*$, 进而得到

$$
\begin{aligned}
&\nabla f(\boldsymbol{x}^*)-\sum_{i\in\mathcal{I}_g^*}\lambda_i\nabla g_i(\boldsymbol{x}^*)+\sum_{j\in\mathcal{I}_h^*}\lambda_j\nabla h_j(\boldsymbol{x}^*)+\\
&\quad\sum_{t=1}^{l}\lambda_t[H_t(\boldsymbol{x}^*)\nabla G_t(\boldsymbol{x}^*)+G_t(\boldsymbol{x}^*)\nabla H_t(\boldsymbol{x}^*)]=\boldsymbol{0},\qquad \forall i\in\mathcal{I}_g^*:\lambda_i\geqslant 0.
\end{aligned}
$$

2 QP 问题及算法

本节以文献 [13] 为背景, 在标准非线性规划问题的基础上引入了存零约束, 同时给出相应的二次规划子问题以及模型对应的算法.

设$\boldsymbol{x}^k$是原问题的近似解, 向量$\boldsymbol{\lambda}_\varepsilon^k,\boldsymbol{\lambda}_s^k$和非负向量$\boldsymbol{\lambda}_\iota^k$为该优化问题在$\boldsymbol{x}^k$点最优 Lagrange 乘子的一个估计. 因此建立 MPSC 的二次规划子问题:

$$
\begin{cases}
\min\ \ f(\boldsymbol{x}^k)+[\nabla f(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}+\dfrac{1}{2}\boldsymbol{d}^{\mathrm{T}}\nabla_{xx}L(\boldsymbol{x}^k,\boldsymbol{\lambda}_\iota^k,\boldsymbol{\lambda}_\varepsilon^k,\boldsymbol{\lambda}_s^k)\boldsymbol{d},\\
\quad\text{s.t.}\ \ g_i(\boldsymbol{x}^k)+[\nabla g_i(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}\geqslant 0,\qquad i\in\{1,2,\cdots,p\},\\
\qquad\quad h_j(\boldsymbol{x}^k)+[\nabla h_j(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}=0,\qquad j\in\{1,2,\cdots,q\},\\
\qquad\quad G_t(\boldsymbol{x}^k)H_t(\boldsymbol{x}^k)+[G_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+H_t(\boldsymbol{x}^k)\nabla \boldsymbol{G}_t(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}=0,\qquad t\in\{1,2,\cdots,l\}.
\end{cases}
\tag{5}
$$

对子问题产生的搜索方向$\boldsymbol{d}$无论步长怎样选取, 都不能保证新的迭代点为原问题的可行点. 为此, 我们引进一个势函数来确定步长, 为使目标函数值下降, 同时又使迭代点接近可行, 我们使用ℓ_1精确罚函数:

$$
P(\boldsymbol{x},\pi)=f(\boldsymbol{x})+\sum_{j=1}^{q}\pi_j|h_j(\boldsymbol{x})|+\sum_{t=1}^{l}\pi_t|G_t(\boldsymbol{x})H_t(\boldsymbol{x})|+\sum_{i=1}^{p}\pi_i\max\{0,-g_i(\boldsymbol{x})\}.
\tag{6}
$$

基于问题 (5), 首先给出算法的具体步骤.

算法

步 0 给定初始点$(\boldsymbol{x}^0,\boldsymbol{\lambda}_\iota^0,\boldsymbol{\lambda}_\varepsilon^0,\boldsymbol{\lambda}_s^0)$, 对称正定矩阵$\boldsymbol{B}^0\in\mathbb{R}^{n\times n}$. 计算初始步各约束梯度:

$$
\boldsymbol{D}_{\mathcal{J}}^0=[\nabla g(\boldsymbol{x}^0)]^{\mathrm{T}},\ \boldsymbol{D}_{\mathcal{E}}^0=\begin{bmatrix}[\nabla h(\boldsymbol{x}^0)]^{\mathrm{T}}\\ [G_1(\boldsymbol{x}^0)\nabla H_1(\boldsymbol{x}^0)+H_1(\boldsymbol{x}^0)\nabla G_1(\boldsymbol{x}^0)]^{\mathrm{T}}\\ \vdots\\ [G_l(\boldsymbol{x}^0)\nabla H_l(\boldsymbol{x}^0)+H_l(\boldsymbol{x}^0)\nabla G_l(\boldsymbol{x}^0)]^{\mathrm{T}}\end{bmatrix},
$$

$$
\boldsymbol{D}^0=\begin{bmatrix}\boldsymbol{D}_{\mathcal{J}}^0\\ \boldsymbol{D}_{\mathcal{E}}^0\end{bmatrix},\ \boldsymbol{A}^0=\begin{bmatrix}h(\boldsymbol{x}^0)\\ G_1(\boldsymbol{x}^0)H_1(\boldsymbol{x}^0)\\ \vdots\\ G_l(\boldsymbol{x}^0)H_l(\boldsymbol{x}^0)\end{bmatrix}.
$$

选择参数$\eta\in(0,1/2)$, $\rho\in(0,1)$, 设容许误差为$0\leqslant\varepsilon_1,\varepsilon_2\ll 1$. 令$k:=0$.

