Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 19, No. 4, pp. 625–643, 2023

RESEARCH ON OPTIMALITY CONDITIONS FOR SSOCCVI PROBLEM

Li Wang, Mengdi Zheng, Juhe Sun and Hao Liu

Abstract: In this paper, we study the optimality conditions for the stochastic second-order cone constrained variational inequality problem. The sample average approximation method is used first to approximate the expected value function and get a approximation problem called SAA problem. The Karush-Kuhn-Tucher system of SAA problem is studied, and the second-order sufficient condition is further defined. Besides, we prove that the Jacobian matrix of equation system's operator is non-singular. Finally, the Newton's algorithm is established and two numerical examples are gave, the results prove that the algorithm is effective.

Key words: *stochastic variational inequality, second-order sufficient condition, numercial examples*

Mathematics Subject Classification: *60H35, 65K99*

1 Introduction

At present, stochastic programming has become one of the research hot spots of scholars, see [1, 2, 3, 4, 5], [13, 15, 16, 18, 21] and the references therein. As an important branch of stochastic programming, stochastic variational inequality has many applications in practical problems. The main methods to solve stochastic variational inequality problem are sample path optimization method [1], sample average approximation method [2, 3, 4], stochastic approximation method [5]. The optimality conditions of stochastic variational inequality are studied in some literatures, see [6, 7, 8, 9]. However, there are few studies on the second-order sufficient condition of stochastic variational inequality problem. Therefore, based on our previous research work [17], we generalize the problem and study the optimality conditions of second-order cone constrained stochastic variational inequality problem.

The stochastic second-order cone constrained variational inequality problem (abbrevd. SSOCCVI), which is to find $x \in C$, such that

$$\langle E[f(x,\xi)], y - x\rangle \geq 0, \forall y \in C, \tag{1.1}$$

where $\langle \cdot, \cdot \rangle$ represents the Euclidean inner product, $E[f(x,\xi)] = \int_\Omega f(x,\xi)dP(\xi)$ is the expectation of ξ, the set C is

$$C = \{x \in \mathcal{R}^n |\ E[h(x,\eta)] = 0, -E[g(x,\zeta)] \in \mathcal{K}\}, \tag{1.2}$$

where $\mathcal{R}^n$ is a real space with n-dimension, and $\mathcal{K}$ is second-order cone with m-dimension. $f(\cdot,\xi)$ is local *Lipschitz* continuous for any ξ, for each fixed η and ζ, $h(\cdot,\eta) : \mathcal{R}^n \to \mathcal{R}^l$, and $g(\cdot,\zeta) : \mathcal{R}^n \to \mathcal{R}^m$ are both continuous differentiable convex functions. ξ, η, ζ are stochastic variables defined in probability space $(\Omega, \mathcal{F}, P)$, where $\mathcal{F}$ is sigma algebra, P is the probability measure on Ω, and

$$\mathcal{K} = \mathcal{K}^{m_1} \times \mathcal{K}^{m_2} \times \cdots \times \mathcal{K}^{m_q}, \tag{1.3}$$

where $q \geq 1$, $m_i \geq 1$, $i = 1, 2, \ldots, q$ and $m_1 + m_2 + \cdots m_q = m$. $\mathcal{K}^{m_i}$ is a second-order cone with m_i-dimension.

2 Preliminaries

In order to better carry out the research, this chapter gives the basic definitions, properties and conclusions required for the research.

Clarify the symbols used in this paper first. Given a mapping $F : \mathcal{R}^n \to \mathcal{R}^m$, we use $JF(x)$ and $\nabla F(x)$ to denote the Jacobian and gradient of F, respectively, and $JF(x) = \nabla F(x)^T$, moreover, $\nabla^2 F(x)$ means Hessian matrix. Given a sequence $\{t_n\} \in \mathcal{R}$, we use $t_n \downarrow 0$ to express $\{t_n\}$ is monotone decreasing and converges to 0. The distance from a point x to a set K is defined by

$$dist(x, K) := inf\{\|x - y\| | \forall y \in K\}. \tag{2.1}$$

Besides, we use $linK$ to denote the linear subspace generated by K.

If a subset $\mathcal{K}^m$ of $\mathcal{R}^m$ satisfies

$$\mathcal{K}^m = \{x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1} | x_1 \geq \|x_2\|\}, \tag{2.2}$$

then $\mathcal{K}^m$ is called second-order cone with m-dimension.

For any two vectors $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$ and $y = (y_1, y_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$, the Jordan product is defined by $x \cdot y = (x^T y, x_1 y_2 + x_2 y_1)$. For any $x \in \mathcal{R}^m$, $x^2 = x \cdot x$, which belongs to $\mathcal{K}^m$. The square root of $x \in \mathcal{K}^m$ is also well defined, which is denoted by $x^{\frac{1}{2}}$ or $\sqrt{x}$, and $x = (x^{\frac{1}{2}})^2$ or $x = (\sqrt{x})^2$, similarly $|x| = \sqrt{x^2}$. Besides, for any vector $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$, the determinant and trace are defined by $det(x) = x_1^2 - ||x_2||^2$, $tr(x) = 2x_1$, generally, $det(x \cdot y) \neq det(x)det(y)$, and only when $x = \alpha y, \alpha \in \mathcal{R}$, the equal sign holds. In addition, a vector $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$ is called invertible, if $det(x) \neq 0$, and its inverse satisfies $x \cdot x^{-1} = e, e = (1, 0, ..., 0)^T \in \mathcal{R}^m$. It is not difficult to deduce that $x^{-1} = \frac{tr(x)e - x}{det(x)}$.

Any vector $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$, has the spectral decomposition

$$x = \rho_1 u^{(1)} + \rho_2 u^{(2)}, \tag{2.3}$$

where ρ_1, ρ_2 are the spectral values, given by

$$\rho_i = x_1 + (-1)^i \|x_2\|, \tag{2.4}$$

and $u^{(1)}, u^{(2)}$ are spectral vectors corresponding to ρ_1, ρ_2, with formulas

$$u_i = \begin{cases} \frac{1}{2}(1, (-1)^i \frac{x_2}{\|x_2\|}), & if\ x_2 \neq 0, \\ \frac{1}{2}(1, (-1)^i w), & if\ x_2 = 0. \end{cases} \tag{2.5}$$

for $i = 1, 2$, where w is an arbitrary vector in $\mathcal{R}^{m-1}$, and $\|w\| = 1$. Obviously, the spectral decomposition is unique if and only if $x_2 \neq 0$.

Proposition 2.1. *Any a vector $x = (x_1, x_2) \in \mathcal{R} \times \mathcal{R}^{m-1}$, the spectral values are ρ_1, ρ_2 and corresponding spectral vectors are $u^{(1)}, u^{(2)}$, then the following properties hold.*

(a) $x \in \mathcal{K}^m$, *if and only if,* $\rho_1 \geq 0$*;*

(b) *Under the Jordan product,* $u^{(1)}$, $u^{(2)}$ *are orthogonal, and have the same mode length* $\frac{1}{\sqrt{2}}$, *i.e.,*

$$u^{(1)}u^{(2)} = 0, \|u^{(1)}\| = \|u^{(2)}\| = \frac{1}{\sqrt{2}};$$

(c) $u^{(1)}, u^{(2)}$ *have idempotent property under the Jordan product, i.e.,* $u^{(i)}u^{(i)} = u^{(i)}$*;*

(d) $x^2 = \rho_1^2 u^{(1)} + \rho_2^2 u^{(2)} \in \mathcal{K}^m$, *and* $x^2 \in \mathcal{K}^m$*;*

(e) *If* $x \in \mathcal{K}^m$, *then* $x^{\frac{1}{2}} = \sqrt{\rho_1}u^{(1)} + \sqrt{\rho_2}u^{(2)}$, *and* $x^{\frac{1}{2}} \in \mathcal{K}^m$*;*
(f) $det(x) = \rho_1\rho_2$, $tr(x) = \rho_1 + \rho_2$, $\|x\|^2 = \frac{\rho_1^2+\rho_2^2}{2}$.

The revelent proof see [10], [20].

Definition 2.2. A mapping $\phi : \mathcal{R}^n \times \mathcal{R}^n \to \mathcal{R}^n$ is called complementary function if and only if

$$x \in \mathcal{K}, y \in \mathcal{K}, x^T y = 0 \Longleftrightarrow \phi(x, y) = 0. \tag{2.6}$$

Suppose x has the spectral decomposition as (3), the projection $\Pi_{\mathcal{K}^m}(x)$ of x onto $\mathcal{K}^m$ is

$$\Pi_{\mathcal{K}^m}(x) = \max\{0, \rho_1\}u^{(1)} + \max\{0, \rho_2\}u^{(2)}.$$

Actually, plugging in ρ_i and $u^{(i)}$ given in (2.4) and (2.5), yields

$$\Pi_{\mathcal{K}^m}(x) = \begin{cases} \frac{1}{2}(1 + \frac{x_1}{\|x_2\|})(\|x_2\|, x_2), & if\ |x_1| < \|x_2\|, \\ (x_1, x_2), & if\ \|x_2\| \leq x_1, \\ 0, & if\ \|x_2\| \leq -x_1. \end{cases} \tag{2.7}$$

Lemma 2.3 below gives the derivative of projection $\Pi_{\mathcal{K}^m}(x)$. The boundary, interior and closure of $\mathcal{K}^m$, are denoted as bd$\mathcal{K}^m$, int$\mathcal{K}^m$ and cl$\mathcal{K}^m$, respectively.

