

半定锥约束变分不等式的二阶充分性条件

卢 丹, 张 杰, 刘书宇

辽宁师范大学, 数学学院, 辽宁 大连

收稿日期: 2024年10月21日; 录用日期: 2024年11月15日; 发布日期: 2024年11月22日

摘 要

半定锥约束变分不等式问题是目前研究热门话题——锥约束变分不等式问题的其中一部分, 在优化理论、经济学和工程应用中具有重要意义。在本文中, 我们给出了半定锥约束变分不等式问题的二阶充分性条件。我们首先回顾了变分不等式的基本理论概念及其在半定锥约束下的具体形式, 然后通过构造相应的拉格朗日函数, 将问题的解与二阶充分性条件关联起来, 最后我们给出了相关定理的具体内容及其证明, 确保在给定约束条件下解的存在性与最优性。

关键词

半定锥约束, 变分不等式, 二阶充分性条件

Second-Order Sufficient Conditions for Semi-Definite Constrained Variational Inequalities

Dan Lu, Jie Zhang, Shuyu Liu

School of Mathematics, Liaoning Normal University, Dalian Liaoning

Received: Oct. 21st, 2024; accepted: Nov. 15th, 2024; published: Nov. 22nd, 2024

Abstract

The semi-definite cone-constrained variational inequality problem is currently a hot topic of research and is a part of the broader cone-constrained variational inequality issues, which hold significant importance in optimization theory, economics, and engineering applications. In this paper, we present the second-order sufficient conditions for the semi-definite cone-constrained variational inequality problem. We first review the fundamental theoretical concepts of variational inequalities and their specific forms under semi-definite cone constraints. Then, by constructing the corresponding

Lagrangian function, we establish a connection between the solutions of the problem and the second-order sufficient conditions. Finally, we provide the specific content and proof of the relevant theorems to ensure the existence and optimality of the solutions under the given constraints.

Keywords

Semi-Definite Cone Constraint, Variational Inequality, Second-Order Sufficient Conditions

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

变分不等式(Variational Inequalities, VI)是一类重要的数学问题,自 20 世纪 60 年代起,Lions、Browder、Stampacchia、KyFan、Lemke、Cottle 和 Dantzing 等人提出并创立变分不等式问题及其相关基本理论,此后经过许多数学家所进行的杰出工作,使变分不等式问题的相关理论取得了重要发展,并广泛应用于博弈论、金融学 and 力学等许多领域。随着半定编程的发展,研究者们开始关注涉及半定约束的变分不等式问题,根据中国运筹学会编著[1]一文,上个世纪九十年代,二阶锥规划和半定规划成为国际优化领域的主要研究热点之一。这表明了半定锥约束变分不等式问题在那个时期开始受到广泛关注。进一步地,[2]一文详细介绍了半定规划的发展与研究现状、基本理论、相关算法以及变分不等式问题的相关知识与投影类算法。文章还提到,在半定规划问题满足严格可行的约束条件下,利用得到的 KKT 条件,将求解半定规划问题等价地转化为求解变分不等式问题。这说明了半定锥约束变分不等式问题的研究不仅涉及理论基础,还包括具体的算法实现。这些研究内容进一步丰富了半定锥约束变分不等式问题的算法研究领域。综上所述,半定锥约束变分不等式问题的发展历程可以追溯到上个世纪九十年代,并且在随后的研究中,不仅涉及了理论基础的探讨,还包括了多种算法的开发和应用。这些研究为解决实际问题提供了重要的理论和技术支持。

本文旨在研究半定锥约束变分不等式问题的二阶充分性条件,这是因为在解决这类问题时,我们常常需要考虑解的性质,尤其是在最优性条件的研究中。二阶充分性条件为我们提供了一种更为严格的解的性质判定方式,尤其是在存在多重解的情形下。二阶充分性条件能够有效地帮助我们筛选出真正的最优解。因此,本文旨在研究半定锥约束变分不等式问题的二阶充分性条件具有重要的理论意义和应用价值。