步 1 求解子问题

$$
\begin{aligned}
&\min\ \ \frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}^k\boldsymbol{d}+[\nabla f(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d},\\
&Q(\boldsymbol{x}^k,\boldsymbol{B}^k)\qquad \text{s.t.}\ \ g(\boldsymbol{x}^k)+\boldsymbol{D}_{\mathcal{J}}^k\boldsymbol{d}\geqslant 0,\\
&\qquad\qquad\qquad\qquad\ \boldsymbol{A}^k+\boldsymbol{D}_{\mathcal{E}}^k\boldsymbol{d}=\boldsymbol{0}
\end{aligned}
$$

的最优解$\boldsymbol{d}^k$.

步 2 若$\| \boldsymbol{d}^k \| \leqslant \varepsilon_1$, 且$\| (g(\boldsymbol{x}^k))_- \|_1 + \| h(\boldsymbol{x}^k) \|_1 + \| (G_1(\boldsymbol{x}^k)H_1(\boldsymbol{x}^k), \cdots, G_l(\boldsymbol{x}^k)H_l(\boldsymbol{x}^k))^{\mathrm{T}} \|_1 \leqslant \varepsilon_2$, 停止计算, 得到原问题的一个近似的下降方向.

步 3 将ℓ_1精确罚函数作为价值函数$P(\boldsymbol{x},\pi)$, 选择罚参数π^k, 使得$\boldsymbol{d}^k$是该函数在$\boldsymbol{x}^k$处的下降方向.

步 4 Armijo 搜索. 令m^k是使下列不等式成立的最小非负整数m:

$$P(\boldsymbol{x}^k + \rho^m \boldsymbol{d}^k, \pi^k) - P(\boldsymbol{x}^k, \pi^k) \leqslant \eta \rho^m P'(\boldsymbol{x}^k, \pi; \boldsymbol{d}^k).$$

令$\alpha^k := \rho^{m^k}, \boldsymbol{x}^{k+1} := \boldsymbol{x}^k + \alpha^k \boldsymbol{d}^k$.

步 5 计算下一步迭代的各约束梯度

$$\boldsymbol{D}_{\mathcal{J}}^{k+1} = [\nabla g(\boldsymbol{x}^{k+1})]^{\mathrm{T}},\ \boldsymbol{D}_{\mathcal{E}}^{k+1} = \begin{bmatrix} [\nabla h(\boldsymbol{x}^{k+1})]^{\mathrm{T}} \\ [G_1(\boldsymbol{x}^{k+1})\nabla H_1(\boldsymbol{x}^{k+1}) + H_1(\boldsymbol{x}^{k+1})\nabla G_1(\boldsymbol{x}^{k+1})]^{\mathrm{T}} \\ \vdots \\ [G_l(\boldsymbol{x}^{k+1})\nabla H_l(\boldsymbol{x}^{k+1}) + H_l(\boldsymbol{x}^{k+1})\nabla G_l(\boldsymbol{x}^{k+1})]^{\mathrm{T}} \end{bmatrix},$$

$$\boldsymbol{D}^{k+1} = \begin{bmatrix} \boldsymbol{D}_{\mathcal{J}}^{k+1} \\ \boldsymbol{D}_{\mathcal{E}}^{k+1} \end{bmatrix},$$

以及对应的最小二乘乘子

$$\begin{bmatrix} \boldsymbol{\lambda}_\iota^{k+1} \\ \boldsymbol{\lambda}_\varepsilon^{k+1} \\ \boldsymbol{\lambda}_s^{k+1} \end{bmatrix} = [\boldsymbol{D}^{k+1}(\boldsymbol{D}^{k+1})^{\mathrm{T}}]^{-1} \boldsymbol{D}^{k+1} \nabla f^{k+1}.$$

步 6 校正矩阵$\boldsymbol{B}^k$为$\boldsymbol{B}^{k+1}$. 令

$$\boldsymbol{s}^k = \alpha^k \boldsymbol{d}^k,\ \boldsymbol{y}^k = \nabla_{\boldsymbol{x}} L(\boldsymbol{x}^{k+1}, \boldsymbol{\lambda}_\iota^{k+1}, \boldsymbol{\lambda}_\varepsilon^{k+1}, \boldsymbol{\lambda}_s^{k+1}) - \nabla_{\boldsymbol{x}} L(\boldsymbol{x}^k, \boldsymbol{\lambda}_\iota^{k+1}, \boldsymbol{\lambda}_\varepsilon^{k+1}, \boldsymbol{\lambda}_s^{k+1}),$$

$$\boldsymbol{B}^{k+1} = \boldsymbol{B}^k - \frac{\boldsymbol{B}^k \boldsymbol{s}^k (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k}{(\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k \boldsymbol{s}^k} + \frac{\boldsymbol{z}^k (\boldsymbol{z}^k)^{\mathrm{T}}}{(\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{z}^k},$$

