Expected residual minimization method for uncertain variational inequality problems

Cunlin Li[a], Zhifu Jia[b,*], Lin Zhang[b]

[a]*School of Management, North Minzu University, Yinchuan, Ningxia 750021, China.*

[b]*School of Mathematics and Information Science, North Minzu University, Yinchuan, Ningxia 750021, China.*

Communicated by Y. H. Yao

Abstract

This paper considers an uncertain variational inequality problem (UVIP). We first establish UVIP as an optimization problem (ERM model) which minimizes the expected residual of the so-called regularized gap function. Then, we make some assumptions about a UVIP subclass in which the function involved is affine. Thus the priority in our paper is to discuss the properties of the ERM problem and comprehensive convergence analysis under uncertainty theory. In the end, we make a conclusion. ©2017 All rights reserved.

Keywords: Uncertain variational inequalities, uncertainty theory, properties of the ERM problem, convergence.
2010 MSC: 49J53, 49J40, 65K10, 90C99.

1. Introduction

The variational inequality problem (VIP for short) is such important in mathematics that trigger many scholars to research. nowadays, people mainly studies three aspects of VIP including solution to the existence and uniqueness, effective solution algorithms and application such as optimization theory, economics, engineering, game theory and networks and the like [4, 10–14]. From the classical VIP of R^n, we here consider the finite dimensional Euclidean space R^n, finding a point $x \in S \subset R^n$ such that

$$(y - x)^T f(x) \geqslant 0, \quad x \in S,$$

where S is a nonempty closed convex subset of R^n, and $f : S \to R^n$ is a mapping. If there exists an element x^* satisfying the variational inequality, then x^* is called a solution of it for every $y \in S$, which is denoted by $VIP(f, S)$. The set of solutions to this problem is denoted by $Sol(S, f)$. $VIP(f, S)$ is transformed to the following complementarity problem:

$$x^* \geqslant 0, \quad f(x^*) \geqslant 0, \quad (x^*)^T f(x^*) = 0.$$

This problem has been used to study various equilibrium models in economics, operations research, transportation and regional science and many algorithms have been developed. For more details, see the monograph [4] and the references therein. In particular, Fukushima [5, 6] defined a regularized gap function

$$g(x) := \max_{y \in S}\{(x-y)^\mathsf{T} f(x) - \frac{\alpha}{2}\|x-y\|_G^2\},$$

where α is a positive parameter, G is an $n \times n$ symmetric positive-definite matrix, and $\|.\|_G$ means the G-norm defined by $\|x\|_G = \sqrt{x^\mathsf{T} G x}$ for $x \in R^n$. It has been shown that $g(x) \geqslant 0$, for every $x \in S$, for $x \in S$, $g(x) = 0$ if and only if x solves VIP(f, S). Based on these facts, VIP(f, S) is transformed into the following minimization problem:

$$\begin{aligned} \min \ & g(x), \\ \text{s.t. } & x \in S. \end{aligned} \tag{1.1}$$

As stated in [5, 6], $H(x) := \mathrm{Proj}_{S,G}(x - \alpha^{-1}G^{-1}f(x))$ is the unique solution of $g(x)$, where $\mathrm{Proj}_{S,G}$ denotes the projection operator onto S under the G-norm. Moreover, if f is continuously differentiable, then g is also continuously differentiable and

$$\nabla g(x) = f(x) - [\nabla_x f(x) - \alpha G](H(x) - x).$$

Note that what the underlying mapping f in (1.1) reflects is deterministic, it does not involve uncertainties. However, there are many instances in practice that the underlying problem not only involves deterministic data, but also contains some uncertain factors in those data. For example, during the transmission of one signal, say sound signal from the base station through relays to the user, it may be interfered by other signals. Therefore, many researchers make a great deal of research for the stochastic variational inequality problem (SVIP for short) in which they can reflect those uncertainties. Finding a vector $x^* \in S \subset R^n$ such that

$$(x-x^*)^\mathsf{T} F(x^*, \omega) \geqslant 0, \quad \forall x \in S, \ \omega \in \Omega \text{ a.s.,}$$