Lemma 2.3. *The metric projection operator* $\Pi_{\mathcal{K}^m}(\cdot)$ *is directionally differentiable at* x *for any* $d \in \mathcal{R}^m$, *and*

$$\Pi'_{\mathcal{K}^m}(x; d) = \begin{cases} J\Pi_{\mathcal{K}^m}(x)d, & if\ x \in \mathcal{R}^m \backslash \{\mathcal{K}^m \cup -\mathcal{K}^m\}, \\ d, & if\ x \in \operatorname{int} \mathcal{K}^m, \\ d - 2[u^{(1)T}d]_- u^{(1)}, & if\ x \in bd\mathcal{K}^m \backslash \{0\}, \\ 0, & if\ x \in -\operatorname{int} \mathcal{K}^m, \\ 2[u^{(2)T}d]_+ u^{(2)}, & if\ x \in -bd\mathcal{K}^m \backslash \{0\}, \\ \Pi_{\mathcal{K}^m}(d), & if\ x = 0. \end{cases} \tag{2.8}$$

where

$$J\Pi_{\mathcal{K}^m}(x) = \frac{1}{2} \begin{pmatrix} 1 & \frac{{x_2}^T}{\|x_2\|} \\ \frac{x_2}{\|x_2\|} & I + \frac{x_1}{\|x_2\|} I - \frac{x_1}{\|x_2\|} \cdot \frac{x_2 {x_2}^T}{\|x_2\|^2} \end{pmatrix}. \tag{2.9}$$

Indeed, $[u^{(1)}d^T]_- := \min\{0, u^{(1)T}d\}, [u^{(2)}d^T]_+ := \max\{0, u^{(2)T}d\}$.

Definition 2.4. For a closed set $K \subseteq \mathcal{R}^n$ and a point $x \in K$, the relevant cones, which can be found in [11], are defined by

(1) The tangent cone

$$T_K(x) := \limsup_{t \downarrow 0} \frac{K - x}{t}, \tag{2.10}$$

(2) The regular normal cone

$$\hat{N}_K(x) := \{ v \in \mathcal{R}^n \,|\, \langle v, y - x \rangle \leq o(\|y - x\|), \forall y \in K \}, \tag{2.11}$$

(3) The normal cone

$$N_K(x) := \lim_{y \xrightarrow{K} x} \sup \hat{N}_K(x)\,. \tag{2.12}$$

Notice that : if K is a closed convex set, then $T_K(x) = cl(K + \mathcal{R}x)$, $\hat{N}_K(x) = N_K(x) = T_K(x)^\circ = \{v \in K^\circ | \langle v, x \rangle \leq 0\}$, where K° represents the polar cone of K.

Lemma 2.5. *The tangent cone and second-order tangent cones of $\mathcal{K}^m$ at $x \in \mathcal{K}^m$ are expressed, respectively, by*

$$T_{\mathcal{K}^m}(x) = \begin{cases} \mathcal{R}^m, & if \;\; x \in \operatorname{int} \mathcal{K}^m, \\ \mathcal{K}^m, & if \;\; x = 0, \\ \{d = (d_1, d_2) \in \mathcal{R} \times \mathcal{R}^{m-1} | \langle d_2, x_2 \rangle - x_1 d_1 \leq 0\}, & if \;\; x \in bd\mathcal{K}^m \backslash \{0\}. \end{cases} \tag{2.13}$$

and

$$T^2_{\mathcal{K}^m}(x, d) = \begin{cases} \mathcal{R}^m, & if \;\; x \in \operatorname{int} T_{\mathcal{K}^m}(x), \\ T_{\mathcal{K}^m}(d), & if \;\; x = 0, \\ \{w = (w_1, w_2) \in \mathcal{R} \times \mathcal{R}^{m-1} | \langle w_2, x_2 \rangle - w_1 x_1 \leq d_1^2 - \|d_2\|^2\}, & otherwise. \end{cases} \tag{2.14}$$

Lemma 2.6 ([22])**.** *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed samples taken from random variable X, with mean μ. Suppose that the moment generating function $M(t) = E[e^{tX_i}] < \infty$, then for any $z > \mu$, the equation holds*

$$\lim_{n \to \infty} \frac{1}{n} \log P\{\frac{1}{n} \sum_{i=1}^{n} X_i \geq z\} = -I(z), \tag{2.15}$$

and the rate function is $I(z) = \sup_{t \in \mathcal{R}} \{tz - logM(t)\}$. The lemma is called the Cramér's Large Deviation Theorem.

3 Sample Average Approximation

There are three methods to solve stochastic variational inequality problem (SVIP): sample path optimization (SPO), sample average approximation (SAA) and stochastic approximation (SA). In this paper, we adopt SAA method to solve SSOCCVI problem. Some basic assumptions used in the process of applying the method can be found in [2], the main idea is to generate N independent and identically distributed random samples $\xi_1, \xi_2, ..., \xi_N$, and use their sample mean function

$$F^N(x) = \frac{1}{N}\sum_{i=1}^{N} f(x, \xi_i),$$

to approximate expected value function $E[f(x, \xi)]$.

Definition 3.1. The moment generating function of random variable $f(x, \xi)$ is defined as

$$M(t) = E[e^{tf(x,\xi)}], \tag{3.1}$$

thus, the moment generating function of $f(x, \xi) - E[f(x, \xi)]$ is denoted by

$$M_x(t) = E[e^{t(f(x,\xi)-E[f(x,\xi)])}]. \tag{3.2}$$

Assumption 3.1. (a) For any $x \in C$, the moment generating function $M_x(t)$ is finite w.r.t. t in a certain neighborhood of zero;

(b) There is a metric function $k : \Omega \to R_+$, such that, for any $\xi \in \Omega$, $x^{'}, x \in X$, there is

$$|f(x^{'}, \xi) - f(x, \xi)| \leq k(\xi)\|x^{'} - x\|; \tag{3.3}$$

(c) The moment generating function $M_k(t) = E[e^{tk(\xi)}]$ is finite w.r.t. t in a certain neighborhood of zero.

Proposition 3.2. *Let $K \subset R^n$ and X is a compact subset of K. If Assumption 3.1 holds, then for any $\varepsilon > 0$, there are $c(\varepsilon) > 0$, $\beta(\varepsilon) > 0$ (independent of N), such that*

$$Prob\{\sup_{x\in X} |F^N(x) - E[f(x, \xi)]| \geq \varepsilon\} \leq c(\varepsilon)e^{-N\beta(\varepsilon)}. \tag{3.4}$$

Proof. By Cramér's Large Deviation Theorem, we can obtain that for any $x \in X$, $\varepsilon > 0$,

$$Prob\{F^N(x) - E[f(x, \xi)] \geq \varepsilon\} \leq exp\{-NI_x(\varepsilon)\}, \tag{3.5}$$

again, by Cramér's Large Deviation Theorem, we have

$$Prob\{F^N(x) - E[f(x, \xi)] \leq -\varepsilon\} \leq exp\{-NI_x(-\varepsilon)\}, \tag{3.6}$$

thus

$$Prob\{|F^N(x) - E[f(x, \xi)]| \geq \varepsilon\} \leq exp\{-NI_x(\varepsilon)\} + exp\{-NI_x(-\varepsilon)\}. \tag{3.7}$$

The Assumption 3.1(a) satisfies the condition of Cramer's Large Deviation Theorem, and

$$I_x(z) := \sup_{t\in R}\{zt - logM_x(t)\}, \tag{3.8}$$

denotes the rate function of random variable $f(x, \xi) - E[f(x, \xi)]$, then by the definition of rate function, $I_x(\varepsilon)$, $I_x(-\varepsilon)$ are positive for any $x \in X$. The Assumption 3.2(b) illustrates

$$|E[f(x^{'}, \xi)] - E[f(x, \xi)]| \leq L\|x^{'} - x\|, \tag{3.9}$$

where $L := E[k(\xi)]$ is finite by the (c), and the inequality holds based on the points below: by law of large numbers, the mean of random variables converges to the expectation with probability one, hence, when $N \to \infty$, the inequality can be converted into

$$|\frac{1}{N}\sum_{i=1}^{N} f(x^{'}, \xi_i) - \frac{1}{N}\sum_{i=1}^{N} f(x, \xi_i)| \leq \frac{1}{N}\sum_{i=1}^{N} k(\xi_i)\|x^{'} - x\|,$$

equivalently

$$|\sum_{i=1}^{N} f(x^{'}, \xi_i) - \sum_{i=1}^{N} f(x, \xi_i)| \leq \sum_{i=1}^{N} k(\xi_i)\|x^{'} - x\|.$$