其中

$$\boldsymbol{z}^k = \theta^k \boldsymbol{y}^k + (1-\theta^k) \boldsymbol{B}^k \boldsymbol{s}^k,$$

参数θ^k定义为

$$\theta^k = \begin{cases} 1, & (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{y}^k \geqslant 0.2 (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k \boldsymbol{s}^k, \\ \dfrac{0.8 (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k \boldsymbol{s}^k}{(\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k \boldsymbol{s}^k - (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{y}^k}, & (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{y}^k < 0.2 (\boldsymbol{s}^k)^{\mathrm{T}} \boldsymbol{B}^k \boldsymbol{s}^k. \end{cases}$$

步 7 令$k := k+1$, 转步 1.

为后续研究, 我们给出子问题$Q(\boldsymbol{x}, \boldsymbol{B})$的最优系统:

$$\begin{cases} \nabla f(\boldsymbol{x}) + \boldsymbol{B}\boldsymbol{d} - \sum\limits_{i \in \mathcal{I}_g^*} \lambda_i \nabla g_i(\boldsymbol{x}) + \sum\limits_{j \in \mathcal{I}_h^*} \lambda_j \nabla h_j(\boldsymbol{x}) + \\ \qquad \sum\limits_{t=1}^{l} \lambda_t [G_t(\boldsymbol{x}) \nabla H_t(\boldsymbol{x}) + H_t(\boldsymbol{x}) \nabla G_t(\boldsymbol{x})] = \boldsymbol{0}, \qquad \lambda_i \geqslant 0, \\ g_i(\boldsymbol{x}) + [\nabla g_i(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{d} \geqslant 0, \qquad i \in \{1, 2, \cdots, p\}, \\ h_j(\boldsymbol{x}) + [\nabla h_j(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{d} = 0, \qquad j \in \{1, 2, \cdots, q\}, \\ \lambda_i (g_i(\boldsymbol{x}) + [\nabla g_i(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{d}) = 0, \qquad i \in \{1, 2, \cdots, p\}, \\ G_t(\boldsymbol{x}) H_t(\boldsymbol{x}) + [G_t(\boldsymbol{x}) \nabla H_t(\boldsymbol{x}) + H_t(\boldsymbol{x}) \nabla G_t(\boldsymbol{x})]^{\mathrm{T}} \boldsymbol{d} = 0, \qquad t \in \{1, 2, \cdots, l\}, \end{cases} \tag{7}$$

其中$(\boldsymbol{d}, \boldsymbol{\lambda})$为原始-对偶平稳点.

3 SQP 算法的全局收敛性

本节我们将在给定的假设条件下分析第 2 节中算法的收敛性.

引理 1 设$f(\boldsymbol{x})$和$g_i(\boldsymbol{x})(i \in \{1, 2, \cdots, p\})$, $h_j(\boldsymbol{x})(j \in \{1, 2, \cdots, q\})$, $G_t(\boldsymbol{x})$, $H_t(\boldsymbol{x})(t \in \{1, 2, \cdots, l\})$连续可微, 若在可行点$\boldsymbol{x}$处 MPSC-LICQ 成立, 算法产生的序列$\{\boldsymbol{x}^k\}$和$\{\boldsymbol{d}^k\}$是有界的, 并且$\lim_{k \in \mathcal{N}_0, k \to \infty} \boldsymbol{x}^k = \boldsymbol{x}^*$, 令$\boldsymbol{\lambda}^k := (\boldsymbol{\lambda}_\iota^k, \boldsymbol{\lambda}_\varepsilon^k, \boldsymbol{\lambda}_s^k)$, 则序列$\{\boldsymbol{\lambda}^k\}$

有界.