where Ω is the underlying sample space, $F : R^n \times \Omega \to R^n$ is a mapping, and a.s. is the abbreviation for almost surely under the given probability measure. SVIP based on probability theory was studied in references such as [1], etc. And its applications can be found in inventory or pricing competition among several firms that provide substitutable goods or services, a supply chain network model, stochastic user equilibrium traffic network and wireless network. A basic premise of using probability theory is that the actual frequency is close enough to probability. In other words, if we want to use probability we must get a mass of historical data. However in some cases, it is very difficult for us to get enough data such that we estimate the probability distribution through statistics. For example, it is impossible for us to estimate a probability distribution of a new stock or devastating military experiments. In this situation, it makes sense to invite some experts to evaluate the belief degree that the event will happen. Here the belief degree represents the degree that we believe the indeterminacy quantity falls into the left side of a given point about the indeterminacy phenomena. There also exist the cases that people's viewpoints may vary from person to person for the same event; even for the same person, his/her view to the event may vary from time to time. Hence, in order to deal with those uncertainty phenomena, Liu [8] found a new theory - uncertainty theory which is quite different from probability theory. Chen and Zhu [3] presented the uncertain variational inequality problem (UVIP for short) based on uncertainty theory. Finding a vector $x^* \in S \subset R^n$ such that

$$(x-x^*)^\mathsf{T} F(x^*, \xi) \geqslant 0, \quad \forall x \in S, \ \xi \in \Xi \text{ a.s.,} \tag{1.2}$$

where Ξ is the uncertain space, $F : R^n \times \Xi \to R^n$ is a mapping, and a.s. is the abbreviation for almost surely under the given uncertain measure. There is no doubt that both of them are natural extensions of

deterministic VIP. Therefore, in this paper, our focus is on developing model to solve UVIP. The remainder of this paper is organized as follows. Some basic concepts and properties were reviewed in Section 2, including uncertain variable, uncertain expectation, inverse uncertainty distribution. In Section 3, the ERM model is established based on uncertainty theory. In Section 4, we give some properties including the differentiability of the function θ and some results for level sets and error bounds. Then, the convergence analysis is investigated in Section 5. Finally, we discuss conclusions in Section 6.

2. Preliminaries

In this section, we will introduce some fundamental concepts and properties concerning uncertain variables, expectation, inverse uncertainty distribution. Let Γ be a nonempty set, and $\mathscr{L}$ a σ-algebra over Γ. Each element Λ in $\mathscr{L}$ is called an event and assigned a number $\mathscr{M}\{\Lambda\}$ to indicate the belief degree with which we believe Λ will happen. In order to deal with belief degrees rightly, Liu put forward the following three axioms:

Axiom 1. (Normality Axiom) $\mathscr{M}\{\Gamma\} = 1$ for the universal set Γ;

Axiom 2. (Duality Axiom) $\mathscr{M}\{\Lambda\} + \mathscr{M}\{\Lambda^c\} = 1$ for any event Λ;

Axiom 3. (Subadditivity Axiom) For every countable sequence of events $\Lambda_1, \Lambda_2, \cdots$, we have

$$\mathscr{M}\left\{\bigcup_{i=1}^{\infty} \Lambda_i\right\} \leqslant \sum_{i=1}^{\infty} \mathscr{M}\{\Lambda_i\}.$$

Definition 2.1 (Liu [8]). The set function $\mathscr{M}$ is called an uncertain measure if it satisfies the normality, duality, and subadditivity axioms.

The triplet $(\Gamma, \mathscr{L}, \mathscr{M})$ is called an uncertainty space. Furthermore, the product uncertain measure on the product σ-algebra $\mathscr{L}$ was defined by Liu as follows:

Axiom 4. (Product Axiom) Let $(\Gamma_k, \mathscr{L}_k, \mathscr{M}_k)$ be uncertainty spaces for $k = 1, 2, \cdots$. The product uncertain measure $\mathscr{M}$ is an uncertain measure satisfying

$$\mathscr{M}\left\{\prod_{k=1}^{\infty} \Lambda_k\right\} = \bigwedge_{k=1}^{\infty} \mathscr{M}_k\{\Lambda_k\},$$

where Λ_k are arbitrary events chosen from $\mathscr{L}_k$ for $k = 1, 2, \cdots$, respectively.

Definition 2.2 (Liu [8]). An uncertain variable is a measurable function ξ from an uncertainty space $(\Gamma, \mathscr{L}, \mathscr{M})$ to the set of real numbers, for any Borel set B of real numbers, the set

$$\{\xi \in B\} = \{\gamma \in \Gamma \mid \xi(\gamma) \in B\},$$

is an event.

