

存零约束优化问题的对偶问题

罗美铃¹, 李高西¹, 吴春²

(1. 重庆工商大学数学与统计学院, 重庆 400067)

(2. 重庆师范大学数学科学学院, 重庆 401331)

摘要: 本文研究了近年提出的一类新优化问题存零约束优化问题, 因存零约束的存在, 使得求解最优解较困难. 因此, 本文针对存零约束优化问题, 利用对偶理论提出了问题的 Wolfe 型对偶模型. 在凸性和严格凸性假设下, 获得了 Wolfe 对偶的弱、强、逆、限制逆和严格逆对偶结果. 并进行了实例论证.

关键词: 非线性规划; 存零约束; 对偶问题; 广义凸性

MR(2010) 主题分类号: 90C30; 90C25 中图分类号: O221.2

文献标识码: A 文章编号: 0255-7797(2023)04-0347-09

1 引言

本文考虑下述非线性规划问题:

$$\begin{aligned} \min & f(x) \\ \text{s.t.} & g_i(x) \geq 0, \quad i = 1, \dots, p, \\ & h_j(x) = 0, \quad j = 1, \dots, q, \\ & G_t(x)H_t(x) = 0, \quad t = 1, \dots, l, \end{aligned} \quad (1.1)$$

其中, 函数 $f(x), g_1, \dots, g_p, h_1, \dots, h_q, G_1, \dots, G_l, H_1, \dots, H_l: \mathbb{R}^n \rightarrow \mathbb{R}$ 均连续可微. 在问题 (1.1) 任意可行点 x 处, 对任意固定的 $t, G_t(x)$ 和 $H_t(x)$ 至少有一个为零, 即两个函数的乘积为零, 我们称这样的等式约束为存零约束, 问题 (1.1) 则为“存零约束”优化问题, 简记为 MPSC. 该模型首先由 Mehrlitz^[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 (1.1), 因此有必要讨论问题 (1.1) 的相关理论. 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 (1.1) 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、消失约束的优化问题密切相关, 见文献 [2]. Mehrlitz^[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, 然后提出了 MPSC 特定的约束规范, 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

*收稿日期: 2022-09-05 接收日期: 2022-11-08

基金项目: 国家自然科学基金项目 (11901068); 重庆市自然科学基金 (cstc2019jcyj-msxmX0390); 重庆工商大学研究生创新型科研项目 (yjscxx2022-112-184); 重庆工商大学科研项目 (ZDPTTD201908).

作者简介: 罗美铃 (1998-), 女, 重庆万州, 研究生, 主要研究方向: 最优化理论及算法.

通讯作者: 吴春 (1976-), 男, 重庆巴南, 副教授, 主要研究方向: 值分布及偏微分方程理论.

E-mail: wuchun@cqnu.edu.cn.

范.Yang 等人^[3] 得出了满足增广拉格朗日问题二阶必要最优性条件的点序列的极限点是原互补约束优化问题的强平稳点.Singh 等人^[4] 针对均衡约束多目标优化问题建立了 Wolfe 型对偶和 Mond-Weir 型对偶模型, 并在伪凸性和拟凸性假设下建立了弱对偶和强对偶定理.

对偶性在研究优化问题时非常重要. 经典的 Wolfe 对偶是由学者 Wolfe 于 1961 年引入. 目前, 研究问题的对偶模型已经被用作解决各种优化问题的工具, 如半无限规划问题^[5], 变分不等式问题^[6], 分数式编程问题^[7], 凸优化问题^[8] 等. 陈世国^[9] 对不可微多目标规划在伪不变凸和拟不变凸时给出了弱有效解和有效解的 Mond-Weir 型对偶理论. 之后, 陈世国等人^[10] 针对含有锥约束的多目标变分问题, 得到其广义对称对偶性.Mishra^[11] 在一些假设下, 建立了消失约束优化问题与对偶问题之间的弱、强、逆、限制逆和严格逆对偶性结果. 本文基于文献 [11], 拟研究 MPSC 的 Wolfe 型对偶问题.

本文余下组成部分为: 第 2 节给出本文研究所需的基本定义. 第 3 节提出 Wolfe 型对偶模型, 在凸性和严格凸性假设下, 获得原问题的 Wolfe 对偶问题的弱、强、逆、限制逆和严格逆对偶结果.