On the other hand, $|\sum_{i=1}^{N} f(x^{'}, \xi_i) - \sum_{i=1}^{N} f(x, \xi_i)| \leq \sum_{i=1}^{N} |f(x^{'}, \xi_i) - f(x, \xi_i)|$. Thus, we need only to proof the inequality below

$$\sum_{i=1}^{N} |f(x^{'}, \xi_i) - f(x, \xi_i)| \leq \sum_{i=1}^{N} k(\xi_i)\|x^{'} - x\|, \tag{3.10}$$

the condition for the inequality to hold is that Assumption 3.1(b) holds, and the explanation is complete. Moreover,

$$|F^N(x^{'}) - F^N(x)| \leq k^N(\xi)\|x^{'} - x\|, \tag{3.11}$$

where $k^N(\xi) = \frac{1}{N}\Sigma_{i=1}^{N} k(\xi_i)$. Assumption 3.1(c) indicates that $M_k(t) = E[e^{tk(\xi)}]$ is finite in a neighborhood of zero, which satisfies the condition of Cramér's Large Deviation Theorem, thus, for any $L^{'} > L$, $\exists \lambda > 0$, such that

$$Prob\{k^N(\xi) > L^{'}\} \leq exp\{-N\lambda\}. \tag{3.12}$$

In order to proceed further, we define a sequence $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_M\} \in X$, which satisfies for any $x \in X$, there is $\bar{x}_j$ $(j = 1, 2, ..., M)$ such that $\|x - \bar{x}_j\| \leq \nu(\nu > 0)$.

Consider $Z_j := F^N(\bar{x}_j, \xi) - E[f(\bar{x}_j, \xi)]$, $j = 1, 2, ..., M$. It is clear that the event $\{\max_{1\leq j\leq M} |Z_j| \geq \varepsilon\}$ is equal to the union of the events $\{|Z_j| \geq \varepsilon\}$, hence

$$Prob\{\max_{1\leq j\leq M} |Z_j| \geq \varepsilon\} \leq \sum_{j=1}^{M} Prob(|Z_j| \geq \varepsilon), \tag{3.13}$$

combining with (3.7), we achieve

$$Prob\{\max_{1\leq j\leq M} |Z_j| \geq \varepsilon\} \leq \sum_{j=1}^{M} Prob(|Z_j| \geq \varepsilon) \leq 2\sum_{j=1}^{M} exp\{-N[I_{\bar{x}_j}(\varepsilon) \wedge I_{\bar{x}_j}(-\varepsilon)]\}. \tag{3.14}$$

Remark $j(x) \in argmin_{1\leq j\leq M}\|x - \bar{x}_j\|$. From the definition of the sequence $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_M\} \in X$, we get $\|x - \bar{x}_{j(x)}\| \leq \nu$, for any $x \in X$, then

$$\begin{aligned}|F^N(x) - E[f(x, \xi)]| &\leq |F^N(x) - F^N(\bar{x}_{j(x)})| + |F^N(\bar{x}_{j(x)}) - E[f(\bar{x}_{j(x)}, \xi)]| \\ &+ |E[f(\bar{x}_{j(x)}, \xi)] - E[f(x, \xi)]| \\ &\leq k^N(\xi)\nu + |F^N(\bar{x}_{j(x)}) - E[f(\bar{x}_{j(x)}, \xi)]| + L\nu.\end{aligned}$$

Now, let ν such that $L\nu = \frac{\varepsilon}{4}$, i.e., $\nu = \frac{\varepsilon}{4L}$, then

$$Prob\{\sup_{x\in X} |F^N(x) - E[f(x,\xi)]| \geq \varepsilon\} \leq Prob\{k^N\nu + \max_{1\leq j\leq M} |F^N(\bar{x}_j) - E[f(\bar{x}_j,\xi)]| \geq \frac{3\varepsilon}{4}\}, \tag{3.15}$$

moreover, by (3.12), we get

$$Prob\{k^N(\xi) > \frac{\varepsilon}{2}\} \leq exp\{-N\lambda\}, \tag{3.16}$$

for some $\lambda > 0$, therefore, combined with (3.14),(3.15),(3.16), we obtain

$$\begin{aligned} Prob\{\sup_{x\in X} |F^N(x) - E[f(x,\xi)] \geq \varepsilon|\} &\leq exp\{-N\lambda\} + Prob\{\max_{1\leq j\leq M} |F^N(\bar{x}_j) - E[f(\bar{x}_j,\xi)]| \geq \frac{\varepsilon}{4}\} \\ &\leq exp\{-N\lambda\} + 2\sum_{j=1}^{M} exp\{-N[I_{\bar{x}_j}(\frac{\varepsilon}{4}) \wedge I_{\bar{x}_j}(-\frac{\varepsilon}{4})]\}. \end{aligned} \tag{3.17}$$

Since the above choice of the sequence dose not rely on the sample, and both $I_{\bar{x}_j}(\frac{\varepsilon}{4})$, $I_{\bar{x}_j}(-\frac{\varepsilon}{4})$ are positive, the proof is complete. □

4 Second-order sufficient condition and non-singularity theorem

In this section, we use SAA method to transform the SSOCCVI problem into SAA problem and give the KKT system of the SAA problem. Using the natural residual (NR) function, we transform the KKT system into an equation system. Based on first-order necessary condition, we define the second-order sufficient condition of the SAA problem. Finally, the non-singularity of the Jacobian matrix of the equation system's operator is proved.

It is known that SSOCCVI problem (1.1)-(1.2) is equivalent to the generalized equation

$$-E[f(x,\xi)] \in N_C(x), \tag{4.1}$$

where $N_C(x)$ denotes the normal cone of the set C at x, and

$$N_C(x) = \{J_xE[h(x,\eta)]^T\mu + J_xE[g(x,\zeta)]^T\lambda | \mu \in \mathcal{R}^l, \lambda \in \mathcal{K}, \langle\lambda, E[g(x,\zeta)]\rangle = 0\}. \tag{4.2}$$

Thus, the KKT system of (1.1)-(1.2) is

$$\begin{cases} L(x,\mu,\lambda) = 0, \\ E[h(x,\eta)] = 0, \\ E[g(x,\zeta)]^T\lambda = 0, -E[g(x,\zeta)] \in \mathcal{K}, \lambda \in \mathcal{K}. \end{cases} \tag{4.3}$$

where $L(x,\mu,\lambda)$ is the Lagrangian function of (1.1)-(1.2), and

$$L(x,\mu,\lambda) = E[f(x,\xi)] + J_xE[h(x,\eta)]^T\mu + J_xE[g(x,\zeta)]^T\lambda. \tag{4.4}$$

By SAA method, problem (1.1)-(1.2) can be converted into the problem: find $x \in C^N$, such that

$$(y-x)^TF^N(x) \geq 0, \forall y \in C^N, \tag{4.5}$$

and

$$C^N = \{x \in \mathcal{R}^n | h^N(x) = 0, -g^N(x) \in \mathcal{K}\}, \tag{4.6}$$

where $F^N(x) = \frac{1}{N}\Sigma_{i=1}^N f(x,\xi_i)$, $h^N(x) = \frac{1}{N}\Sigma_{i=1}^N h(x,\eta_i)$, $g^N(x) = \frac{1}{N}\Sigma_{i=1}^N g(x,\zeta_i)$. ξ_i, η_i, ζ_i, $i = 1, 2, \dots N$, are independent and identically distributed samples of ξ, η, ζ. We call (1.1)-(1.2) the original problem and (4.5)-(4.6) the SAA problem.

Similarly, (4.5)-(4.6) is equivalent to the generalized equation

$$-F^N(x) \in N_{C^N}(x), \tag{4.7}$$

and $N_{C^N}(x) = \{J_x h^N(x)^T\mu + J_x g^N(x)^T\lambda | \mu \in \mathcal{R}^l, \lambda \in \mathcal{K}, \langle\lambda, g^N(x)\rangle = 0\}$, hence, the KKT system of (4.5)-(4.6) is

$$\begin{cases} F^N(x) + J_x h^N(x)^T\mu + J_x g^N(x)^T\lambda = 0, \\ h^N(x) = 0, \\ g^N(x)^T\lambda = 0, \lambda \in \mathcal{K}, -g^N(x) \in \mathcal{K}. \end{cases} \tag{4.8}$$

To facilitate our research, for each $\mathcal{K}^{m_i}$, we define

$$\mathcal{K}^{m_i} = \{(x_{i1}, x_{i2}, \dots, x_{im_i})^T \in \mathcal{R}^{m_i} | x_{i1} \geq \|(x_{i2}, \dots, x_{im_i})\|\}, \tag{4.9}$$

then $g^N(x)^T\lambda = 0$, can be denoted as

$$g_{m_i}^N(x)^T\lambda_{m_i} = 0, \lambda_{m_i} \in \mathcal{K}^{m_i}, -g_{m_i}^N(x) \in \mathcal{K}^{m_i}, i = 1, 2, \dots, q, \tag{4.10}$$

where $\{\lambda_{m_i} = (\tilde{\lambda}_{m_i}, \bar{\lambda}_{m_i}) \in \mathcal{R} \times \mathcal{R}^{m_i-1} | \tilde{\lambda}_{m_i} \geq \|\bar{\lambda}_{m_i}\|\}$ and $\{g_{m_i}^N = (\tilde{g}_{m_i}^N, \bar{g}_{m_i}^N) \in \mathcal{R} \times \mathcal{R}^{m_i-1} | \tilde{g}_{m_i}^N \geq \|\bar{g}_{m_i}^N\|\}$.