Theorem 2.3. *Let $\xi_1, \xi_2, \cdots, \xi_n$ be uncertain variables, and f be a real-valued measurable function. Then $\xi = f(\xi_1, \xi_2, \cdots, \xi_n)$ is an uncertain variable defined by*

$$\xi(\gamma) = f(\xi_1(\gamma), \xi_2(\gamma), \cdots, \xi_n(\gamma)), \quad \forall \gamma \in \Gamma.$$

Definition 2.4. (Liu [8]) Suppose ξ is an uncertain variable. Then the uncertainty distribution of ξ is defined by

$$\Phi(x) = \mathscr{M}\{\xi \leqslant x\}$$

for any real number x.

An uncertainty distribution $\Phi(x)$ is said to be regular if its inverse function $\Phi^{-1}(\alpha)$ exists and is unique for each $\alpha \in (0,1)$. Inverse uncertainty distribution plays an important role in the operations of independent uncertain variables. In the following, the concept of inverse uncertainty distribution will be presented.

Definition 2.5 (Liu [8]). Suppose ξ is an uncertain variable with regular uncertainty distribution $\Phi(x)$. Then the inverse function $\Phi^{-1}(\alpha)$ is called the inverse uncertainty distribution of ξ.

Theorem 2.6 (Liu [9]). *A function $\Phi^{-1}(\alpha) : (0,1) \to \mathfrak{R}$ is an inverse uncertainty distribution if and only if it is a continuous and strictly increasing function with respect to α.*

The operational law of independent uncertain variables was given by [2] in order to calculate the uncertainty distribution of a strictly increasing or decreasing function of uncertain variables. Before introducing the operational law, the concept of independence of uncertain variables is presented as follows.

Definition 2.7 (Liu [9]). The uncertain variables $\xi_1, \xi_2, \cdots, \xi_n$ are said to be independent if

$$\mathcal{M}\left\{\bigcap_{i=1}^{n}(\xi_i \in B_i)\right\} = \bigwedge_{i=1}^{n}\mathcal{M}\{\xi_i \in B_i\}$$

for any Borel sets $B_1, B_2, \cdots, B_n$.

Theorem 2.8 (Liu [9]). *Let $\xi_1, \xi_2, \cdots, \xi_n$ be independent uncertain variables with continuous uncertainty distributions $\Phi_1, \Phi_2, \cdots, \Phi_n$, respectively. If the function $f(x_1, x_2, \cdots, x_n)$ is strictly increasing with respect to $x_1, x_2, \cdots, x_m$ and strictly decreasing with respect to $x_{m+1}, x_{m+2}, \cdots, x_n$, then the uncertain variable*

$$\xi = f(\xi_1, \xi_2, \cdots, \xi_n),$$

has an uncertainty distribution

$$\Phi(x) = \sup_{f(x_1,x_2,\cdots,x_n)=x}\left(\min_{1\leqslant i\leqslant m}\Phi_i(x)\wedge\min_{m+1\leqslant i\leqslant n}(1-\Phi_i(x))\right).$$

Moreover, if $\Phi_1, \Phi_2, \cdots, \Phi_n$ are regular, then ξ has an inverse uncertainty distribution

$$\Phi^{-1}(\alpha) = f(\Phi_1^{-1}(\alpha), \cdots, \Phi_m^{-1}(\alpha), \Phi_{m+1}^{-1}(1-\alpha)\cdots, \Phi_n^{-1}(1-\alpha)).$$

For ranking uncertain variables, the concept of expected value was proposed by Liu [8] as follows.

Definition 2.9 (Liu [8]). Let ξ be an uncertain variable. Then the expected value of ξ is defined by

$$E[\xi] = \int_{0}^{+\infty}\mathcal{M}\{\xi \geqslant x\}dx - \int_{-\infty}^{0}\mathcal{M}\{\xi \leqslant x\}dx,$$

provided that at least one of the two integrals is finite.

Theorem 2.10 (Liu [8]). *Let ξ be an uncertain variable with uncertainty distribution Φ. If the expected value exists, then*

$$E[\xi] = \int_{-\infty}^{+\infty}t d\Phi(t).$$

3. Establishment of ERM model

We considered an expected residual minimization uncertain affine variational inequality problem on the basis of uncertainty theory. Following [5, 6], we define a regularized gap function.