2 预备知识

为便于后续研究, 定义指标集:

$$\begin{aligned}\mathcal{I}_g &= \{i \in \{1, \dots, p\} : g_i(x) = 0\}, \mathcal{I}_h = \{1, \dots, q\}, \\ \mathcal{I}_G &= \{t \in \{1, \dots, l\} : G_t(x) = 0 \wedge H_t(x) \neq 0\}, \\ \mathcal{I}_H &= \{t \in \{1, \dots, l\} : G_t(x) \neq 0 \wedge H_t(x) = 0\}, \\ \mathcal{I}_{GH} &= \{t \in \{1, \dots, l\} : G_t(x) = 0 \wedge H_t(x) = 0\}.\end{aligned}$$

其中, $\{\mathcal{I}_G, \mathcal{I}_H, \mathcal{I}_{GH}\}$ 是 $\{1, \dots, l\}$ 上互不相交的划分. 定义 MPSC 的拉格朗日函数为

$$L(x, \lambda, \lambda_e, \mu, \nu) = f(x) - \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \lambda_j h_j(x) + \sum_{t=1}^l \mu_t G_t(x) + \sum_{t=1}^l \nu_t H_t(x).$$

下面将给出一些关于 MPSC 的平稳性和约束规范的定义.

定义 2.1^[1] 设 $x \in \mathbb{R}^n$ 为 MPSC 的可行点, 如果存在乘子满足

$$\begin{aligned}0 &= \nabla f(x) - \sum_{i \in \mathcal{I}_g} \lambda_i \nabla g_i(x) + \sum_{j \in \mathcal{I}_h} \lambda_j \nabla h_j(x) + \sum_{t=1}^l [\mu_t \nabla G_t(x) + \nu_t \nabla H_t(x)], \\ \forall i \in \mathcal{I}_g : \lambda_i &\geq 0, \forall t \in \mathcal{I}_G : \nu_t = 0, \quad \forall t \in \mathcal{I}_H : \mu_t = 0, \forall t \in \mathcal{I}_{GH} : \mu_t \nu_t = 0,\end{aligned}\quad (2.1)$$

则称 x 为 Mordukhovich 平稳点, 简记为 M 平稳点.

定义 2.2^[1] 设 $x \in \mathbb{R}^n$ 为 MPSC 的可行点, 如果在 x 点处约束函数的梯度

$$\begin{aligned}\nabla g_i(x), i \in \mathcal{I}_g, & \quad \nabla h_j(x), j \in \mathcal{I}_h, \\ \nabla G_t(x), t \in \mathcal{I}_G \cup \mathcal{I}_{GH}, & \quad \nabla H_t(x), t \in \mathcal{I}_H \cup \mathcal{I}_{GH},\end{aligned}$$

线性独立, 则称 MPSC-LICQ 在 x 点处成立.

在建立对偶理论时, 凸性和广义凸性概念具有重要作用, 因此, 给出定义如下:

定义 2.3^[12] 设 $S \subseteq \mathbb{R}^n$ 为任一非空集合, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 连续可微. 则称 f 在 $x^* \in S$ 点处是凸的, 当且仅当对于任意的 $x \in S$, 有

$$f(x) - f(x^*) \geq \langle \nabla f(x^*), x - x^* \rangle.$$

若 $x \neq x^*$ 时, 以上定义的严格不等式成立, 则称 f 在 $x^* \in S$ 点处是严格凸的.

定义 2.4^[12] 设 $S \subseteq \mathbb{R}^n$ 为任一非空集合, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 连续可微. 则称 f 在 $x^* \in S$ 点处是拟凸的, 当且仅当对于任意的 $x \in S$, 有

$$f(x) \leq f(x^*) \Rightarrow \langle \nabla f(x^*), x - x^* \rangle \leq 0.$$

定义 2.5^[12] 设 $S \subseteq \mathbb{R}^n$ 为任一非空集合, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 连续可微. 则称 f 在 $x^* \in S$ 点处是伪凸的, 当且仅当对于任意的 $x \in S$, 有

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0 \Rightarrow f(x) \geq f(x^*).$$

定义 2.6^[12] 设 $S \subseteq \mathbb{R}^n$ 为任一非空集合, $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 连续可微. 则称 f 在 $x^* \in S$ 点处是严格伪凸的, 当且仅当对于任意的 $x \in S$ 且 $x \neq x^*$, 有

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0 \Rightarrow f(x) > f(x^*).$$