证明　用反证法, 假设序列$\{\lambda^k\}$无界, 则存在$\mathcal{N}_1\subseteq\mathcal{N}_0$, 使得

$$\|\lambda^k\|\to\infty,\qquad k\in\mathcal{N}_1.$$

根据子问题$Q(\boldsymbol{x}^k,\boldsymbol{B}^k)$最优系统第一个等式有

$$\begin{aligned}\boldsymbol{0}=&\nabla f(\boldsymbol{x}^k)+\boldsymbol{B}^k\boldsymbol{d}^k-\sum_{i\in\mathcal{I}_g^*}\lambda_i^k\nabla g_i(\boldsymbol{x}^k)+\sum_{j\in\mathcal{I}_h^*}\lambda_j^k\nabla h_j(\boldsymbol{x}^k)+\\&\sum_{t=1}^{l}\lambda_t^k[G_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+H_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k)]=\\&\nabla f(\boldsymbol{x}^k)+\boldsymbol{B}^k\boldsymbol{d}^k-\sum_{i\in\mathcal{I}_g^*}\lambda_i^k\nabla g_i(\boldsymbol{x}^k)+\sum_{j\in\mathcal{I}_h^*}\lambda_j^k\nabla h_j(\boldsymbol{x}^k)+\\&\sum_{t=1}^{l}\lambda_t^kG_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+\sum_{t=1}^{l}\lambda_t^kH_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k)=\\&\nabla f(\boldsymbol{x}^k)+\boldsymbol{B}^k\boldsymbol{d}^k-\sum_{i\in\mathcal{I}_g^*}\lambda_i^k\nabla g_i(\boldsymbol{x}^k)+\sum_{j\in\mathcal{I}_h^*}\lambda_j^k\nabla h_j(\boldsymbol{x}^k)+\\&\sum_{t\in\mathcal{I}_H^*}\lambda_t^kG_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+\sum_{t\in\{1,2,\cdots,l\}\backslash\mathcal{I}_H^*}\lambda_t^kG_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+\\&\sum_{t\in\mathcal{I}_G^*}\lambda_t^kH_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k)+\sum_{t\in\{1,2,\cdots,l\}\backslash\mathcal{I}_G^*}\lambda_t^kH_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k).\end{aligned}$$

等式两边同时乘以$1/\|\lambda^k\|$并取极限, 我们设$(\lambda_t^k,\lambda_\varepsilon^k,\lambda_s^k)/\|\lambda^k\|$的极限点为$(\widetilde{\lambda}_t,\widetilde{\lambda}_\varepsilon,\widetilde{\lambda}_s)$, 此时的极限点不为$\boldsymbol{0}$, 且

$$-\sum_{i\in\mathcal{I}_g^*}\widetilde{\lambda}_i\nabla\boldsymbol{g}_i(\boldsymbol{x}^*)+\sum_{j\in\mathcal{I}_h^*}\widetilde{\lambda}_j\nabla\boldsymbol{h}_j(\boldsymbol{x}^*)+\sum_{t\in\mathcal{I}_H^*}\widetilde{\lambda}_tG_t(\boldsymbol{x}^*)\nabla\boldsymbol{H}_t(\boldsymbol{x}^*)+\sum_{t\in\mathcal{I}_G^*}\widetilde{\lambda}_tH_t(\boldsymbol{x}^*)\nabla\boldsymbol{G}_t(\boldsymbol{x}^*)=\boldsymbol{0}.$$

上述等式与 MPSC-LICQ 矛盾. 因此$\{\lambda^k\}$有界.

类似于文献 [13] 中的定理 11.2.1 可得如下定理.

定理 2　设$f(\boldsymbol{x})$和$g_i(\boldsymbol{x})(i\in\{1,2,\cdots,p\})$, $h_j(\boldsymbol{x})(j\in\{1,2,\cdots,q\})$, $G_t(\boldsymbol{x}),H_t(\boldsymbol{x})(t\in\{1,2,\cdots,l\})$ 连续可微, 矩阵 $\boldsymbol{B}$ 正定, 若$(\boldsymbol{d},\lambda)$为子问题$Q(\boldsymbol{x},\boldsymbol{B})$的 KKT 点对, 其中$\boldsymbol{d}\neq\boldsymbol{0}$, 引理 1 成立, 且有$|\lambda_\tau|\leqslant\pi_\tau$, $\tau\in\{1,2,\cdots,p\}\bigcup\{1,2,\cdots,q\}\bigcup\{1,2,\cdots,l\}$, 则$P'(\boldsymbol{x},\pi;\boldsymbol{d})<0$.

证明　定义指标集:

$$\begin{aligned}&\mathcal{J}^<=\{i\in\{1,2,\cdots,p\}|g_i(\boldsymbol{x})<0\},\ \mathcal{E}^<=\{j\in\{1,2,\cdots,q\}|h_j(\boldsymbol{x})<0\},\\&\mathcal{J}^==\{i\in\{1,2,\cdots,p\}|g_i(\boldsymbol{x})=0\},\ \mathcal{E}^==\{j\in\{1,2,\cdots,q\}|h_j(\boldsymbol{x})=0\},\\&\mathcal{J}^>=\{i\in\{1,2,\cdots,p\}|g_i(\boldsymbol{x})>0\},\ \mathcal{E}^>=\{j\in\{1,2,\cdots,q\}|h_j(\boldsymbol{x})>0\},\\&\mathcal{S}^<=\{t\in\{1,2,\cdots,l\}|G_t(\boldsymbol{x})H_t(\boldsymbol{x})<0\},\\&\mathcal{S}^==\{t\in\{1,2,\cdots,l\}|G_t(\boldsymbol{x})H_t(\boldsymbol{x})=0\},\\&\mathcal{S}^>=\{t\in\{1,2\cdots,l\}|G_t(\boldsymbol{x})H_t(\boldsymbol{x})>0\}.\end{aligned}$$