Definition 3.1. A regularized gap function satisfying $g : R^n \times \Xi \to R$ for uncertain variational inequality problem (denoted by UVIP(F,S)) is defined such as

$$g(x, \xi) := \max_{y\in S}\{(x-y)^{\mathsf{T}}F(x,\xi) - \frac{\alpha}{2}\|x-y\|_G^2\}. \tag{3.1}$$

Lemma 3.2. *Let S be a closed convex set in R^n. Then for each $f(x, \xi) \in R^n$, where $f : R^n \times \Xi \to R^n$, there is a unique point $z \in S$ such that $\|f(x, \xi) - z\|_G \leqslant \|f(x, \xi) - y\|_G$, for all $y \in S$, z is the orthogonal projection of x on the set S with respect to the Euclidean G norm, that is,*

$$z = \mathrm{Proj}_{S,G} f(x, \xi) = \arg \min_{y \in S} \|f(x, \xi) - y\|_G.$$

Proof. Let (x, ξ) be fixed and let $t \in S$. Minimizing $\|f(x, \xi) - y\|_G$ over all $y \in S$ is equivalent to minimizing the same function over all $y \in S$ such that $\|f(x, \xi) - y\|_G \leqslant \|f(x, \xi) - t\|_G$, S is a compact set. The function g defined by $g(y) = \|f(x, \xi) - y\|_G^2$ is continuous. Existence of a minimizing z follows because a continuous function on a compact set always attains its minimum. To prove that z is unique, we only need to prove that the square of the Euclidean G norm is a strictly convex function. Because of its Hessian matrix is equal to 2G. Here G is an $n \times n$ symmetric positive-definite matrix. Hence, g is strictly convex and its minimum is unique. $\square$

Theorem 3.3. *The problem $\max\limits_{y \in S}\{(x - y)^\mathsf{T} F(x, \xi) - \frac{\alpha}{2}\|x - y\|_G^2\}$ is essentially equivalent to the below problem*

$$\min_{y \in S} \|y - (x - \alpha^{-1} G^{-1} F(x, \xi)\|_G^2\}.$$

Proof.

$$\max_{y \in S}\{(x - y)^\mathsf{T} F(x, \xi) - \frac{\alpha}{2}\|x - y\|_G^2\}$$

$$\Leftrightarrow \max_{y \in S}\{2\alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) - \|x - y\|_G^2\}$$

$$\Leftrightarrow \max_{y \in S}\{2\alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) - \|x - y\|_G^2 - \alpha^{-2} F^\mathsf{T}(x, \xi) G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{\|x - y\|_G^2 - 2\alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) + \alpha^{-2} F^\mathsf{T}(x, \xi) G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{\|x - y\|_G^2 - 2\alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) + \alpha^{-2} F^\mathsf{T}(x, \xi) G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{\|x - y\|_G^2 - \alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) - \alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) + \alpha^{-2} F^\mathsf{T}(x, \xi) G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{(x - y)^\mathsf{T} G(x - y) - \alpha^{-1}(x - y)^\mathsf{T} F(x, \xi) - \alpha^{-1} F^\mathsf{T}(x, \xi)(x - y) + \alpha^{-2} F^\mathsf{T}(x, \xi) G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{(x - y)^\mathsf{T} G - \alpha^{-1} F^\mathsf{T}(x, \xi)\}\{(x - y) - \alpha^{-1} G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{(x - y)^\mathsf{T} - \alpha^{-1} F^\mathsf{T}(x, \xi) G^{-1}\} G\{(x - y) - \alpha^{-1} G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S}\{(x - y)^\mathsf{T} - (\alpha^{-1} G^{-1} F(x, \xi))^\mathsf{T}\} G\{(x - y) - \alpha^{-1} G^{-1} F(x, \xi)\}$$

$$\Leftrightarrow \min_{y \in S} \|(x - y) - \alpha^{-1} G^{-1} F(x, \xi))\|_G^2$$

$$\Leftrightarrow \min_{y \in S} \|y - (x - \alpha^{-1} G^{-1} F(x, \xi))\|_G^2.$$