2. 预备知识

2.1. 半定锥约束变分不等式问题

用 S^p 表示 $p \times p$ 阶的实对称矩阵空间,在 S^p 中定义内积 $\langle A, B \rangle = \text{Tr}(AB), \forall A, B \in S^p$, 此内积引导的范数即 Frobenius 范数 $\|A\| = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2}$ 。用 S_+^p 表示 $p \times p$ 阶的对称半正定矩阵构成的锥,记 S_-^p 表示 $p \times p$ 阶的对称半负定矩阵构成的锥,可以验证: $S_-^p = -S_+^p$ 。

非线性半定规划问题: $\min_{x \in Q} f(x)$, s.t. $G(x) \in S_+^p$, 其中 $f: R^n \rightarrow R$ 是实值函数, $G: R^n \rightarrow S^p$ 是一矩阵

值函数, $Q \in R^n$ 是一闭凸集。

基于上述描述, 我们可定义 SDCCVI 问题。

定义 2.1 半定锥约束变分不等式问题(SDCCVIP): 给定一个映射 $F: R^n \rightarrow R^n$, 对于给定的子集 C 与半定锥 S_+^p , 要找到一点 $x \in C$, 使得对 $\forall y \in C$, 满足 $\langle F(x), y-x \rangle \geq 0$, 其中 $C = \{x \in R^n | G(x) \in S_+^p\}$ 。

2.2. 切锥与二阶切锥

定义 2.2 本文用 $T_{S_+^p}(x^*) = \begin{cases} S^p, & \text{if } x \in \text{int } S_+^p \\ S_+^p, & \text{if } x = 0 \\ \{\langle w, x^* \rangle \geq 0 | w \in S^p\}, & \text{if } x \in \text{bd } S_+^p \setminus \{0\} \end{cases}$ 来描述 S_+^p 在 $x^* \in S_+^p$ 处的切锥; 用

$T_{S_+^p}^2(x^*, d) = \begin{cases} S^p, & \text{if } x \in T_{S_+^p}(x^*) \\ T_{S_+^p}(x^*), & \text{if } x = 0 \\ \{\langle w, d \rangle \geq \|d\|^2 | w \in S^p\}, & \text{otherwise} \end{cases} = T^2$ 来描述 S_+^p 在 $x^* \in S_+^p$ 处沿方向 d 的二阶切锥。见[3]

[4]。

2.3. 二阶正则

定义 2.3.1 称集合 K 在点 $y \in K$ 沿方向 $d \in T_k(y)$ 关于线性映射 $M: X \rightarrow Y$ 是二阶正则的:

若对任何具有下述形式的 $y_n \in K: y_n = y + t_n d + \frac{1}{2} t_n^2 r_n$, 其中 $t_n \downarrow 0$, $r_n = M w_n + a_n$, 其中 $\{a_n\}$ 是 Y 中的收敛序列, $\{w_n\}$ 是 X 中的满足 $t_n w_n \rightarrow 0$ 的序列, 则下述条件成立: $\lim_{n \rightarrow \infty} d(r_n, T_{S_+^p}^2(y, d)) = 0$ 。见[3]。

3. 二阶充分性条件定理及其证明

接下来我们来描述 SDCCVI 问题的二阶充分性条件定理及其证明, 该证明同 Sun 在[4]中定理 2 类似。

给出定理之前, 补充以下定义。

定义 3.1 点 x^* 到 S_+^p 的距离: $\text{dist}(x^*, S_+^p) := \inf \{\|x^* - y\|, \forall y \in S_+^p\}$ 。

定义 3.2 支撑函数: 记 $\sigma(x, C) = \max \{\langle x, v \rangle : v \in C\}$ 为集合 C 的支撑函数。

定义 3.3 本文用 $\wedge(x^*)$ 记作满足 kkt 条件的 Lagrange 乘子, 将 $t_n \downarrow 0$ 表示 $\{t_n\}$ 是单调递减趋于 0 的序列。

定义 3.4 x^* 处的临界锥与极限法锥: $C(x^*) = \{d | d \in T_C(x^*), d \perp F(x^*)\}$; $N_{S_+^p}(x^*) := \limsup_{x \xrightarrow{S_+^p} x^*} \hat{N}_{S_+^p}(x)$,

其中 $\hat{N}_{S_+^p}(x^*) := \{v \in S^p | \langle v, y - x^* \rangle \leq o(\|y - x^*\|), \forall y \in S_+^p\}$ 。