3 Wolfe 型对偶

本节根据可行点 $x \in X$ 定义 MPSC 的 Wolfe 型对偶问题 (WD) 如下:

$$\begin{aligned} \max_{y, \lambda_l, \lambda_e, \mu, \nu} \quad & L(y, \lambda_l, \lambda_e, \mu, \nu) \quad \text{s.t.} \quad \nabla L(y, \lambda_l, \lambda_e, \mu, \nu) = 0, \\ & \lambda_i \geq 0, i \notin \mathcal{I}_g, \quad \mu_t = 0, t \in \mathcal{I}_H, \quad \nu_t = 0, t \in \mathcal{I}_G, \quad \mu_t \nu_t = 0, t \in \mathcal{I}_{GH}. \end{aligned}$$

定义 $\text{WD}(x)$ 所有可行点构成的集合为

$$\Omega(x) := \left\{ (y, \lambda_l, \lambda_e, \mu, \nu) : \begin{array}{l} \nabla L(y, \lambda_l, \lambda_e, \mu, \nu) = 0, \\ \lambda_i \geq 0, i \notin \mathcal{I}_g, \\ \mu_t = 0, t \in \mathcal{I}_H, \\ \nu_t = 0, t \in \mathcal{I}_G, \\ \mu_t \nu_t = 0, t \in \mathcal{I}_{GH}. \end{array} \right\}.$$

此时, 定义集合 $\Omega(x)$ 在 \mathbb{R}^n 上的投影如下

$$P\Omega(x) := \left\{ y \in \mathbb{R}^n : (y, \lambda_l, \lambda_e, \mu, \nu) \in \Omega(x) \right\}.$$

为区分 MPSC 与 WD, 本节考虑由 WD 定义的另一对偶问题

$$\max_{y, \lambda_l, \lambda_e, \mu, \nu} \quad L(y, \lambda_l, \lambda_e, \mu, \nu) \quad \text{s.t.} \quad (y, \lambda_l, \lambda_e, \mu, \nu) \in \bigcap_{x \in X} \Omega(x).$$

其中定义 $\Omega = \bigcap_{x \in X} \Omega(x)$ 是 WD 所有可行点的集合, $P\Omega$ 是集合 Ω 在 \mathbb{R}^n 上的投影. 为方便起见, 定义以下指标集:

$$\begin{aligned} \mathcal{I}_g^+ &= \{i \in \{1, \dots, p\} : \lambda_i > 0\}, \\ \mathcal{I}_h^+ &= \{j \in \{1, \dots, q\} : \lambda_j > 0\}, \mathcal{I}_h^- = \{j \in \{1, \dots, q\} : \lambda_j < 0\}, \\ \mathcal{I}_G^+ &= \{t \in \{1, \dots, l\} : \mu_t > 0\}, \mathcal{I}_G^- = \{t \in \{1, \dots, l\} : \mu_t < 0\}, \\ \mathcal{I}_H^+ &= \{t \in \{1, \dots, l\} : \nu_t > 0\}, \mathcal{I}_H^- = \{t \in \{1, \dots, l\} : \nu_t < 0\}. \end{aligned} \quad (3.1)$$

定理 3.1(弱对偶) 设 $x \in X, (y, \lambda_i, \lambda_e, \mu, \nu) \in \Omega$. 如果下列条件中有一个成立:

- (i) $L(\cdot, \lambda_i, \lambda_e, \mu, \nu)$ 在 $y \in X \cup P\Omega$ 处为凸.
(ii) $f, -g_i (i \in \mathcal{I}_g^+), h_j (j \in \mathcal{I}_h^+), -h_j (j \in \mathcal{I}_h^-), G_t (t \in \mathcal{I}_G^+), -G_t (t \in \mathcal{I}_G^-), H_t (t \in \mathcal{I}_H^+), -H_t (t \in \mathcal{I}_H^-)$ 在 $y \in X \cup P\Omega$ 处为凸.

则 $f(x) \geq L(y, \lambda_i, \lambda_e, \mu, \nu)$.

证 (i) 反证法. 假设

$$f(x) < f(y) - \sum_{i=1}^p \lambda_i g_i(y) + \sum_{j=1}^q \lambda_j h_j(y) + \sum_{t=1}^l \mu_t G_t(y) + \sum_{t=1}^l \nu_t H_t(y). \quad (3.2)$$