$\square$

The optimal solution of (3.1) is precisely $\mathrm{Proj}_{S,G}(x - \alpha^{-1} G^{-1} F(x, \xi))$, the projection of the point $x - \alpha^{-1} G^{-1} F(x, \xi)$ is onto the set S with respect to the norm $\|\cdot\|_G$. Thus for each x, we uniquely determine the optimal solution of (3.1). Hereafter we shall denote it by $H(x, \xi)$ for notational simplicity. Namely we define the mapping $H : R^n \times \Xi \to R^n$ as $H(x, \xi) = \mathrm{Proj}_{S,G}(x - \alpha^{-1} G^{-1} F(x, \xi))$. So it is not difficult to see that for any $x \in R^n$ and any $\xi \in \Xi$,

$$g(x, \xi) = (x - H(x, \xi))^\mathsf{T} F(x, \xi) - \frac{\alpha}{2}\|x - H(x, \xi)\|_G^2, \tag{3.2}$$

where

$$H(x, \xi) := \text{Proj}_{S,G}(x - \alpha^{-1}G^{-1}F(x, \xi)), \quad \forall x \in S, \ \xi \in \Xi \text{ a.s.},$$

where Ξ is uncertain space, $F : R^n \times \Xi \to R^n$ is a mapping.

Theorem 3.4. *Let S be a closed convex set. Then $z = \text{Proj}_{S,G} f(x, \xi)$ if and only if*

$$z^T(y - z) \geqslant f^T(x, \xi)(y - z), \quad \forall y \in S,$$

or

$$(z - f(x, \xi))^T(y - z) \geqslant 0, \quad \forall y \in S,$$

where $f : R^n \times \Xi \to R^n$.

Proof. Note that $z = \text{Proj}_{S,G} f(x, \xi)$ is the minimizer of $g(y)$ over all $y \in S$. Since $\bigtriangledown g(y) = -2G(f(x, \xi) - y)$, the result follows from the optimality conditions for constrained optimization problems. $\square$

Proposition 3.5. *For each $x \in R^n$, let $H(x, \xi)$ be the unique optimal solution of problem (3.1), or equivalently, let $H : R^n \times \Xi \to R^n$ be the mapping defined by Theorem 3.3. Then x solves the variational inequality problem (1.2) if and only if x^* is a fixed point of the mapping H, i.e., $x^* = H(x^*, \xi)$.*

Proof. Suppose that x^* is a solution of the variational inequality, i.e., $F(x^*, \xi)^T(x - x^*) \geqslant 0$. Multiplying the above inequality by $-\alpha^{-1}G^{-1} < 0$, and adding $x^{*T}(x - x^*)$ to both sides of the resulting inequality, we obtain

$$x^{*T}(x - x^*) \geqslant [x^* - \alpha^{-1}G^{-1}F(x^*, \xi)]^T(x - x^*).$$

From Theorem 3.4 and Lemma 3.2, we conclude that

$$x^* = \text{Proj}_{S,G}(x - \alpha^{-1}G^{-1}F(x, \xi)).$$

Conversely, if $x^* = \text{Proj}_{S,G}(x - \alpha^{-1}G^{-1}F(x, \xi))$, for $\alpha^{-1}G^{-1} > 0$, then

$$x^{*T}(x - x^*) \geqslant [x^* - \alpha^{-1}G^{-1}F(x^*, \xi)]^T(x - x^*),$$

therefore $F(x^*, \xi)^T(x - x^*) \geqslant 0$. $\square$

The next theorem, we establish the equivalence of the variational inequality problem (1.1) and the optimization problem (3.2).

Theorem 3.6. *Let the function $g(x, \xi) : R^n \times \Xi \to R^n$ be defined by (3.1). Then $g(x, \xi) \geqslant 0$, for every $x \in S$, $\xi \in \Xi$, and $g(x, \xi) = 0$ if and only if x solves UVIP(f, S). Hence x solves the optimization problem (3.2) if and only if it solves the variational inequality problem (1.2).*

Proof.