并根据文献 [13] 中的引理 11.2.1, 我们有

$$\begin{aligned}P'(\boldsymbol{x},\pi;\boldsymbol{d})=&[\nabla f(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}-\sum_{j\in\mathcal{E}^<}\pi_j[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\sum_{j\in\mathcal{E}^=}\pi_j|[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}|+\sum_{j\in\mathcal{E}^>}\pi_j[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}-\\&\sum_{i\in\mathcal{J}^<}\pi_i[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\sum_{i\in\mathcal{J}^=}\pi_i\max\{0,-[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}\}-\\&\sum_{t\in\mathcal{S}^<}\pi_t[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\\&\sum_{t\in\mathcal{S}^=}\pi_t|[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}|+\\&\sum_{t\in\mathcal{S}^>}\pi_t[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}.\end{aligned}$$

利用子问题$Q(\boldsymbol{x},\boldsymbol{B})$的最优性条件:

$$\nabla f(\boldsymbol{x})+\boldsymbol{B}\boldsymbol{d}-\sum_{i=1}^{p}\lambda_i\nabla g_i(\boldsymbol{x})+\sum_{j=1}^{q}\lambda_j\nabla h_j(\boldsymbol{x})+\sum_{t=1}^{l}\lambda_t[G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})+H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})]=\boldsymbol{0}.$$

并在上述等式两边同时乘以$\boldsymbol{d}$得

$$\begin{aligned}P'(\boldsymbol{x},\pi;\boldsymbol{d})=&-\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{d}+\sum_{j\in\mathcal{E}^{<}}(-\lambda_j-\pi_j)[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\sum_{j\in\mathcal{E}^{=}}(-\lambda_j[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\pi_j|[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}|)+\\&\sum_{j\in\mathcal{E}^{>}}(-\lambda_j+\pi_j)[\nabla h_j(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\sum_{i\in\mathcal{J}^{<}}(\lambda_i-\pi_i)[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\\&\sum_{i\in\mathcal{J}^{=}}[\lambda_i[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\pi_i\max\{0,-[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}\}]+\sum_{i\in\mathcal{J}^{>}}\lambda_i[\nabla g_i(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\\&\sum_{t\in\mathcal{S}^{<}}(-\lambda_t-\pi_t)[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\\&\sum_{t\in\mathcal{S}^{=}}(-\lambda_t[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}+\pi_t|[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}|)+\\&\sum_{t\in\mathcal{S}^{>}}(-\lambda_t+\pi_t)[H_t(\boldsymbol{x})\nabla G_t(\boldsymbol{x})+G_t(\boldsymbol{x})\nabla H_t(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{d}.\end{aligned}$$

由于$\boldsymbol{d}$为子问题$Q(\boldsymbol{x},\boldsymbol{B})$的解, 根据题设和互补条件分析$P'(\boldsymbol{x},\pi;\boldsymbol{d})$的右端第二项开始的各级数或小于等于零或等于零, 进而有

$$P'(\boldsymbol{x},\pi;\boldsymbol{d})\leqslant-\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}\boldsymbol{d}<0.$$ □

下面将给出对应算法的全局收敛性结果.

定理 3 设$f(\boldsymbol{x})$和$g_i(\boldsymbol{x})(i\in\{1,2,\cdots,p\})$, $h_j(\boldsymbol{x})(j\in\{1,2\cdots,q\})$, $G_t(\boldsymbol{x})$, $H_t(\boldsymbol{x})(i\in\{1,2,\cdots,l\})$连续可微, 引理 1 成立, 且存在正常数$m$与$M$使对任意的$\boldsymbol{d}\in\mathbb{R}^n$, 有

$$m\|\boldsymbol{d}\|^2\leqslant\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}^k\boldsymbol{d}\leqslant M\|\boldsymbol{d}\|^2,\qquad\forall k\geqslant1.$$

则在 Armijo 步长规则下, 算法产生的点列或终止于原问题的 S-平稳点, 或其聚点是 S-平稳点.