因为 x 和 $(y, \lambda_i, \lambda_e, \mu, \nu)$ 分别为 MPSC 和 WD 的可行点, 显然有

$$\begin{aligned} -g_i(x) < 0, \lambda_i \geq 0, \forall i \notin \mathcal{I}_g, \quad g_i(x) = 0, \lambda_i \in \mathbb{R}, \forall i \in \mathcal{I}_g, \quad h_j(x) = 0, \lambda_j \in \mathbb{R}, \forall j \in \mathcal{I}_h, \\ G_t(x) = 0, \mu_t \in \mathbb{R}, \forall t \in \mathcal{I}_G, \quad G_t(x) \neq 0, \mu_t = 0, \forall t \in \mathcal{I}_H, \\ H_t(x) = 0, \nu_t \in \mathbb{R}, \forall t \in \mathcal{I}_H, \quad H_t(x) \neq 0, \nu_t = 0, \forall t \in \mathcal{I}_G, \\ G_t(x) = H_t(x) = 0, \mu_t = 0 \vee \nu_t = 0, \forall t \in \mathcal{I}_{GH}. \end{aligned}$$

则有

$$- \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \lambda_j h_j(x) + \sum_{t=1}^l \mu_t G_t(x) + \sum_{t=1}^l \nu_t H_t(x) \leq 0. \quad (3.3)$$

将 (3.2) 和 (3.3) 相加, 有

$$\begin{aligned} f(x) - \sum_{i=1}^p \lambda_i g_i(x) + \sum_{j=1}^q \lambda_j h_j(x) + \sum_{t=1}^l \mu_t G_t(x) + \sum_{t=1}^l \nu_t H_t(x) \\ < f(y) - \sum_{i=1}^p \lambda_i g_i(y) + \sum_{j=1}^q \lambda_j h_j(y) + \sum_{t=1}^l \mu_t G_t(y) + \sum_{t=1}^l \nu_t H_t(y) \end{aligned}$$

即是

$$L(x, \lambda_i, \lambda_e, \mu, \nu) < L(y, \lambda_i, \lambda_e, \mu, \nu). \quad (3.4)$$

根据 $L(\cdot, \lambda_i, \lambda_e, \mu, \nu)$ 的凸性可得

$$L(y, \lambda_i, \lambda_e, \mu, \nu) + \langle \nabla L(y, \lambda_i, \lambda_e, \mu, \nu), x - y \rangle \leq L(x, \lambda_i, \lambda_e, \mu, \nu). \quad (3.5)$$

由于 $(y, \lambda_l, \lambda_e, \mu, \nu)$ 是 WD 的可行点, 即 $\nabla L(y, \lambda_l, \lambda_e, \mu, \nu) = 0$, 则 (3.5) 可化简为 $L(y, \lambda_l, \lambda_e, \mu, \nu) < L(x, \lambda_l, \lambda_e, \mu, \nu)$. 这与 (3.4) 矛盾, 因此结论得证.

(ii) 根据定理中的第二个题设条件, 易得 $L(\cdot, \lambda_l, \lambda_e, \mu, \nu)$ 在 $y \in X \cup P\Omega$ 处为凸, 再结合 (i) 的证明, 即可得证.

定理 3.2 (强对偶) 令 $x^* \in X$ 是 MPSC 的局部最小解, MPSC-LICQ 在 x^* 处成立. 则存在拉格朗日乘子 $(\lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 使得 $(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的可行点, 且

$$-\sum_{i=1}^p \lambda_i^* g_i(x^*) + \sum_{j=1}^q \lambda_j^* h_j(x^*) + \sum_{t=1}^l \mu_t^* G_t(x^*) + \sum_{t=1}^l \nu_t^* H_t(x^*) = 0. \quad (3.6)$$

另外, 如果下列条件中有一个成立:

(i) $L(\cdot, \lambda_e, \lambda_l, \mu, \nu)$ 在 $y \in X \cup P\Omega$ 处为凸.

(ii) $f, -g_i (i \in \mathcal{I}_g^+(x^*)), h_j (j \in \mathcal{I}_h^+(x^*)), -h_j (j \in \mathcal{I}_h^-(x^*)), G_t (t \in \mathcal{I}_G^+(x^*)), -G_t (t \in \mathcal{I}_G^-(x^*)), H_t (t \in \mathcal{I}_H^+(x^*)), -H_t (t \in \mathcal{I}_H^-(x^*))$ 在 $y \in X \cup P\Omega$ 处为凸.

则 $(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的全局最大解且 $f(x^*) = L(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$.