Based on the definition of complementary function given in (2.6), the KKT system of (4.5)-(4.6) can be converted to

$$\Phi_{NR}^N(x,\mu,\lambda) = \begin{pmatrix} L^N(x,\mu,\lambda) \\ -h^N(x) \\ \phi_{NR}^N(-g_{m_1}^N, \lambda_{m_1}) \\ \vdots \\ \phi_{NR}^N(-g_{m_q}^N, \lambda_{m_q}) \end{pmatrix} = 0. \tag{4.11}$$

Further, we choose a semi-smooth natural residual (NR) function in [10], i.e.,

$$\phi_{NR}(x,y) = x - \Pi_{\mathcal{K}^m}(x-y), \tag{4.12}$$

hence, (4.11) is equivalent to

$$\Phi_{NR}^N(x,\mu,\lambda) = \begin{pmatrix} L^N(x,\mu,\lambda) \\ -h^N(x) \\ -g^N(x) - \Pi_{\mathcal{K}}(-g^N(x) - \lambda) \end{pmatrix} = 0, \tag{4.13}$$

where $\Pi_{\mathcal{K}}(-g^N(x) - \lambda) = \left(\Pi_{\mathcal{K}^{m_1}}(-g_{m_1}^N(x) - \lambda_{m_1})^T, \dots, \Pi_{\mathcal{K}^{m_q}}(-g_{m_q}^N(x) - \lambda_{m_q})^T\right)^T$, and for $d \in \mathcal{R}^m$, $\Pi'_{\mathcal{K}}(-g^N(x) - \lambda; d) = diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N(x) - \lambda_{m_i}; d_{m_i})\}_{i=1}^q$.

In conclusion, we convert the KKT condition of (4.5)-(4.6) into the equation $\Phi_{NR}^N(x,\mu,\lambda) = 0$. Next, we are devoted to exploring the second-order sufficient condition of (4.5)-(4.6).

Assume that x^* is a local optimal solution of (4.5)-(4.6), the Robinson's constraint qualification

$$\begin{pmatrix} \nabla h^N(x^*)^T \\ -\nabla g^N(x^*)^T \end{pmatrix} \mathcal{R}^n + T_{\{0_l\}\times\mathcal{K}}(h^N(x^*), -g^N(x^*)) = \mathcal{R}^l \times \mathcal{R}^m, \tag{4.14}$$

holds at x^*. The first-order necessary condition is

$$\langle F^N(x^*), d\rangle \geq 0, \forall d \in T_{C^N}(x^*), \tag{4.15}$$

where

$$T_{C^N}(x^*) = \left\{d|\nabla h^N(x^*)^T d = 0, -\nabla g^N(x^*)^T d \in T_{\mathcal{K}}(-g^N(x^*))\right\}. \tag{4.16}$$

$N_{C^N}(x^*) = \nabla h^N(x^*)\mathcal{R}^l + \left\{\nabla g^N(x^*)\lambda| - \lambda \in N_{\mathcal{K}}(-g^N(x^*))\right\}$. For convenience, for $z = (z_{m_1}, \ldots, z_{m_q}) \in \mathcal{R}^m$, we can write $N_{\mathcal{K}}(z) := N_{\mathcal{K}^{m_1}}(z_{m_1}) \times N_{\mathcal{K}^{m_2}}(z_{m_2}) \times \cdots \times N_{\mathcal{K}^{m_q}}(z_{m_q})$ and $N_{\mathcal{K}^{m_i}}(z_{m_i}) := \{u_{m_i} \in \mathcal{R}^{m_i}|\langle u_{m_i}, v - z_{m_i}\rangle \leq 0\}, \forall v \in \mathcal{K}^{m_i}$.

Definition 4.1 ([12])**.** The critical cone of set C^N at x^* is

$$\mathcal{C}(x^*) = \{d|d \in T_{C^N}(x^*), d \perp F^N(x^*)\}. \tag{4.17}$$

Definition 4.2. Let x^* be a local optimal solution of SAA problem (4.5)-(4.6) and lagrange multiplier set $\Lambda(x^*) = \{(\mu, \lambda)\}$ is nonempty and compact . If $JF^N(x^*)$ is positive semi-definite and Robinson's constraint qualification holds at x^*, then the second-order sufficient condition of (4.5)-(4.6) is

$$\sup_{(\mu,\lambda)\in\Lambda(x^*)} \{\left\langle J_x L^N(x^*, \mu, \lambda)d, d\right\rangle - \delta^*(\lambda|T^2_{\mathcal{K}}(-g^N(x^*), -\nabla g^N(x^*)^T d))\} > 0, \forall d \in \mathcal{C}(x^*)\backslash\{0\}. \tag{4.18}$$

where

$$\delta^*\left(\lambda|T^2_{\mathcal{K}}(-g^N(x^*), -\nabla g^N(x^*)^T d)\right) = \begin{cases} 0, & if \;\; \lambda \in N_{\mathcal{K}}(-g^N(x^*)), \\ +\infty, & otherwise. \end{cases} \tag{4.19}$$

Remark 4.3. Since x^* is a local optimal solution of (4.5)-(4.6), which implies for any $\varepsilon > 0$,

$$\left\langle F^N(x^*), x - x^*\right\rangle \geq 0, \forall x \in \mathbb{B}_\varepsilon(x^*) \cap C^N, \tag{4.20}$$

equivalently,

$$x^* \in \arg\min\left\{\left\langle F^N(x^*), x - x^*\right\rangle \mid x \in \mathbb{B}_\varepsilon(x^*) \cap C^N\right\}. \tag{4.21}$$

Moreover, when $JF^N(x^*)$ is positive semi-definite, it's known that (4.21) holds if and only if

$$x^* \in \arg\min\left\{\left\langle F^N(x^*), x - x^*\right\rangle + \left\langle JF^N(x^*)(x - x^*), x - x^*\right\rangle \mid x \in \mathbb{B}_\varepsilon(x^*) \cap C^N\right\}. \tag{4.22}$$

Hence, problem (4.20) is equivalent to

$$\begin{aligned} \min \quad & \left\langle F^N(x^*), x - x^*\right\rangle + \frac{1}{2}\left\langle JF^N(x^*)(x - x^*), x - x^*\right\rangle \\ s.t. \quad & x \in \mathbb{B}_\varepsilon(x^*) \cap C^N \end{aligned} \tag{4.23}$$

where $C^N = \{x \in \mathcal{R}^n|h^N(x) = 0, -g^N(x) \in \mathcal{K}\}$. The knowledge needed to prove the second-order sufficient condition of (4.23) refer to [12].

Theorem 4.4. *Assume that (x^*, μ^*, λ^*) is a KKT point of (4.5)-(4.6), then $J\Phi^N_{NR}(x^*, \mu^*, \lambda^*)$ is non-singular if and only if*

(i) $\Lambda(x^*) \neq \phi$;

(ii) *The second-order sufficient condition (4.18) holds;*

(iii) $-\lambda^* \in intN_{\mathcal{K}}(-g^N(x^*))$ *holds;*

(iv) *The constraint non-degeneracy condition holds, i.e.,*

$$\begin{pmatrix} \nabla h^N(x^*)^T \\ -\nabla g^N(x^*)^T \end{pmatrix} \mathcal{R}^n + linT_{\{0_l\}\times\mathcal{K}}(h^N(x^*), -g^N(x^*)) = \mathcal{R}^l \times \mathcal{R}^m. \tag{4.24}$$

Proof. It is clear that

$$J\Phi_{NR}^N(x,\mu,\lambda) = \begin{pmatrix} J_xL^N(x,\mu,\lambda) & J_xh^N(x)^T & J_xg^N(x)^T \\ J_xh^N(x) & 0 & 0 \\ \Phi_1 & 0 & \Phi_2 \end{pmatrix}, \tag{4.25}$$

where

$$\Phi_1 = -\nabla g_{m_i}^N(x)\left(I - diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N(x) - \lambda_{m_i}; d_{m_i})\}_{i=1}^q\right), \tag{4.26}$$

$$\Phi_2 = diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N(x) - \lambda_{m_i}; d_{m_i})\}_{i=1}^q. \tag{4.27}$$