证明 若存在指标k使$\boldsymbol{d}^k=\boldsymbol{0}$, 则由子问题$Q(\boldsymbol{x}^k,\boldsymbol{B}^k)$的最优性条件得

$$\begin{cases}\nabla f(\boldsymbol{x}^k)+\boldsymbol{B}^k\boldsymbol{d}^k-\sum\limits_{i\in\mathcal{I}_g^*}\lambda_i^k\nabla g_i(\boldsymbol{x}^k)+\sum\limits_{j\in\mathcal{I}_h^*}\lambda_j^k\nabla h_j(\boldsymbol{x}^k)+\\\qquad\sum\limits_{t=1}^{l}\lambda_t^k[G_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+H_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k)]=\boldsymbol{0},\qquad\lambda_i^k\geqslant0,\\g_i(\boldsymbol{x}^k)+[\nabla g_i(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}^k\geqslant0,\qquad i\in\{1,2,\cdots,p\},\\h_j(\boldsymbol{x}^k)+[\nabla h_j(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}^k=0,\qquad j\in\{1,2,\cdots,q\},\\\lambda_i^k(g_i(\boldsymbol{x}^k)+[\nabla g_i(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}^k)=0,\qquad i\in\{1,2,\cdots,p\},\\G_t(\boldsymbol{x}^k)H_t(\boldsymbol{x}^k)+[G_t(\boldsymbol{x}^k)\nabla H_t(\boldsymbol{x}^k)+H_t(\boldsymbol{x}^k)\nabla G_t(\boldsymbol{x}^k)]^{\mathrm{T}}\boldsymbol{d}^k=0,\qquad t\in\{1,2,\cdots,l\},\end{cases}\tag{8}$$

从而$\boldsymbol{x}^k$为原问题 (1) 的 KKT 点, 算法终止.

若对任意的k, 均有$\boldsymbol{d}^k\neq\boldsymbol{0}$, 则算法产生无穷点列$\{\boldsymbol{x}^k\}$. 设$\boldsymbol{x}^*$为其一聚点, 由 MPSC-LICQ 和引理 1 中$\{\boldsymbol{\lambda}^k\}$的有界性, 不妨设

$$\lim_{\substack{k\in N_0\\k\to\infty}}\boldsymbol{x}^k=\boldsymbol{x}^*,\ \lim_{\substack{k\in N_0\\k\to\infty}}\boldsymbol{B}^k=\boldsymbol{B}^*,\ \lim_{\substack{k\in N_0\\k\to\infty}}\boldsymbol{\lambda}^k=\boldsymbol{\lambda}^*.$$

再由式 (8), $\nabla f(\boldsymbol{x}),\nabla g_i(\boldsymbol{x}),\nabla h_j(\boldsymbol{x}),\nabla G_t(\boldsymbol{x})$和$\nabla H_t(\boldsymbol{x})$的连续以及 $\boldsymbol{B}^*$的正定性, 可知$\{\boldsymbol{d}^k\}_{k\in N_0}$收敛, 记极限为$\boldsymbol{d}^*$. 则$(\boldsymbol{d}^*,\boldsymbol{\lambda}^*)$是子问题$Q(\boldsymbol{x}^*,\boldsymbol{B}^*)$的 KKT 点对.

下面用反证法证明$\boldsymbol{d}^*=\boldsymbol{0}$. 若$\boldsymbol{d}^*\neq\boldsymbol{0}$, 由 Armijo 步长规则

$$P(\boldsymbol{x}^k+\alpha^k\boldsymbol{d}^k,\pi)\leqslant P(\boldsymbol{x}^k,\pi)+\eta\alpha^k[\nabla P(\boldsymbol{x}^k,\pi)]^{\mathrm{T}}\boldsymbol{d}^k$$

有

$$P(\boldsymbol{x}^*,\pi)\leqslant P(\boldsymbol{x}^k,\pi)+\eta\sum_{i=k}^{\infty}\alpha_i[\nabla P(\boldsymbol{x}^i,\pi)]^{\mathrm{T}}\boldsymbol{d}^i.$$

当对于充分大的$k\in\mathcal{N}_0$, 且根据定理 2 有

$$P(\boldsymbol{x}^*,\pi)\leqslant P(\boldsymbol{x}^*,\pi)+\eta\lim_{k\to\infty}\sum_{i=k}^{\infty}\alpha_i[\nabla P(\boldsymbol{x}^i,\pi)]^{\mathrm{T}}\boldsymbol{d}^i<P(\boldsymbol{x}^*,\pi).$$

此矛盾说明$\boldsymbol{d}^*=\boldsymbol{0}$, 从而$\boldsymbol{x}^*$为原问题 (1) 的 KKT 点. 由定理 1, S 平稳条件与 KKT 条件等价, 所以该点也为 S-平稳点.

4 数值实验

本文在数值实验过程中以 MATLAB 作为主要处理工具, 其电脑型号及配置为 Lenovo 小新 Pro14HU2021x, 第十一代智能英特尔酷睿 i5 处理器, 16 GB 运行内存. 同时基于文献 [1] 中存零约束的优化问题实例, 来分析 SQP 算法的数值实验结果.