证 由于 x^* 是 MPSC 的局部最小解且 MPSC-LICQ 在该点处成立, 再根据定义 2.1, 对于点 x^* , 存在拉格朗日乘子 $(\lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 使得 (2.1) 成立, 因此 $(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的可行点. 再根据 (2.1) 的后四个条件和 (3.1) 易得 (3.6). 由定理 3.1, 有

$$f(x^*) \geq L(y, \lambda_l, \lambda_e, \mu, \nu), \forall (y, \lambda_l, \lambda_e, \mu, \nu) \in \Omega(x^*). \quad (3.7)$$

进一步将 (3.6) 与 (3.7) 相加有 $L(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*) \geq L(y, \lambda_l, \lambda_e, \mu, \nu), \forall (y, \lambda_l, \lambda_e, \mu, \nu) \in \Omega(x^*)$. 即是 $(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 为 WD 的全局最大解. 同时, MPSC 的局部最优值等价于 WD 的全局最优值.

定理 3.3 (逆对偶) 令 x 是 MPSC 的非空可行集, $(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的可行点并有以下成立

$$\begin{aligned} -\lambda_i^* g_i(y^*) &\geq 0, i \in \{1, \dots, p\}, & \lambda_j^* h_j(y^*) &= 0, j \in \{1, \dots, q\}, \\ \mu_t^* G_t(y^*) &= 0, t \in \{1, \dots, l\}, & \nu_t^* H_t(y^*) &= 0, t \in \{1, \dots, l\}. \end{aligned} \quad (3.8)$$

另外, 如果下列条件中有一个成立:

(i) $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为凸.

(ii) $f, -g_i (i \in \mathcal{I}_g^+), h_j (j \in \mathcal{I}_h^+), -h_j (j \in \mathcal{I}_h^-), G_t (t \in \mathcal{I}_G^+), -G_t (t \in \mathcal{I}_G^-), H_t (t \in \mathcal{I}_H^+), -H_t (t \in \mathcal{I}_H^-)$ 在 $y^* \in X \cup P\Omega$ 处为凸. 则 y^* 为 MPSC 的全局最小解.

证 反证法. 假设 y^* 不是 MPSC 的全局最小解, 即存在 $\hat{x} \in X$ 使得

$$f(\hat{x}) < f(y^*). \quad (3.9)$$

(i) 因为 \hat{x} 和 $(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 分别为 MPSC 和 WD 的可行点, 由定理中的假设易得

$$\begin{aligned} -g_i(\hat{x}) &< 0, \lambda_i^* \geq 0, \forall i \notin \mathcal{I}_g(\hat{x}), & g_i(\hat{x}) &= 0, \lambda_i^* \in \mathbb{R}, \forall i \in \mathcal{I}_g(\hat{x}), \\ h_j(\hat{x}) &= 0, \lambda_j^* \in \mathbb{R}, \forall j \in \mathcal{I}_h(\hat{x}), & G_t(\hat{x}) &= 0, \mu_t^* \in \mathbb{R}, \forall t \in \mathcal{I}_G(\hat{x}), \\ G_t(\hat{x}) &\neq 0, \mu_t^* = 0, \forall t \in \mathcal{I}_H(\hat{x}), & H_t(\hat{x}) &= 0, \nu_t^* \in \mathbb{R}, \forall t \in \mathcal{I}_H(\hat{x}), \\ H_t(\hat{x}) &\neq 0, \nu_t^* = 0, \forall t \in \mathcal{I}_G(\hat{x}), & G_t(\hat{x}) &= H_t(\hat{x}) = 0, \mu_t^* = 0 \vee \nu_t^* = 0, \forall t \in \mathcal{I}_{GH}(\hat{x}). \end{aligned}$$

进一步推得

$$\begin{aligned} & -\sum_{i=1}^p \lambda_i^* g_i(\hat{x}) + \sum_{j=1}^q \lambda_j^* h_j(\hat{x}) + \sum_{t=1}^l \mu_t^* G_t(\hat{x}) + \sum_{t=1}^l \nu_t^* H_t(\hat{x}) \\ \leq & -\sum_{i=1}^p \lambda_i^* g_i(y^*) + \sum_{j=1}^q \lambda_j^* h_j(y^*) + \sum_{t=1}^l \mu_t^* G_t(y^*) + \sum_{t=1}^l \nu_t^* H_t(y^*). \end{aligned} \quad (3.10)$$

将 (3.9) 和 (3.10) 相加有 $L(\hat{x}, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*) < L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$. 根据 $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y \in X \cup P\Omega$ 处为凸, 可得

$$\langle \nabla L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*), \hat{x} - y^* \rangle < 0,$$

这与约束 $\nabla L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*) = 0$ 矛盾, 因此结论得证.

(ii) 根据定理中的第二个题设条件, 易得 $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为凸, 再结合 (i) 的证明, 即可得证.