$$
\begin{aligned}
g(x, \xi) &= (x - H(x, \xi))^T F(x, \xi) - \frac{\alpha}{2} \|x - H(x, \xi)\|_G^2 \\
&= \frac{\alpha}{2} \{2\alpha^{-1}(x - H(x, \xi))^T F(x, \xi) - \|x - H(x, \xi)\|_G^2\} \\
&= \frac{\alpha}{2} \{2\alpha^{-1}(x - H(x, \xi))^T F(x, \xi) - \|x - H(x, \xi)\|_G^2\} \\
&= \frac{\alpha}{2} \{2\alpha^{-1}(x - H(x, \xi))^T F(x, \xi) - \|x - H(x, \xi)\|_G^2 \\
&\quad - \alpha^{-2}F^T(x, \xi)G^{-1}F(x, \xi) + \alpha^{-2}F^T(x, \xi)G^{-1}F(x, \xi)\} \\
&= \frac{\alpha}{2} \{-(\|x - H(x, \xi)\|_G^2 - 2\alpha^{-1}(x - H(x, \xi))^T F(x, \xi) + \alpha^{-2}F^T(x, \xi)G^{-1}F(x, \xi)) \\
&\quad + \alpha^{-2}F^T(x, \xi)G^{-1}F(x, \xi)\}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{\alpha}{2}\{-\|(x - H(x,\xi)) - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2 + \alpha^{-2}F^T(x,\xi)G^{-1}F(x,\xi)\} \\
&= \frac{\alpha}{2}\{\alpha^{-2}F^T(x,\xi)G^{-1}F(x,\xi) - \|(x - H(x,\xi)) - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2\} \\
&= \frac{\alpha}{2}\{F^T(x,\xi)\alpha^{-1}G^{-1}G\alpha^{-1}G^{-1}F(x,\xi) - \|H(x,\xi) - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2\} \\
&= \frac{\alpha}{2}\{\|\alpha^{-1}G^{-1}F(x,\xi)\|_G^2 - \|H(x,\xi) - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2\} \\
&= \frac{\alpha}{2}\{\|x - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2 - \|H(x,\xi) - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G^2\}.
\end{aligned}
$$

Since $\|x - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G$ equals the G-distance between x and $x - \alpha^{-1}G^{-1}F(x,\xi)$,

$$
\|H(x,\xi) - (x - \alpha^{-1}G^{-1}F(x,\xi))\|_G,
$$

is the G-distance between $x - \alpha^{-1}G^{-1}F(x,\xi)$ and its projection $H(x,\xi)$ onto S, we have $g(x,\xi) \geqslant 0$, for every $x \in S$, $\xi \in \Xi$.. Moreover, the definition of $H(x,\xi)$ implies that those distances are equal, i.e., $g(x,\xi) = 0$ if and only if $x = H(x,\xi)$. This along with Proposition 3.5 proves the first part of the theorem. □

In what follows, we suppose that $g(x,\cdot)$ is integrable on Ξ for each $x \in S$. Motivated by the work of Chen and Fukushima [1], we propose the following deterministic formulation for UVIP(F,S), called it as the ERM problem, in which we try to find a vector $x \in S$ that minimizes an expected residual,

$$
\min \ \theta(x) := E[g(x,\xi)] = \int_\Xi g(x,\xi)d\Phi(\xi), \tag{3.3}
$$
$$
\text{s.t. } x \in S.
$$

Here, E stands for the expectation with respect to the uncertain variable $\xi \in \Xi$ and $\Phi(\xi)$ stands for the uncertain distribution function. If Ξ has only one realization, then problem (3.3) reduces to (1.2). In the rest of this paper, we assume that F is affine, that is,

$$
F(x,w) := M(\xi)x + q(\xi),
$$

where $M : \Xi \to R^{n \times n}$ and $q : \Xi \to R^n$ are continuous and integrable on Ξ. We further suppose that

$$
\|M(\xi)\| + \|q(\xi)\| < +\infty,
$$

we can easily get

$$
(\|M(\xi)\| + \|q(\xi)\| + 1)^2 < +\infty, \tag{3.4}
$$

where $\|\cdot\|$ means the Euclidean norm. Moreover, we denote

$$
f(x) := \bar{M}x + \bar{q},
$$

with

$$
\bar{M} := E[M(\xi)],
$$
$$
\bar{q} := E[q(\xi)].
$$

Note that

$$
\sqrt{\lambda_{\min}}\|x\| \leqslant \|x\|_G \leqslant \sqrt{\lambda_{\max}}\|x\|, \tag{3.5}
$$

where $\lambda_{\min}$ and $\lambda_{\max}$ indicate the smallest and largest eigenvalues of G, respectively. Therefore, if $\bar{M}$ is positive definite, then f is strongly monotone in the sense that

$$
(x - y)^T(f(x) - f(y)) \geqslant \mu_{\min}\lambda_{\max}^{-1}\|x - y\|_G^2, \tag{3.6}
$$

where $\mu_{\min}$ is the smallest eigenvalue of $\frac{M^T + M}{2}$.