To simplify the proof of the theorem, in the following process, we let $J\Phi_{NR}^N := J\Phi_{NR}^N(x^*,\mu^*,\lambda^*)$, $Jh^N := Jh^N(x^*)$, $Jg^N := Jg^N(x^*)$, $J_xL^N := J_xL^N(x^*,\mu^*,\lambda^*)$, thus, we get

$$J\Phi_{NR}^N \begin{pmatrix} dx \\ d\mu \\ d\lambda \end{pmatrix} = \begin{pmatrix} J_xL^N dx + (Jh^N)^T d\mu + (Jg^N)^T d\lambda \\ Jh^N dx \\ M \end{pmatrix}, \tag{4.28}$$

where

$$\begin{aligned} M &= \left(I - diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N - \lambda_{m_i}^*; d_{m_i})\}_{i=1}^q\right)^T (-Jg_{m_i}^N)dx \\ &\quad + \left(diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N - \lambda_{m_i}^*; d_{m_i})\}_{i=1}^q\right)^T d\lambda \\ &= -Jg_{m_i}^N dx - \left(diag\{\Pi'_{\mathcal{K}^{m_i}}(-g_{m_i}^N - \lambda_{m_i}^*; d_{m_i})\}_{i=1}^q\right)^T (-Jg_{m_i}^N dx - d\lambda) \\ &= -Jg^N dx - \Pi'_{\mathcal{K}}(-g^N - \lambda^*; -Jg^N dx - d\lambda). \end{aligned} \tag{4.29}$$

Hence, we gain

$$J\Phi_{NR}^N \begin{pmatrix} dx \\ d\mu \\ d\lambda \end{pmatrix} = \begin{pmatrix} J_xL^N dx + (Jh^N)^T d\mu + (Jg^N)^T d\lambda \\ Jh^N dx \\ -Jg^N dx - \Pi'_{\mathcal{K}}(-g^N - \lambda^*; -Jg^N dx - d\lambda) \end{pmatrix}. \tag{4.30}$$

Notice that $J\Phi_{NR}^N(x,\mu,\lambda)$ is non-singular means that $J\Phi_{NR}^N(x,\mu\,\lambda)\begin{pmatrix} dx \\ d\mu \\ d\lambda \end{pmatrix} = 0$, is equivalent to $dx = 0, d\mu = 0, d\lambda = 0$, for any $(dx, d\mu, d\lambda)^T \in \mathcal{R}^n \times \mathcal{R}^l \times \mathcal{R}^m$. Firstly, from 2nd and 3nd of (4.30), we have

$$\begin{cases} Jh^N dx = 0, \\ -Jg^N dx = \Pi'_{\mathcal{K}}(-g^N - \lambda^*; -Jg^N dx - d\lambda). \end{cases} \tag{4.31}$$

which means $dx \in \mathcal{C}(x^*)$. On the other side, the first and second formulas of (4.30) mean

$$\left\langle J_x L^N dx, dx \right\rangle + \left\langle Jg^N dx, d\lambda \right\rangle = 0. \tag{4.32}$$

To continue, we cite the index sets:

$$\begin{aligned} I^* &= \left\{ i \middle| -g^N_{m_i}(x^*) \in \operatorname{int} \mathcal{K}^{m_i}, i = 1, \dots, q \right\}; \\ B^* &= \left\{ i \middle| -g^N_{m_i}(x^*) \in bd\mathcal{K}^{m_i}, g^N_{m_i}(x^*) \neq 0 \right\}; \\ Z^* &= \left\{ i | g^N_{m_i}(x^*) = 0 \right\}. \end{aligned} \tag{4.33}$$

Notice that

$$\mathcal{C}_{\mathcal{K}}(-g^N) = \left\{ d \in \mathcal{R}^n \middle| -Jg^N d \in T_{\mathcal{K}}(-g^N) \right\}, \tag{4.34}$$

and

$$T_{\mathcal{K}}(-g^N) = \left\{ d \middle| \begin{array}{l} -J\tilde{g}^N_{m_i} d - \frac{J\bar{g}^N_{m_i} d}{\tilde{g}^N_{m_i}} \geq 0, \qquad i \in B^* \\ -(J\tilde{g}^N_{m_i})^T d + J\bar{g}^N_{m_i} d \geq 0, i \in Z^* \end{array} \right\}. \tag{4.35}$$

Due to $\lambda \perp -g^N$, then

$$\lambda = \left\{ \lambda \middle| \begin{array}{l} \lambda_{m_i} = 0, i \in I^* \\ \lambda_{m_i} = \sigma(-\tilde{g}^N_{m_i}(x^*), \bar{g}^N_{m_i}(x^*)), \sigma > 0, i \in B^* \\ \lambda_{m_i} \in \operatorname{int} \mathcal{K}^{m_i}, i \in Z^* \end{array} \right\}. \tag{4.36}$$

In conclusion

$$\left[-g^N - \lambda^* \right]_i = \begin{cases} -g^N_{m_i}(x^*) \in int\mathcal{K}^{m_i}, i \in I^* \\ \left((1-\sigma)(-\tilde{g}^N_{m_i}(x^*)), (1+\sigma)(-\bar{g}^N_{m_i}(x^*)) \right), i \in B^* \\ \lambda_{m_i} \in \operatorname{int} \mathcal{K}^{m_i}, i \in Z^* \end{cases} \tag{4.37}$$

Indeed condition (iii) means

$$\mathcal{C}(x^*) = \left\{ d \middle| \begin{array}{l} Jh^N d = 0, -Jg^N_{m_i} d = 0, i \in Z^* \\ -Jg^N_{m_i} d \in T_{\mathcal{K}}(-g^N_{m_i}), \left\langle \lambda_{m_i}, -Jg^N_{m_i} d \right\rangle = 0, i \in B^* \end{array} \right\}, \tag{4.38}$$

since $\mathcal{C}(x^*)$ is a linear space, we get

$$\delta^* \left(\lambda \left| T^2_{\mathcal{K}} \left(-g^N, -Jg^N d \right) \right. \right) = \sum_{i \in B^*} \frac{\tilde{\lambda}_{m_i}}{-\tilde{g}^N_{m_i}} \left[\left\| J\tilde{g}^N_{m_i} dx \right\|^2 - \left\| J\bar{g}^N_{m_i} dx \right\|^2 \right]. \tag{4.39}$$

Case (1). If $i \in B^*$, by (4.30) and Lemma 2.3, we have

$$\begin{aligned} &\Pi'_{\mathcal{K}^m} \left(-g^N_{m_i} - \lambda^*_{m_i}; -Jg^N_{m_i} dx - d\lambda_{m_i} \right) \\ &= \frac{1}{2} \begin{pmatrix} 1 & w_i^T \\ w_i & \frac{2}{1+\sigma} I - \frac{1-\sigma}{1+\sigma} w_i w_i^T \end{pmatrix} \left(-Jg^N_{m_i} dx - d\lambda_{m_i} \right) \\ &= A_i(-Jg^N_{m_i} dx - d\lambda_{m_i}) = -Jg^N_{m_i} dx, \end{aligned} \tag{4.40}$$

where $\omega_i = \frac{-\bar{g}^N_{m_i}}{\|\bar{g}^N_{m_i}\|}$. Now we need to prove $dx \in \mathcal{C}(x^*)$ and

$$-J\tilde{g}^N_{m_i} dx \geq \frac{(\bar{g}^N_{m_i})^T J\bar{g}^N_{m_i} dx}{\|\bar{g}^N_{m_i}\|}. \tag{4.41}$$

From $-\tilde{g}^N_{m_i} = \|\bar{g}^N_{m_i}\|$, we gain

$$\lambda^*_{m_i} = \begin{pmatrix} -\sigma\tilde{g}^N_{m_i} \\ +\sigma\bar{g}^N_{m_i} \end{pmatrix} = -\sigma\tilde{g}^N_{m_i}\begin{pmatrix} 1 \\ -\omega_i \end{pmatrix}, \tag{4.42}$$

where $\|\omega_i\| = 1$, $\omega_i = \frac{\bar{g}^N_{m_i}}{\tilde{g}^N_{m_i}} = \frac{-\bar{g}^N_{m_i}}{\|\bar{g}^N_{m_i}\|}$. In summary, we gain

$${\lambda^*_{m_i}}^T A_i = (1 - \|w_i\|^2, {w_i}^T - \frac{2}{1+\sigma}{w_i}^T + \frac{1-\sigma}{1+\sigma}{w_i}^T\|w_i\|^2) = (0, 0). \tag{4.43}$$