定理 3.1 设 x^* 为 SDCCVI 的一个可行解, $\Lambda(x^*) = \{\lambda\}$ 是非空且紧的, 若 $JF(x^*)$ 正半定且

$0 \in \text{int}\{G(x^*) + JG(x^*)R^n - S_+^p\}$, 则 $\sup_{\lambda \in \Lambda(x^*)} \left\{ \langle J_x L(x^*, \lambda) d, d \rangle + d^T H(x^*, \lambda) d \right\} > 0, \forall d \in C(x^*) \setminus \{0\}$ 为 SDCCVI 的二阶充分性条件。

其中 $\otimes$ 是 Kronecker 积, $H(x^*, \lambda) = -2 \left(\frac{\partial G(x^*)}{\partial x} \right)^T \left(\lambda \otimes G(x^*)^T \right) \left(\frac{\partial G(x^*)}{\partial x} \right)$, $\frac{\partial G(x^*)}{\partial x}$ 为 $P^2 \times n$ Jacobi

矩阵, $\frac{\partial G(x^*)}{\partial x} := [\text{vec}G_1(x^*), \dots, \text{vec}G_n(x^*)]$, $G_i(x^*) = \frac{\partial G(x^*)}{\partial x_i}$, $\text{vec}(A)$ 记由 A 的列拉直得到的向量。

证明:

设 x^* 是问题 SDCCVI 的一个可行解, 由 $JF(x^*)$ 正半定的, 则存在 $\varepsilon > 0$ 使得 $\langle F(x^*), x - x^* \rangle \geq 0$, $\forall x \in B_\varepsilon(x^*) \cap C$, 其中 $B_\varepsilon(x^*)$ 为 x^* 的 ε 邻域。

这等价于

$$x \in \arg \min \{ \langle F(x^*), x - x^* \rangle \mid x \in B_\varepsilon(x^*) \cap C \} \quad (1)$$

由于 $JF(x^*)$ 正半定则

(1)成立当且仅当

$$x \in \arg \min \{ \langle F(x^*), x - x^* \rangle + JF(x^*)(x - x^*), (x - x^*) \mid x \in B_\varepsilon(x^*) \cap C \} \quad (2)$$

成立。

考虑这样的优化问题

$$\begin{cases} \min \langle F(x^*), x - x^* \rangle + \frac{1}{2} \langle JF(x^*)(x - x^*), x - x^* \rangle \\ \text{s.t. } x \in B_\varepsilon(x^*) \cap C \end{cases} \quad (3)$$

显然

x^* 为(3)的稳定点当且仅当

$$0 \in F(x^*) + JF(x^*)(x - x^*) + N_{B_\varepsilon(x^*) \cap C}(x^*) \quad (4)$$

其中

$$N_{B_\varepsilon(x^*) \cap C}(x^*) = N_{B_\varepsilon(x^*)}(x^*) + N_{C(x^*)}(x^*) = N_{C(x^*)}(x^*) \quad (5)$$

由(4)和(5)知 $0 \in F(x^*) + N_{C(x^*)}(x^*)$ 。

因此当 x^* 为 SDCCVI 问题的可行解时, x^* 也是(3)的稳定点。

接下来考虑(3)的临界锥 $C_p(x^*)$ 与 SDCCVI 的临界锥 $C(x^*)$ 之间的关系: 事实上 $C(x^*) = C_p(x^*)$ 。

这是因为

$$C_p(x^*) = \left\{ d \in \mathbb{R}^n \mid \begin{pmatrix} JG(x^*)d \\ d \end{pmatrix} \in T_{S_+^p \times B_\varepsilon(x^*)}(G(x^*), x^*), \langle d, F(x^*) + JF(x^*)(x - x^*) \rangle = 0 \right\}$$

且 $T_{S_+^p \times B_\varepsilon(x^*)}(G(x^*), x^*) = T_{S_+^p}(G(x^*)) \times T_{B_\varepsilon(x^*)}(x^*) = T_{S_+^p}(G(x^*)) \times \mathbb{R}^n$ 。