4. Properties of the function θ

We first discuss the differentiability of the function θ

Lemma 4.1. *Suppose that $f(\xi, x)$ is a measurable and integrable function of ξ, for each x in (a, b). Let*

$$\psi(x) = \int f(\xi, x) d\phi(\xi).$$

Suppose that for $\xi \in \Xi$, $f(\xi, x)$ has a derivative in (a, b). For the given $f(\xi, x)$ and $g(\xi)$, we suppose further that

$$|f'(\xi, x)| \leqslant g(\xi)$$

for $\xi \in A$ and $x \in (a, b)$, where g is integrable. Then $\psi(x)$ has derivative $\int f'_x(\xi, x) d\phi(\xi)$ on (a, b).

Proof. To prove this part, consider a fixed x. If $\xi \in \Xi$, then by the mean-value theorem,

$$\frac{f(\xi, x + h) - f(\xi, x)}{h} = f'_x(\xi, s),$$

where s lies between x and $x + h$. The ratio on the left goes to $f'_x(\xi, x)$ as $h \to 0$ and $f'_x(\xi, x)$ is by hypothesis dominated by the integrable function $g(\xi)$. Therefore,

$$\frac{\psi(x + h) - \psi(x)}{h} = \int \frac{f(\xi, x + h) - f(\xi, x)}{h} d\phi(\xi) \to \int f'_x(\xi, x) d\phi(\xi).$$

The condition involving g in this part can be weakened. It suffices to assume that for each x there is an integrable $g(\xi, x)$ such that $|f'_x(\xi, x)| \leqslant g(\xi, x)$ for $\xi \in A$ and all s in some neighborhood of x. $\qquad \square$

Theorem 4.2. *Both g and θ are differentiable with respect to x. In particular, for any $x \in S$, we have*

$$\nabla \theta(x) = E[\nabla_x g(x, \xi)]. \tag{4.1}$$

Proof. Since F is differentiable with respect to x, in a similar way to [5], we can show that g is differentiable with respect to x and

$$\nabla_x g(x, \xi) = F(x, \xi) - (\nabla_x F(x, \xi) - \alpha G)(H(x, \xi) - x). \tag{4.2}$$

On the other hand, note that $g(x, \xi) \geqslant 0$ for any $x \in S$ and $\xi \in \Xi$. We then have from (3.2) and (3.5) that

$$\frac{\alpha}{2} \|x - H(x, \xi)\|_G^2 \leqslant (x - H(x, \xi))^\mathsf{T} F(x, \xi)$$

$$\leqslant \|(x - H(x, \xi))^\mathsf{T} F(x, \xi)\|$$

$$\leqslant \|(x - H(x, \xi))\| \|F(x, \xi)\|$$

$$\leqslant \frac{1}{\sqrt{\lambda_{\min}}} \|x - H(x, \xi)\|_G \|F(x, \xi)\|.$$

It follows that

$$\|x - H(x, \xi)\|_G \leqslant \frac{2}{\alpha \sqrt{\lambda_{\min}}} \|F(x, \xi)\|,$$

and hence

$$\|x - H(x, \xi)\| \leqslant \frac{1}{\sqrt{\lambda_{\min}}} \|x - H(x, \xi)\|_G \leqslant \frac{2}{\alpha \lambda_{\min}} \|F(x, \xi)\|. \tag{4.3}$$

Then, we have from (4.2) and (4.3) that

$$\|\nabla_x g(x, \xi)\| = \|F(x, \xi) + (\nabla_x F(x, \xi) - \alpha G)(x - H(x, \xi))\|$$

$$\leqslant \|F(x, \xi)\| + \|(\nabla_x F(x, \xi) - \alpha G)\| \|x - H(x, \xi)\|$$

$$\leqslant \|F(x,\xi)\| + (\|\nabla_x F(x,\xi)\| + \alpha\|G\|)\|x - H(x,\xi)\|$$

$$\leqslant (1 + \frac{2}{\alpha\lambda_{min}}(\|M(\xi) + \alpha\|G\|))\|F(x,\xi)\|$$

$$= (1 + \frac{2}{\alpha\lambda_{min}}(\|M(\xi) + \alpha\|G\|))\|M(\xi)x + q(\xi)\|$$

$$\leqslant (1 + \frac{2}{\alpha\lambda_{min}}(\|M(\xi) + \alpha\|G\|))(\|x\| + 1)(\|M(\xi)\| + \|q(\xi)\|)^2$$

$$\leqslant \frac{2(\|x\| + 1)}{\alpha\lambda_{min}}(\|M(\xi)\| + \|q(\xi)\|)^2 + (1 + \frac{2\|G\|}{\lambda_{min}})(\|x\| + 1)(\|M(\xi)\| + \|q(\xi)\|).$$