By (4.43) and (4.40), we derive

$$\left\langle \lambda^*_{m_i}, -Jg^N_{m_i}dx \right\rangle = 0, \tag{4.44}$$

which implies $dx \in \mathcal{C}(x^*)$. On the other hand, by (4.40)

$$\begin{aligned} & A_i(-Jg^N_{m_i}dx - d\lambda_{m_i}) = -Jg^N_{m_i}dx \\ \Leftrightarrow\ & (A_i - I)(-Jg^N_{m_i}dx) = A_i d\lambda_{m_i} \\ \Leftrightarrow\ & (1, -{w_i}^T)\begin{pmatrix} -\frac{1}{2} & \frac{1}{2}{w_i}^T \\ \frac{1}{2}w_i & \frac{-\sigma}{1+\sigma}I - \frac{1}{2}\frac{1-\sigma}{1+\sigma}w_i{w_i}^T \end{pmatrix}\begin{pmatrix} -J\tilde{g}^N_{m_i}dx \\ -\nabla\bar{g}^N_{m_i}dx \end{pmatrix} \\ =\ & (1, -{w_i}^T)\begin{pmatrix} \frac{1}{2} & \frac{1}{2}{w_i}^T \\ \frac{1}{2}w_i & \frac{1}{1+\sigma}I - \frac{1}{2}\frac{1-\sigma}{1+\sigma}w_i{w_i}^T \end{pmatrix}\begin{pmatrix} d\tilde{\lambda}_{m_i} \\ d\bar{\lambda}_{m_i} \end{pmatrix}, \end{aligned} \tag{4.45}$$

hence, we deduce

$$(-1, \frac{1}{2}\omega_i^T + \frac{\sigma}{1+\sigma}\omega_i^T + \frac{1}{2}\frac{1-\sigma}{1+\sigma}\omega_i^T)\begin{pmatrix} -J\tilde{g}^N_{m_i}dx \\ -J\bar{g}^N_{m_i}dx \end{pmatrix} = 0, \tag{4.46}$$

equivalently

$$(-1, \omega_i^T)\begin{pmatrix} -J\tilde{g}^N_{m_i}dx \\ -J\bar{g}^N_{m_i}dx \end{pmatrix} = 0, \tag{4.47}$$

thus

$$-J\tilde{g}^N_{m_i}dx = \frac{(\bar{g}^N_{m_i})^T J\bar{g}^N_{m_i}dx}{\|\bar{g}^N_{m_i}\|}. \tag{4.48}$$

which implies (4.41) holds.

Case (2). If $i \in Z^*$, i.e., $g^N_{m_i}(x^*) = 0$, $\lambda_{m_i} \in int\mathcal{K}^{m_i}$. It is easy to obtain that

$$\Pi'_{\mathcal{K}^{m_i}}(0 - \lambda_{m_i}; -Jg^N_{m_i}dx - d\lambda_{m_i}) = -Jg^N_{m_i}dx, \tag{4.49}$$

which indicates $-Jg^N_{m_i}dx = 0$.

Case (3). If $i \in I^*$, i.e., $g^N_{m_i}(x^*) \in int\mathcal{K}^{m_i}$, $\lambda_{m_i} = 0$, which indicates

$$\Pi'_{\mathcal{K}^{m_i}}(-g^N_{m_i}; -Jg^N_{m_i}dx - d\lambda_{m_i}) = -Jg^N_{m_i}dx - d\lambda_{m_i} = -Jg^N_{m_i}dx, \tag{4.50}$$

thus $d\lambda_{m_i} = 0$.

Combining the above discussion, we acquire $dx \in \mathcal{C}(x^*)$ means

$$\begin{cases} Jg_{m_i}^N dx = 0, i \in Z^* \\ \tilde{g}_{m_i}^N J\tilde{g}_{m_i}^N dx = (\bar{g}_{m_i}^N)^T J\bar{g}_{m_i}^N dx, i \in B^* \end{cases} \tag{4.51}$$

Further, (4.30) is equivalent to

$$\begin{aligned} 0 &= \left\langle dx, J_x L^N dx + (Jh^N)^T d\mu + (Jg^N)^T d\lambda \right\rangle \\ &= \left\langle dx, J_x L^N dx \right\rangle - \Sigma_{i \in B^*} \left\langle -Jg_{m_i}^N dx, d\lambda_{m_i} \right\rangle, \end{aligned} \tag{4.52}$$

hence, for $i \in B^*$

$$\begin{aligned} &\left\langle -Jg_{m_i}^N dx, d\lambda_{m_i} \right\rangle = -J\tilde{g}_{m_i}^N dx d\tilde{\lambda}_{m_i} + \left\langle -\nabla \bar{g}_{m_i}^N dx, d\bar{\lambda}_{m_i} \right\rangle \\ &= J\tilde{g}_{m_i}^N dx \cdot \tfrac{(\bar{g}_{m_i}^N)^T}{\|\bar{g}_{m_i}^N\|} d\bar{\lambda}_{m_i} - \left\langle \nabla \bar{g}_{m_i}^N dx, d\bar{\lambda}_{m_i} \right\rangle \\ &= \tfrac{(\bar{g}_{m_i}^N)^T J\bar{g}_{m_i}^N dx}{\|\bar{g}_{m_i}^N\|^2} (\bar{g}_{m_i}^N)^T d\bar{\lambda}_{m_i} - \left\langle \nabla \bar{g}_{m_i}^N dx, d\bar{\lambda}_{m_i} \right\rangle \\ &= \left\langle (-J\bar{g}_{m_i}^N dx)^T \left[I - \tfrac{\bar{g}_{m_i}^N (\bar{g}_{m_i}^N)^T}{\|\bar{g}_{m_i}^N\|^2} \right], d\bar{\lambda}_{m_i} \right\rangle. \end{aligned} \tag{4.53}$$

By (4.45), we get

$$\begin{aligned} &\begin{pmatrix} \tfrac{1}{2} J\tilde{g}_{m_i}^N dx + \tfrac{1}{2} \tfrac{(\bar{g}_{m_i}^N)^T}{\|\bar{g}_{m_i}^N\|} J\bar{g}_{m_i}^N dx \\ \tfrac{1}{2} w_i \left(-J\tilde{g}_{m_i}^N dx - {w_i}^T J\bar{g}_{m_i}^N dx \cdot \tfrac{1-\sigma}{1+\sigma} \right) - \tfrac{\sigma}{1+\sigma} (-J\bar{g}_{m_i}^N dx) \end{pmatrix} \\ = &\begin{pmatrix} \tfrac{1}{2} d\tilde{\lambda}_{m_i} + \tfrac{1}{2} {w_i}^T d\bar{\lambda}_{m_i} \\ \tfrac{1}{2} w_i (d\tilde{\lambda}_{m_i} - \tfrac{1-\sigma}{1+\sigma} {w_i}^T d\bar{\lambda}_{m_i}) + \tfrac{1}{1+\sigma} d\bar{\lambda}_{m_i} \end{pmatrix}, \end{aligned} \tag{4.54}$$

from (4.48) and (4.54), we gain

$$\begin{aligned} &\tfrac{1}{2} w_i \left(-J\tilde{g}_{m_i}^N dx - {w_i}^T J\bar{g}_{m_i}^N dx \cdot \tfrac{1-\sigma}{1+\sigma} \right) + \tfrac{\sigma}{1+\sigma} J\bar{g}_{m_i}^N dx \\ &= \tfrac{1}{2} w_i \left(-J\tilde{g}_{m_i}^N dx + \tfrac{1-\sigma}{1+\sigma} J\tilde{g}_{m_i}^N dx \right) + \tfrac{\sigma}{1+\sigma} J\bar{g}_{m_i}^N dx \\ &= \tfrac{\sigma}{1+\sigma} \left(w_i (-J\tilde{g}_{m_i}^N dx) + J\bar{g}_{m_i}^N dx \right), \end{aligned} \tag{4.55}$$

and

$$\begin{aligned} &\tfrac{1}{2} w_i (d\tilde{\lambda}_{m_i} - \tfrac{1-\sigma}{1+\sigma} {w_i}^T d\bar{\lambda}_{m_i}) + \tfrac{1}{1+\sigma} d\bar{\lambda}_{m_i} \\ &= \tfrac{1}{2} w_i (d\tilde{\lambda}_{m_i} + \tfrac{1-\sigma}{1+\sigma} d\tilde{\lambda}_{m_i}) + \tfrac{1}{1+\sigma} d\bar{\lambda}_{m_i} \\ &= \tfrac{1}{1+\sigma} (w_i d\tilde{\lambda}_{m_i} + d\bar{\lambda}_{m_i}). \end{aligned} \tag{4.56}$$

By (4.54), (4.55) and (4.56), we get

$$\frac{1}{1+\sigma} (w_i d\tilde{\lambda}_{m_i} + d\bar{\lambda}_{m_i}) = \frac{\sigma}{1+\sigma} (w_i (-J\tilde{g}_{m_i}^N dx) + J\bar{g}_{m_i}^N dx), \tag{4.57}$$

which means

$$w_i d\tilde{\lambda}_{m_i} + d\bar{\lambda}_{m_i} = -\sigma \left(w_i J\tilde{g}_{m_i}^N dx - J\bar{g}_{m_i}^N dx \right). \tag{4.58}$$