因此 $C_p(x^*) = \{ d \in \mathbb{R}^n \mid (JG(x^*)d) \in T_{S_+^p}(G(x^*)), \langle d, F(x^*) \rangle = 0 \} = C(x^*)$ 。

易知(3)的 Lagrange 函数为

$$L(x^*, \lambda, \mu) = \langle F(x^*), x - x^* \rangle + \frac{1}{2} \langle JF(x^*)(x - x^*), x - x^* \rangle + \langle -G(x), \lambda \rangle + \langle x, \mu \rangle$$

则

$$\nabla_x L(x^*, \lambda, \mu) = F(x^*) + JF(x^*)(x - x^*) + JG(x^*)\lambda + \mu;$$

$$\nabla_{xx}L(x^*, \lambda) = JF(x^*) + J_x G(x^*)\lambda, \text{ 此处记 } \nabla_{xx}L(x^*, \lambda) = J_x L(x^*, \lambda)。$$

又 S_+^p 在 $G(x^*)$ 沿着 $JG(x^*)d$ 方向对于映射 $JG(x^*), \forall d \in C(x^*)$ 是二阶正则的, 这是因为利用二阶正则性定义

$$y_n = G(x^*) + t_n(JG(x^*)d) + \frac{1}{2}t_n^2 r_n, \forall y_n \in S_+^p$$

其中 $t_n \downarrow 0$, $r_n = JG(x^*)w_n + a_n$, a_n 为一收敛序列以及 $t_n w_n \rightarrow 0 (n \rightarrow +\infty)$ 使得

$$\lim_{n \rightarrow \infty} \text{dist}(r_n, T_{S_+^p}^2(G(x^*), JG(x^*)d)) = 0。$$

基于以上所述, 知对于 $\forall P_n \in S_+^p \times B_\varepsilon(x^*)$, $P_n = \begin{pmatrix} G(x^*) \\ x^* \end{pmatrix} + t_n \begin{pmatrix} JG(x^*)d \\ d \end{pmatrix} + \frac{1}{2}t_n^2 \begin{pmatrix} r_n \\ q_n \end{pmatrix}。$

其中 $\begin{pmatrix} r_n \\ q_n \end{pmatrix} = \begin{pmatrix} JG(x^*)w_n \\ w_n \end{pmatrix} + \begin{pmatrix} a_n \\ b_n \end{pmatrix}$, $\begin{pmatrix} a_n \\ b_n \end{pmatrix}$ 是一个收敛序列以及 $t_n w_n \rightarrow 0 (n \rightarrow +\infty)$,

因此我们得到

$$\lim_{n \rightarrow \infty} \text{dist}(r_n, T_{S_+^p}^2(G(x^*), JG(x^*)d)) = 0$$

且

$$\begin{aligned} & \lim_{n \rightarrow \infty} \text{dist}\left(\begin{pmatrix} r_n \\ q_n \end{pmatrix}, T_{S_+^p \times B_\varepsilon(x^*)}^2\left(\begin{pmatrix} G(x^*) \\ x^* \end{pmatrix}, \begin{pmatrix} JG(x^*)d \\ d \end{pmatrix}\right)\right) \\ &= \lim_{n \rightarrow \infty} \text{dist}\left(\begin{pmatrix} r_n \\ q_n \end{pmatrix}, T_{S_+^p}^2(G(x^*), JG(x^*)d) \times T_{B_\varepsilon(x^*)}^2(x^*, d)\right) \\ &= \lim_{n \rightarrow \infty} \text{dist}(r_n, T_{S_+^p}^2(G(x^*), JG(x^*)d)) = 0. \end{aligned}$$

因此 S_+^p 在 $G(x^*)$ 沿着 $JG(x^*)d$ 方向对于映射 $JG(x^*), \forall d \in C(x^*)$ 是二阶正则的。

因此 $\sup_{\lambda \in \Lambda(x^*)} \left\{ \nabla_{xx}^2 L(x^*, \lambda) - \delta^* \times ((\lambda, \mu), T_{S_+^p \times B_\varepsilon(x^*)}^2(G(x^*), x^*), (JG(x^*)d, d)) \right\} > 0, \forall d \in C_p(x^*) \setminus \{0\}。$