定理 3.4(限制逆对偶) 令 $x^* \in X$ 是 MPSC 的任一可行点, $(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的可行点且满足 $f(x^*) = L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 另外, 如果下列条件中有一个成立:

(i) $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为凸.

(ii) $f, -g_i (i \in \mathcal{I}_g^+(x^*)), h_j (j \in \mathcal{I}_h^+(x^*)), -h_j (j \in \mathcal{I}_h^-(x^*)), G_t (t \in \mathcal{I}_G^+(x^*)), -G_t (t \in \mathcal{I}_G^-(x^*)), H_t (t \in \mathcal{I}_H^+(x^*)), -H_t (t \in \mathcal{I}_H^-(x^*))$ 在 $y^* \in X \cup P\Omega$ 处为凸.

则 x^* 为 MPSC 的全局最小解.

证 反之, 假设 x^* 不是 MPSC 的全局最小解, 则存在 \hat{x} 使得 $f(\hat{x}) < f(x^*)$. 结合定理中的假设有 $f(\hat{x}) < L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$, 这与定理 3.1 的结论矛盾, 因此得证定理 3.4.

定理 3.5(严格逆对偶) 令 $x^* \in X$ 是 MPSC 的局部最小解, 并使得 MPSC-LICQ 在该点处成立. 此外, 假设定理 3.2 的假设条件成立, $(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 是 WD 的全局最大解. 如果下列条件中有一个成立:

(i) $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为严格凸.

(ii) f 是严格凸的, 且 $-g_i (i \in \mathcal{I}_g^+(x^*)), h_j (j \in \mathcal{I}_h^+(x^*)), -h_j (j \in \mathcal{I}_h^-(x^*)), G_t (t \in \mathcal{I}_G^+(x^*)), -G_t (t \in \mathcal{I}_G^-(x^*)), H_t (t \in \mathcal{I}_H^+(x^*)), -H_t (t \in \mathcal{I}_H^-(x^*))$ 在 $y^* \in X \cup P\Omega$ 处为凸. 则 $x^* = y^*$.

证 (i) 同样用反证法. 假设 $x^* \neq y^*$. 则由定理 3.2, 存在拉格朗日乘子 $(\bar{\lambda}_l, \bar{\lambda}_e, \bar{\mu}, \bar{\nu})$ 使得 $(x^*, \bar{\lambda}_l, \bar{\lambda}_e, \bar{\mu}, \bar{\nu})$ 为 WD 的全局最大解, 因此

$$f(x^*) = L(x^*, \bar{\lambda}_l, \bar{\lambda}_e, \bar{\mu}, \bar{\nu}) = L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*). \quad (3.11)$$

根据 x^* 和 $(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 分别是 MPSC 和 WD 的可行点, 容易得到

$$\begin{aligned} & -g_i(x^*) < 0, \lambda_i^* \geq 0, \forall i \notin \mathcal{I}_g(x^*), \quad g_i(x^*) = 0, \lambda_i^* \in \mathbb{R}, \forall i \in \mathcal{I}_g(x^*), \\ & h_j(x^*) = 0, \lambda_j^* \in \mathbb{R}, \forall j \in \mathcal{I}_h(x^*), \quad G_t(x^*) = 0, \mu_t^* \in \mathbb{R}, \forall t \in \mathcal{I}_G(x^*), \\ & G_t(x^*) \neq 0, \mu_t^* = 0, \forall t \in \mathcal{I}_H(x^*), \quad H_t(x^*) = 0, \nu_t^* \in \mathbb{R}, \forall t \in \mathcal{I}_H(x^*), \\ & H_t(x^*) \neq 0, \nu_t^* = 0, \forall t \in \mathcal{I}_G(x^*), \quad G_t(x^*) = H_t(x^*) = 0, \mu_t^* = 0 \vee \nu_t^* = 0, \forall t \in \mathcal{I}_{GH}(x^*). \end{aligned}$$

从而有

$$-\sum_{i=1}^p \lambda_i^* g_i(x^*) + \sum_{j=1}^q \lambda_j^* h_j(x^*) + \sum_{t=1}^l \mu_t^* G_t(x^*) + \sum_{t=1}^l \nu_t^* H_t(x^*) \leq 0. \quad (3.12)$$

将 (3.11) 与 (3.12) 相加有

$$L(x^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*) \leq L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*).$$

结合以上不等式以及 $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为严格凸, 推知

$$\langle \nabla L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*), x^* - y^* \rangle < 0.$$

这与约束 $\nabla L(y^*, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*) = 0$ 矛盾, 因此结论得证.