Notice that

$$\begin{aligned} w_i d\tilde{\lambda}_{m_i} + d\bar{\lambda}_{m_i} &= (I - w_i {w_i}^T) d\bar{\lambda}_{m_i} \\ &= \left(I - \tfrac{\bar{g}_{m_i}^N (\bar{g}_{m_i}^N)^T}{\|\bar{g}_{m_i}^N\|^2} \right) d\bar{\lambda}_{m_i}, \end{aligned} \tag{4.59}$$

thus, from (4.53), (4.58) and (4.59), we have

$$\begin{aligned} &\left\langle -Jg_{m_i}^N dx, d\lambda_{m_i}\right\rangle = \left\langle -J\bar{g}_{m_i}^N dx, \left(I - \frac{\bar{g}_{m_i}^N(\bar{g}_{m_i}^N)^T}{\left\|\bar{g}_{m_i}^N\right\|^2}\right) d\tilde{\lambda}_{m_i}\right\rangle \\ &= \sigma\left(\left\langle -J\bar{g}_{m_i}^N dx, w_i(-J\tilde{g}_{m_i}^N dx)\right\rangle - \left\|J\bar{g}_{m_i}^N dx\right\|^2\right) \\ &= \sum_{i\in B^*} \frac{\tilde{\lambda}_{m_i}}{-\bar{g}_{m_i}^N}\left(\left\|J\tilde{g}_{m_i}^N dx\right\|^2 - \left\|J\bar{g}_{m_i}^N dx\right\|^2\right) \\ &= \delta^*\left(\lambda \left| T_{\mathcal{K}}{}^2(-g^N; -Jg^N dx)\right.\right), \end{aligned} \tag{4.60}$$

besides, by (4.60) and (4.32), we have

$$\left\langle J_x L^N dx, dx\right\rangle - \delta^*\left(\lambda \left| T_{\mathcal{K}}{}^2(-g^N; -Jg^N dx)\right.\right) = 0. \tag{4.61}$$

Since the second-order sufficient condition holds, we deduce $dx = 0$, further, from (4.30) we reach

$$\nabla h^N d\mu + \nabla g^N d\lambda = 0, \tag{4.62}$$

the third expression of (4.30) and condition (iv) guarantee $d\mu = 0$, $d\lambda = 0$, therefore, $J\Phi_{NR}^N(x^*, \mu^*, \lambda^*)$ is nonsingular. □

5 The algorithm and numerical experiments

In Chapter 4, we have proved the non-singularity of the matrix $J\Phi_{NR}^N(x, \mu, \lambda)$. Based on this, this Chapter will construct Newton's algorithm and give two numerical examples to test the effectiveness of the algorithm.

In order to solve $\Phi_{NR}^N(x, \mu, \lambda) = 0$, define function

$$\psi^N(z) = \frac{1}{2}\|\Phi_{NR}^N(x, \mu, \lambda)\|^2. \tag{5.1}$$

By a simple proof, it can be obtained that under the condition that $f(x, \xi)$, $h(x, \eta)$, $g(x, \zeta)$ are continuously differentiable with respect to x, and $J\Phi_{NR}^N(x, \mu, \lambda)$ is non-singular, then every stationary point of $\psi^N(z)$ is a KKT point of SAA problem.

Algorithm 5.1

Step 0 Given $z^0 = (x^0, \mu^0, \lambda^0) \in R^n \times R^l \times R^m$, let $\sigma > 0$, $s > 1$, and $\gamma \in (0, 1)$. Set $k := 0$;

Step 1 If $z^k = (x^k, \mu^k, \lambda^k)$ is a stable point of ψ^N, stop; otherwise, go to Step 2.

Step 2 Choose an element $H^k \in J\Phi_{NR}^N(z^k)$, and find a direction d^k satisfying

$$\Phi_{NR}^N(z^k) + H^k d^k = 0. \tag{5.2}$$

If (5.2) is unsolvable, or does not satisfy the following condition

$$\nabla\psi^N(z^k)^T d^k \le -\sigma\|d^k\|^s, \tag{5.3}$$

let $d^k := -\nabla\psi^N(z^k)$;

Step 3 Find the smallest non-negative integer ι_k, let $\iota = \iota_k$ such that

$$\psi^N(z^k + 2^{-\iota}d^k) \le \psi^N(z^k) + \gamma 2^{-\iota}\nabla\psi^N(z^k)^T d^k. \tag{5.4}$$

Let $\varrho_k = 2^{-\iota}$;

Step 4 Let $z^{k+1} = z^k + \varrho_k d^k$, $k = k+1$, return to Step 2.

Based on Algorithm 5.1, two examples are gave and the experimental environment is Matlab2018a, Intel Core i5, CPU 2.2GHz. In example 5.1, ξ follows a normal distribution with mean 1 and variance 2. In example 5.2, ξ follows an exponential distribution with $\lambda = 1$. ζ is a random distribution on the interval [0,1], and sample size $N = 1000$.

Example 5.1. Consider the optimization problem

$$\begin{aligned} &\min F(x,\xi) = \left(\frac{1}{2}\xi(x_1^2 - 8x_1), \frac{3}{2}x_2^2 - 2x_2, \xi(-cosx_3 - x_3), \xi(e^{x_4} - 2x_4), x_5^2 + 3x_5\right) \\ &s.t.\ \ C = \{x \in R^5 | - g(x,\zeta) = \zeta x \in \mathcal{K}^5\} \end{aligned} \tag{5.5}$$

Let $f(x,\xi) = \nabla F(x,\xi) = (\xi x_1 - 4\xi, 3x_2 - 2, \xi sinx_3 - \xi, \xi e^{x_4} - 2\xi, 2x_5 + 3)^T$, then the SSOCCVI problem: find $x \in C$, such that

$$\begin{aligned} &\langle E[f(x,\xi)], y - x\rangle \geq 0, \quad \forall y \in C, \\ &C = \{x \in R^5 | - E[g(x,\zeta)] \in \mathcal{K}^5\}. \end{aligned} \tag{5.6}$$

The SAA problem: find $x \in C^N$, such that

$$\begin{aligned} &\langle \frac{1}{N}\sum_{i=1}^{N} f(x,\xi_i), y - x\rangle \geq 0, \quad \forall y \in C^N, \\ &C^N = \{x \in R^5 | - \frac{1}{N}\sum_{i=1}^{N} g(x,\zeta_i) \in \mathcal{K}^5\}. \end{aligned} \tag{5.7}$$

According to the analysis in Chapter 4, The KKT system of (5.7) is equal to $\Phi_{NR}^N(x,\mu,\lambda) = 0$, expressed by

$$\Phi_{NR}^N(x,\mu,\lambda) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N f(x,\xi_i) - \frac{1}{N}\sum_{i=1}^N \zeta_i x \\ -\frac{1}{N}\sum_{i=1}^N \zeta_i x - \Pi_{\mathcal{K}}(-\frac{1}{N}\sum_{i=1}^N \zeta_i x - \lambda) \end{pmatrix} = 0. \tag{5.8}$$

Select initial point $x_0 = (1,0,0,0,0)^T$, $\lambda_0 = (1,0,0,0,0)^T$, let parameters $\sigma = 0.01$, $\gamma = 0.45$. The solution result is shown in Figure 1, and it can be acquired that with the increase of the number of iterations, the curve trajectory gradually become stable, after the 39th iteration, the optimal solution $x^* = (4.0000, 0.6667, 1.5634, 0.6931, -1.5000)^T$ is found, obviously, $x^* \in \mathcal{K}^5$. Further, denote $B := \frac{1}{N}\sum_{i=1}^N \xi_i$,

$$\begin{aligned} J_x L^N(x,\mu,\lambda) &= \frac{1}{N}\sum_{i=1}^{N} f(x,\xi_i) + (J_x \frac{1}{N}\sum_{i=1}^{N}\zeta_i x)^T\lambda \\ &= \frac{1}{N}\sum_{i=1}^{N} f(x,\xi_i) + \frac{1}{N}\sum_{i=1}^{N}\zeta_i (J_x x)^T\lambda \\ &= \begin{pmatrix} B & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & Bcosx_3 & 0 & 0 \\ 0 & 0 & 0 & Be^{x_4} & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \end{aligned} \tag{5.9}$$

By $\xi \sim N(1,2)$, and $B \sim N(1, \frac{4}{N})$, which means eigenvalues of $J_x L^N(x,\mu,\lambda)$ are greater than zero at x^*, thus, $J_x L^N(x,\mu,\lambda)$ is positive definite, which indicates the second-order sufficient condition holds at x^*.