4.1 MPSC 实例

例 1　考虑问题 (文献 [1] 中的例 4.1)

$$\begin{aligned}\min\ & (x_1-2)^2+(x_2-1)^2+(x_3-2)^2,\\ \text{s.t.}\ & x_1^2+x_2^2+x_3^2\leqslant 3,\\ & x_3\leqslant 1,\\ & (x_1-x_2^2)(x_2-x_1^2)=0,\end{aligned}$$

显然, $(1,1,1)$是全局最优解.

例 2　考虑问题 (文献 [1] 中的例 5.1)

$$\begin{aligned}\min\ & (x_1-1)^2+x_2^2,\\ \text{s.t.}\ & x_1\leqslant 0,\\ & -x_2\leqslant 0,\\ & x_1x_2=0,\end{aligned}$$

其有唯一的全局最小解$(0,0)$.

例 3　考虑问题 (文献 [1] 中的例 5.2)

$$\begin{aligned}\min\ & x_1+x_2^2,\\ \text{s.t.}\ & -x_1+x_2\leqslant 0,\\ & x_1x_2=0,\end{aligned}$$

其有唯一的全局最小解$(0,0)$.

例 4　考虑问题 (文献 [1] 中的例 5.3)

$$\begin{aligned}\min\ & x_1+x_2,\\ \text{s.t.}\ & x_1^2-x_2\leqslant 0,\\ & x_1x_2=0,\end{aligned}$$

其有唯一的全局最小解$(0,0)$.

对以上例子进行数值实验, 其中后 3 个例子取相同初始值. 由于每次运行时间不同, 我们对每个例子在同一初始值处分别运行 50 次, 求各项结果的平均值. 表 1 中每列表示意义如下: 第一列表示各算例, $\boldsymbol{x}^*$表示问题的精确解, $\boldsymbol{x}'$表示问题的近似解, S 表示成功率, N 表示平均迭代次数, T 表示平均运行时间.

表 1 中结果显示, 通过 SQP 算法每次都能成功并较快地获得问题的精确解或近似解, 其中同一目标函数对应的约束函数越复杂, 处理该问题的迭代次数和运行时间会相应增多或变长. 这一结果说明采用 SQP 算法来处理 MPSC 是合理可行的.

4.2 投资组合

随后, 我们基于文献 [14] 中的投资组合优化实例对应的存零约束问题, 来分析 SQP 算法的数值实验结

果. 例子如下:

$$
\begin{aligned}
\min\quad & \boldsymbol{x}^{\mathrm{T}}\boldsymbol{Q}\boldsymbol{x},\\
\text{s.t.}\quad & \boldsymbol{e}^{\mathrm{T}}\boldsymbol{x}-1=0,\\
& \rho-\boldsymbol{\mu}^{\mathrm{T}}\boldsymbol{x}\leqslant 0,\\
& \boldsymbol{x}-\boldsymbol{u}\leqslant\boldsymbol{0},\ -\boldsymbol{y}\leqslant\boldsymbol{0},\\
& x_i(x_i-l_i-y_i)=0,\qquad i=1,2,\cdots,n,
\end{aligned}
$$

其中的变量$\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$, 随机生成$\boldsymbol{Q}\in\mathbb{R}^{n\times n}$, $\boldsymbol{\mu}\in\mathbb{R}^n$, $\boldsymbol{l},\boldsymbol{u}\in\mathbb{R}^n$以及$\rho\in\mathbb{R}$, $\boldsymbol{e}$ 表示所有元素均为 1 的列向量. 变量与参数的实际经济含义为: $\boldsymbol{x},\boldsymbol{y}$分别表示所选择投资股票的份额 (比例) 和松弛变量, $\boldsymbol{Q}$ 为方差-协方差矩阵, $\boldsymbol{\mu}$ 为资产的预期收益平均值, $\boldsymbol{l},\boldsymbol{u}$ 为资产的最小最大买入阈值, ρ为期望的回报水平. 以上参数取值为$\rho=0.002$, $l_i=-10, u_i=20(i=1,2,\cdots,n)$, $\boldsymbol{\mu}$是随机生成的列向量, 初始点$\boldsymbol{x}_0$是从第一行开始每$1/(2n)$个元素分别取值为$-1/(2n), 5/(2n)$, 0 和 0 的列向量, $\boldsymbol{Q}$为随机生成的对称正定矩阵. 表 2 中第一列表示样本量, 即$\boldsymbol{x}$的维度, 其余符号如表 1 所述.