(ii) 根据定理中的第二个题设条件, 易得 $L(\cdot, \lambda_l^*, \lambda_e^*, \mu^*, \nu^*)$ 在 $y^* \in X \cup P\Omega$ 处为严格凸, 再结合 (i) 的证明, 即可得证.

接下来, 利用下面的例子来验证 Wolfe 型对偶和相关定理的有效性.

例 3.1 考虑含存零约束的优化问题

$$\begin{aligned} \min_x f(x) &: = (x_1 - 1)^2 + (x_2 + 3)^2 + x_3^6 \\ \text{s.t. } G_1(x) &: = x_1^2 - 2x_1 + x_3 + 1, \\ H_1(x) &: = x_2^2 + 6x_2 - x_3 + 9, \\ G_1(x)H_1(x) &= 0, \end{aligned} \quad (3.13)$$

其可行点的集合 X 为

$$X := \{(x_1, x_2) \in \mathbb{R}^2 \mid (x_1^2 - 2x_1 + x_3 + 1)(x_2^2 + 6x_2 - x_3 + 9) = 0\}.$$

(3.13) 的对偶形式为

$$\begin{aligned} \min_{y, \mu_1, \nu_1} L(y, \mu_1, \nu_1) &= (y_1 - 1)^2 + (y_2 + 3)^2 + y_3^6 + \mu_1(y_1^2 - 2y_1 + y_3 + 1) \\ &\quad + \nu_1(y_2^2 + 6y_2 - y_3 + 9) \\ \text{s.t. } 2(y_1 - 1)(1 + \mu_1) &= 0, \quad 2(y_2 + 3)(1 + \nu_1) = 0, \quad 6y_3^5 + \mu_1 - \nu_1 = 0, \\ \mu_1 = 0, 1 \in \mathcal{I}_H, \quad \nu_1 = 0, 1 \in \mathcal{I}_G, \quad \mu_1\nu_1 = 0, 1 \in \mathcal{I}_{GH}, \end{aligned} \quad (3.14)$$

(1) 令 $x^* = (1, -3, 0)^T \in X$, $(y_1^*, y_2^*, y_3^*, \mu_1^*, \nu_1^*) = (1, -3, 0, 0, 0) \in \Omega(x^*)$, 有

$$f(x^*) = L(y^*, \mu_1^*, \nu_1^*).$$

易验证定理 3.4 的所有条件均成立, 因此 x^* 为 (3.13) 的全局最小解. 故定理 3.4 得以验证.

(2) 由 (3.14) 的前三个等式约束可得, 当 $y_1 = 1, y_2 = -3, y_3 = \sqrt[5]{\frac{\nu_1 - \mu_1}{6}}$. 结合 (3.13) 中的目标函数, 易得到

$$L(y, \mu_1, \nu_1) = \left(\sqrt[5]{\frac{\nu_1 - \mu_1}{6}}\right)^6 - \frac{(\sqrt[5]{\nu_1 - \mu_1})^6}{\sqrt[5]{6}} \leq (x_1 - 1)^2 + (x_2 + 3)^2 + x_3^6 = f(x).$$

当 $\mu_1 = -1, y_1 \in \mathbb{R}, \nu_1 = -1, y_2 \in \mathbb{R}, y_3 = 0$ 时, 同样可得到

$$L(y, \mu_1, \nu_1) = 0 \leq (x_1 - 1)^2 + (x_2 + 3)^2 + x_3^6 = f(x).$$

当 $\mu_1 = -1, y_1 \in \mathbb{R}, y_2 = -3, y_3 = \sqrt[5]{\frac{\nu_1 + 1}{6}}$ 时,

$$L(y, \mu_1, \nu_1) = \left(\sqrt[5]{\frac{\nu_1 + 1}{6}}\right)^6 - \frac{(\sqrt[5]{\nu_1 + 1})^6}{\sqrt[5]{6}} \leq (x_1 - 1)^2 + (x_2 + 3)^2 + x_3^6 = f(x).$$

当 $\nu_1 = -3, y_2 \in \mathbb{R}, y_1 = 1, y_3 = \sqrt[5]{\frac{-\mu_1 - 1}{6}}$ 时,

$$L(y, \mu_1, \nu_1) = \left(\sqrt[5]{\frac{\mu_1 + 1}{6}}\right)^6 - \frac{(\sqrt[5]{\mu_1 + 1})^6}{\sqrt[5]{6}} \leq (x_1 - 1)^2 + (x_2 + 3)^2 + x_3^6 = f(x).$$

综上所述, 定理 3.1 得以验证.