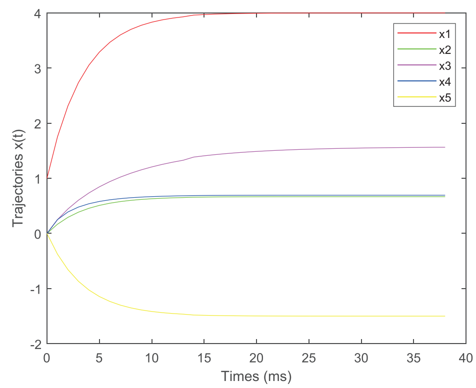

Figure 1: Example 5.1

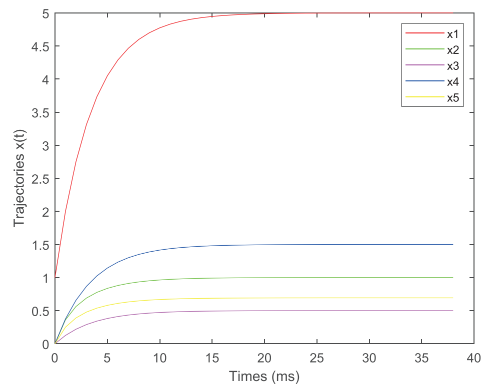

Figure 2: Example 5.2

Example 5.2. Consider the SSOCCVI problem: find $x \in C$, such that

$$\begin{aligned} &\langle E[f(x,\xi)], y-x\rangle \geq 0, \quad \forall y \in C, \\ &C = \{x \in R^5 | -g(x,\zeta) = \zeta x \in \mathcal{K}^5\}, \end{aligned} \tag{5.10}$$

where $f(x,\xi) = (x_1 - 5, \xi 2^{x_2} - 2\xi, 2\xi x_3 - \xi, 2x_4 - 3, \xi e^{x_5} - 2\xi)^T$.

The SAA problem: find $x \in C^N$, such that

$$\begin{aligned} &\langle \frac{1}{N}\sum_{i=1}^{N} f(x,\xi_i), y-x\rangle \geq 0, \quad \forall y \in C^N, \\ &C^N = \{x \in R^5 | -\frac{1}{N}\sum_{i=1}^{N} g(x,\zeta_i) \in \mathcal{K}^5\}. \end{aligned} \tag{5.11}$$

Similarly, the KKT system of (5.11) is equal to the following equation

$$\Phi_{NR}^N(x,\mu,\lambda) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N f(x,\xi_i) - \frac{1}{N}\sum_{i=1}^N \zeta_i x \\ -\frac{1}{N}\sum_{i=1}^N \zeta_i x - \Pi_{\mathcal{K}}(-\frac{1}{N}\sum_{i=1}^N \zeta_i x - \lambda) \end{pmatrix} = 0. \tag{5.12}$$

Select initial point $x_0 = (1,0,0,0,0)^T$, $\lambda_0 = (1,0,0,0,0)^T$, let parameters $\sigma = 0.01$, $\gamma = 0.45$. The result is shown in Figure 2, we can get: as the number of iterations increases, the feasible solution keeps going up until it flattens out, at last, the optimal solution $x^* = (5.0000, 1.0000, 0.5000, 1.5000, 0.6931)^T$ is obtained, obviously, $x^* \in \mathcal{K}^5$. Denote $Y := \frac{1}{N}\sum_{i=1}^N \xi_i$,

$$\begin{aligned} J_x L^N(x,\mu,\lambda) &= \frac{1}{N}\sum_{i=1}^N f(x,\xi_i) + (J_x \frac{1}{N}\sum_{i=1}^N \zeta_i x)^T \lambda \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & Y(2^{x_2} ln2) & 0 & 0 & 0 \\ 0 & 0 & 2Y & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & Ye^{x_5} \end{pmatrix} \end{aligned} \tag{5.13}$$

Due to $\xi \sim Exp(1)$, then $Y \sim N(1, \frac{1}{N})$, obviously, $J_x L^N(x, \mu, \lambda)$ is positive definite. Hence, the second-order sufficient condition is satisfied at x^*.

6 Conclusion

The main goal of this paper is to explore second-order sufficient condition for stochastic second-order cone constrained variational inequality problem. To this end, the following work has been done. In order to solve the difficulty that the expected value is not easy to obtain in practice, we first use the SAA method to approximate the original problem as the SAA problem, and give the rationality. In addition, the Karush-Kuhn-Tucker condition is derived for the SAA problem and it is transformed into a system of equations $\Phi^N_{NR}(x, \mu, \lambda) = 0$. Further, the second-order sufficient condition for the SAA problem is explored, and the non-singularity of $J\Phi^N_{NR}(x, \mu, \lambda)$ is proved. In the end, based on the second-order sufficient condition, we design Newton's algorithm and apply it to numerical examples, the numerical results show that the algorithm is effective.

References

[1] G. Gurkan, A.Y. Ozge and S.M. Robinson, Sample path solution of stochastic variational inequality inequalities, with applications to option pricing, in: Proceeding of the 28th Conference on Winter Simulation-WSC 96, J.M. Charnes D. Morrice, D.T. Brunner, J.J. Swai (eds), 1996.

[2] A. Shapiro, D. Dentcheva and A. Ruszczyński, *Lectures on Stochastic Programming: Modeling and Theory*, SIAM, 2009.

[3] H.F. Xu, Sample average approximation methods for a class of stochastic variational inequality problems, *Asia Pacific Journal of Operational Research* 27 (2010) 103–119.

[4] H. Robbins and S. Monro, A stochastic approximation method, *The Annals of Mathematical Statistics* 22 (1951) 400–407.

[5] K.L. Chuang, On a stochastic approximation method, *The Annals of Mathematical Statistics* 25 (1954) 463–483.

[6] S. Chen, L.P. Pang, F.F. Guo and Z.Q. Xia, Stochastic methods based on Newton method to the stochastic variational inequality problem with constraint conditions, Mathematical and Computer Modelling **55** (2012) 779–784.

[7] M.J. Luo and G.H. Lin, Sample average approximation method for solving a deterministic formulation for box constrained stochastic variational inequality problems, *Asia-Pacific Journal of Operational Research* 29 (2012): 1250014.

[8] S. Chen, L.P. Pang, X.F. Ma and D. Li, SAA method based on modified Newton method for stochastic variational inequality with second-order cone constraints and application in portfolio optimization, *Mathematiocal Methods of Operations Research* 84 (2016) 129–154.

[9] J.H. Sun, J.S. Chen and C.H. Ko, Neural networks for solving second-order cone constrained variational inequality problem, *Computational Optimization and Applications* 51 (2012) 623–648.

[10] M. Fukushima, Z.Q. Luo and P. Tseng, Smoothing functions for second-order cone complimentarity problems, *SIAM Journal on Optimization*12 (2002) 436–460.

[11] R.T. Rockafellar and R.J-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin Heidelberg, 1998.

[12] J.F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Spring-Verlag, New York, 2000.

[13] H.F. Xu, Sample average approximation methods for a class of stochastic variational inequality problems, *Asia Pacific Journal of Operational Research* 27 (2010) 103–119.

[14] Z.P. Yang, J. Zhang, X.D. Zhu, and G.H. Lin, Infeasible interior-point algorithms based on sampling average approximations for a class of stochastic complementarity problems and their applications, *Journal of Computational and Applied Mathematics* 352 (2019) 382–400.

[15] H.Y. Jiang and H.F. Xu, Stochastic approximation approaches to the stochastic variational inequality problem, *IEEE Transactions on Automatic Control* 53 (2008) 1462–1475.

[16] G.H. Lin, M.J. Luo, and D.L. Zhang, Stochastic second-order cone complementarity problems: expected residual minimization formulation and its applications, *Mathematical Programming* 165 (2017) 197–233.

[17] J.H. Sun, W.C. Fu, J.H. Alcantara, and J.S. Chen, A neural network based on the metric projector for solving SOCCVI problem, *IEEE Transaction on Network and Learning Systems* 32 (2021) 2886–2900.

[18] G.H. Lin, M.J. Luo, and J. Zhang, Smoothing and SAA method for stochastic programming problems with non-smooth objective and constraints, *Journal of Global Optimization* 66 (2016) 487–510.

[19] F. Facchinei and J.S. Pang, *Finite Dimensional Variational Inequalities and Complementarity Problems, Volume I*, Springer-verlag, New York, 2003.

[20] U. Faraut and A. Koráanyi, *Analysis on Symmetric Cones*, Oxford University Press, Oxford, UK, 1994.

[21] A. Shapiro and H.F. Xu, Stochastic mathematical programs with equilibrium constraints, modeling and sample average approximation, *Optimization* 57 (2008) 395–418.

[22] P. Mörters, Large deviation theory and applications, Lecture Notes on Large Deviation Theory, University of Bath, Math. Sciences Dept, 2008.

Manuscript received 7 November 2021
revised 24 April 2022
accepted for publication 14 June 2022

Li Wang
School of Science
Shenyang Aerospace University Shenyang 11036, China
E-mail address: liwang211@gmail.com

Mengdi Zheng

School of Science
Shenyang Aerospace University Shenyang 11036, China
E-mail address: ipurple1997@163.com

Juhe Sun

School of Science
Shenyang Aerospace University Shenyang 11036, China
E-mail address: juhesun@163.com

Hao Liu

School of Science
Shenyang Aerospace University Shenyang 11036, China
E-mail address: 1722993580@qq.com