我们可以进一步化简

$$\begin{aligned} & \sup_{\lambda \in \Lambda(x^*)} \left\{ \nabla_{xx}^2 L(x^*, \lambda)(d, d) - \delta^* \times ((\lambda, \mu), T_{S_+^p \times B_\varepsilon(x^*)}^2(G(x^*), x^*), (JG(x^*)d, d)) \right\} \\ &= \sup_{\lambda \in \Lambda(x^*)} \left\{ \nabla_{xx}^2 L(x^*, \lambda)(d, d) - \delta^* \times ((\lambda, \mu), T_{S_+^p}^2(G(x^*), JG(x^*)d) \times T_{B_\varepsilon(x^*)}^2(x^*, d)) \right\} \\ &= \sup_{\lambda \in \Lambda(x^*)} \left\{ \nabla_{xx}^2 L(x^*, \lambda)(d, d) - \delta^* \times ((\lambda, \mu), T_{S_+^p}^2(G(x^*), JG(x^*)d \times \mathbb{R}^n)) \right\} \\ &= \sup_{\lambda \in \Lambda(x^*)} \left\{ J_x L(x^*, \lambda)(d, d) - \delta^* \left(\lambda | T_{S_+^p}^2(G(x^*) + JG(x^*)d) \right) \right\}. \end{aligned}$$

综上所述, SDCCVI 问题的二阶充分性条件为

$$\sup_{\lambda \in \Lambda(x^*)} \left\{ J_x L(x^*, \lambda)(d, d) - \delta^* \left(\lambda | T_{S_+^p}^2(G(x^*) + JG(x^*)d) \right) \right\}, \forall d \in C(x^*) \setminus \{0\}$$

即

$$\sup_{\lambda \in N(x^*)} \left\{ \left\langle J_x L(x^*, \lambda) d, d \right\rangle + d^T H(x^*, \lambda) d \right\} \succ 0, \forall d \in C(x^*) \setminus \{0\}$$

其中 $H(x^*, \lambda) = -2 \left(\frac{\partial G(x^*)}{\partial x} \right)^T \left(\lambda \otimes G(x^*)^T \right) \left(\frac{\partial G(x^*)}{\partial x} \right)$, $\frac{\partial G(x^*)}{\partial x}$ 为 $P^2 \times n$ Jacobi 矩阵。

4. 结论

1) 重要性: 本论文详细描述并证明了半定锥约束变分不等式二阶充分性条件定理及其相关证明, 提供了一种有效的方法来验证该类问题局部最优解的性质。这在许多领域(如控制系统、资源分配、线性回归及金融投资组合优化)中尤为重要, 能够帮助决策者在多维空间中找到最优解。

2) 先进性: 本文结合拉格朗日乘数法和 Hessian 矩阵的分析为优化理论提供了坚实的基础。这种理论框架使得半定锥约束变分不等式问题的求解过程变得更加系统化、更方便。此外, 论文中所涉及的数值分析与计算方法还能够帮助提高相关算法的收敛性与效率, 对获得更加精确的数值解有重要意义。

3) 局限性: 虽然二阶条件能有效判断局部最优性, 但在某些特定情况下, 如非凸问题中, 仅依赖于此条件可能会导致无法找到全局最优解。因此, 在实际应用中需结合其他方法一同使用, 以确保结果的全局最优性。

致 谢

本人感谢我的导师张教授对本论文的指导, 在研究过程中提供了宝贵的建议和支持, 他严谨的治学态度和深厚的专业知识使我创作了这篇论文并收益颇丰。

参考文献

- [1] 中国科学技术协会. 2012-2013 运筹学学科发展报告[Z]. 2014.
- [2] 秦淑英. 半定规划的两种投影类算法[D]: [硕士学位论文]. 呼和浩特: 内蒙古工业大学, 2015.
- [3] 张立卫, 吴佳, 张艺. 变分分析与优化[M]. 北京: 科学出版社, 2013.
- [4] Sun, J., Fu, W., Alcantara, J.H. and Chen, J. (2021) A Neural Network Based on the Metric Projector for Solving SOCCVI Problem. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 2886-2900. <https://doi.org/10.1109/tnnls.2020.3008661>