(3) 因为 $\nabla H_1(x) = (0, 2x_2 + 6, -1)^T$ 和 $\nabla G_1(x) = (2x_1 - 2, 0, 1)^T$, 即 (3.13) 满足 MPSC-LICQ. 由定义 2.1, 对于点 $x^* = (1, -3, 0)^T, y^* = (1, -3, 0)^T$, 存在拉格朗日乘子 $(\mu_1^*, \nu_1^*) = (0, 0)$ 使得 (y^*, μ_1^*, ν_1^*) 为 (3.14) 的可行点且有

$$\mu_1^* G(x^*) + \nu_1^* H(x^*) = 0.$$

因此, (y^*, μ_1^*, ν_1^*) 为 (3.14) 的全局最大解, 同时有 $f(x^*) = L(y^*, \mu_1^*, \nu_1^*)$. 定理 3.2 得证.

4 小结

上述考虑的均是在凸性条件下的对偶刻画, 如果弱化凸性假设, 则一方面, 在证明弱、强和逆对偶性时, 利用拟凸或伪凸性无法根据定义推出矛盾, 进而也无法证得限制逆和严格逆对偶性成立; 另一方面, 为获得上述结论, 可以通过刻画比原 Wolfe 型对偶约束条件更加严格的 Mond-Weir 型对偶问题得到. 由于篇幅受限, 本文只针对存零约束优化问题建立了 Wolfe 型对偶问题, 并分析了对偶问题分别在凸性、严格凸性的假设下, 原问题的对偶问题的弱、强、逆、限制逆和严格逆对偶结果, 并给出例子验证结论成立.

参 考 文 献

- [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] Yang X Q, Huang X X. Convergence analysis of an augmented Lagrangian method for mathematical programs with complementarity constraints[J]. *Nonlinear Analysis Theory Methods and Applications*, 2005, 63(5-7): e2247–e2256.
- [4] Singh K, Mishra S K. On multiobjective mathematical programming problems with equilibrium constraints[J]. *Applied Mathematics and Information Sciences Letters*, 2019, 7(1): 17–25.

- [5] Li S J, Yang X Q, Teo K L. Duality for semi-definite and semi-infinite programming[J]. Optimization, 2003, 52(4): 507–528.
- [6] Bermdez A, Moreno C. Duality methods for solving variational inequalities[J]. Computers and Mathematics with Applications, 1983, 7(1): 43–58.
- [7] Huang X X, Yang X Q. Duality for multiobjective optimization via nonlinear lagrangian functions[J]. Journal Of Optimization Theory And Applications, 2004, 120(1): 111–127.
- [8] 孙祥凯. 不确定信息下凸优化问题的鲁棒解刻划 [J]. 数学物理学报, 2017, 37(2): 257–264.
- [9] 陈世国. 非凸不可微多目标规划的 Mond-Weir 型对偶性 [J]. 数学杂志, 1995, 15(4): 381–384.
- [10] 陈世国, 刘家学. 含有锥约束的多目标变分问题的广义对称对偶性 [J]. 数学杂志, 2011, 31(6): 1145–1151.
- [11] Mishra S K, Singh V, Laha V. On duality for mathematical programs with vanishing constraints[J]. Annals of Operations Research, 2016, 241(1-2): 249–293.
- [12] Cambini A, Martein L. Generalized convexity and optimization: theory and applications[M]. Berlin: Springer, 2009.

DUALITY PROBLEM OF MATHEMATICAL PROGRAMS WITH SWITCHING CONSTRAINTS

LUO Mei-ling¹, LI Gao-xi¹, WU Chun²

(1. College of Mathematics and Statistics, Chongqing Technology and Business University, Chongqing 400067, China)

(2. College of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China)

Abstract: In this paper, a new class of optimization problems proposed in recent years is studied, which makes it difficult to solve the optimal solution due to the existence of switching constraint. Therefore, this paper proposes a Wolfe type dual model by using the duality theory to solve the mathematical programs with switching constraints problem. Under the convexity and strict convexity assumptions, the weak, strong, inverse, restricted inverse and strict inverse duality results of Wolfe's duality are obtained. An example is given to demonstrate.

Keywords: nonlinear programming; switching constraint; duality problem; generalized convexity

2010 MR Subject Classification: 90C30; 90C25