表 1 MPSCs 数值结果

Table 1 MPSCs numerical results

example	$\boldsymbol{x}^*$	$\boldsymbol{x}'$	S	N	T
1	(1,1,1)	(1.0,1.0,1.0)	1	19	0.0070
2	(0,0)	(0.0,0.0)	1	4	0.0043
3	(0,0)	(−0.0,−0.0)	1	17	0.0065
4	(0,0)	(−0.0,0.0)	1	36	0.0082

表 2 投资组合数值结果

Table 2 Portfolio numerical results

n	S	N	T
50	1	24	0.1685
100	1	28	0.6690
200	1	36	3.3620
500	0.96	49	61.2914
800	0.90	59	1387.5846
1000	0.86	63	2336.6510
1200	0.82	69	4425.9890

结果显示, 随着变量维度的不断增加, 相应的迭代次数和运行时间逐渐增加, 与此同时, SQP 算法处理投资组合问题时的成功率虽然逐渐降低, 但依旧能较大概率求解该问题, 这一结果说明 SQP 算法处理含高维变量的问题仍然可行.

5 结论

本文研究了存零约束优化问题的求解方法. 首先基于 MPSC 相应的定义, 讨论了 MPSC 中的 S-平稳点和 KKT 点之间的关系, 然后利用 SQP 算法设计出该类问题对应的算法. 通过构造合理的假设条件, 证明了算法生成的乘子的适定性, 进而得到了算法的全局收敛性质. 最后的数值实例表明提出的算法对求解存零约束优化问题的可行性, 且可推广应用到具有大样本的实际经济问题中. 本文应用的方法对于求解存零约束优化问题提供了一个新的方式.

致谢 本文作者衷心感谢重庆工商大学科研项目 (ZDPTTD201908) 和重庆工商大学研究生创新型科研项目 (yjscxx2022-112-184) 对本文的资助.

参考文献(References):

[1] MEHLITZ P. Stationarity conditions and constraint qualifications for mathematical programs with switching constraints[J]. *Mathematical Programming*, 2020, **181**(1): 149-186.

[2] MEHLITZ P. On the linear independence constraint qualification in disjunctive programming[J]. *Optimization*, 2020, **69**(10): 2241-2277.

[3] LUO Z Q, PANG J S, RALPH D. *Mathematical Programs With Equilibrium Constraints*[M]. Cambridge: Cambridge University Press, 1996.

[4] ACHTZIGER W, KANZOW C. Mathematical programs with vanishing constraints: optimality conditions and constraint qualifications[J]. *Mathematical Programming*, 2008, **114**(1): 69-99.

[5] LIANG Y C, YE J J. New optimality conditions and exact penalty for mathematical programs with switching constraints[J]. *Journal of Optimization Theory and Applications*, 2021, **190**: 1-31.

[6] FLETCHER R, LEYFFER S, RALPH D, et al. Local convergence of SQP methods for mathematical programs with equilibrium constraints[J]. *SIAM Journal on Optimization*, 2006, **17**(1): 259-286.

[7] 朱志斌, 罗志军, 曾吉文. 互补约束均衡问题一个新的磨光技术[J]. 应用数学和力学, 2007, **28**(10): 1253-1260. (ZHU Zhibin, LUO Zhijun, ZENG Jiwen. Complementary constraint equalization problem a new polishing technique[J]. *Applied Mathematics and Mechanics*, 2007, **28**(10): 1253-1260.(in Chinese))

[8] ITO K, KUNISCH K. Augmented Lagrangian-SQP methods for nonlinear optimal constol problems of tracking type[J]. *SIAM Journal on Control and Optimization*, 1996, **34**(3): 874-891.

[9] YU Y H, GAO L. Nonmonotone line search algorithm for constrained minimax problems[J]. *Journal of Optimization Theory and Application*, 2002, **115**: 419-446.

[10] LING C, QI L Q, ZHOU G L, et al. Global convergence of a robust smoothing SQP method for semi-infinite programming[J]. *Journal of Optimization Theory and Application*, 2006, **129**: 147-164.

[11] WRIGHT J. Modifying SQP for degenerate problems[J]. *SIAM Journal on Optimization*, 2002, **13**(2): 470-497.

[12] 朱志斌, 简金宝, 张聪. 非线性互补约束均衡问题的一个SQP算法[J]. 应用数学和力学, 2009, **30**(5): 613-622. (ZHU Zhibin, JIAN Jinbao, ZHANG Cong. An SQP algorithm for mathematical programs with nonlinear complementarity constraints[J]. *Applied Mathematics and Mechanics*, 2009, **30**(5): 613-622.(in Chinese))

[13] 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, 2019. (WANG Yiju, XIU Naihua. *Theory and Method of Nonlinear Optimization*[M]. Beijing: Science Press, 2019. (in Chinese))

[14] FRANGIONI A, GRNTILE C. SDP diagonalizations and perspective cuts for a class of nonseparable MIQP[J]. *Operations Research Letters*, 2007, **35**(2): 181-185.