By Lemma 4.1, we get the function θ is differentiable and (4.1) holds. □

We next investigate the conditions for the boundedness of the level set defined by

$$(L_\theta(\bar{c}))^S := \{x \in S | \theta(x) \leqslant c\},$$

where c is a nonnegative number. In the rest of this section, we suppose that $\bar{M}$ is positive definite and $\mu_{min} > 0$ is the smallest eigenvalue of $\frac{\bar{M}^T + \bar{M}}{2}$. As a result, the function f is strongly monotone and (3.6) holds. It has a unique solution [7] from the strong monotonicity of f. We denote by $x^* \in S$ the solution below. Therefore, we have

$$(x - x^*)^T f(x^*) \geqslant 0, \quad \forall x \in S. \tag{4.4}$$

Theorem 4.3. *Let* $\alpha \in (0, 2\mu_{min}\lambda_{max}^{-1})$. *Then, the level set* $(L_\theta(c))^S$ *is bounded for any* $c \geqslant 0$.

Proof. Suppose that there is a nonnegative number $\bar{c}$ such that $(L_\theta(\bar{c}))^S$ is unbounded. This implies that there exists a sequence $x^k \subseteq (L_\theta(\bar{c}))^S$ such that

$$\lim_{k \to \infty} \|x^k\| = +\infty.$$

We have

$$\bar{c} \geqslant \theta(x^k) = E[\max_{y \in S}\{(x^k - y)^T F(x,\xi) - \frac{\alpha}{2}\|x^k - y\|_G^2\}],$$

from Jensen's inequality we can get

$$E[\max_{y \in S}\{(x^k - y)^T F(x,\xi) - \frac{\alpha}{2}\|x^k - y\|_G^2\}] \geqslant \max_{y \in S}\{(x^k - y)^T E[F(x,\xi)] - \frac{\alpha}{2}\|x^k - y\|_G^2\}.$$

So we further get

$$\bar{c} \geqslant \max_{y \in S}\{(x^k - y)^T E[F(x,\xi)] - \frac{\alpha}{2}\|x^k - y\|_G^2\}$$

$$\geqslant (x^k - x^*)^T E[F(x^k,\xi)] - \frac{\alpha}{2}\|x^k - x^*\|_G^2$$

$$\geqslant (x^k - x^*)^T f(x^k) - \frac{\alpha}{2}\|x^k - x^*\|_G^2.$$

From (3.6) we can get

$$(x^k - x^*)^T(f(x^k) - f(x^*)) \geqslant \mu_{min}\lambda_{max}^{-1}\|x - y\|_G^2.$$

So we have

$$(x^k - x^*)^T f(x^k) - \frac{\alpha}{2}\|x^k - x^*\|_G^2 \geqslant (x^k - x^*)^T f(x^*) + (\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})\|x^k - x^*\|_G^2.$$

From (4.4) we can get

$$(x^k - x^*)^T f(x^*) + (\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})\|x^k - x^*\|_G^2 \geqslant (\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})\|x^k - x^*\|_G^2.$$

Finally we get

$$\bar{c} \geqslant (\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})\|x^k - x^*\|_G^2, \xrightarrow{k \to \infty} +\infty.$$

This is a contradiction. Hence $(L_\theta(c))^S$ is bounded for any number $c \geqslant 0$. $\qquad\square$

We further have the following result.

Theorem 4.4. *Let $\alpha \in (0, 2\mu_{min}\lambda_{max}^{-1})$ and let x^* be the unique solution of $VI(f, S)$. We have*

$$\|x - x^*\|_G \leqslant \sqrt{(\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})^{-1}\theta(x)}.$$

Proof. Let $x \in S$. We can obtain

$$\mu_{min}\lambda_{max}^{-1}\|x - x^*\|_G^2 \leqslant (x - x^*)^T(f(x) - f(x^*))$$
$$\leqslant (x - x^*)^T f(x).$$

Therefore, we have

$$(\mu_{min}\lambda_{max}^{-1} - \frac{\alpha}{2})\|x - x^*\|_G^2 \leqslant (x - x^*)^T f(x) - \frac{\alpha}{2}\|x - x^*\|_G^2$$
$$\leqslant \max_{y \in S}\{(x - y)^T f(x) - \frac{\alpha}{2}\|x^k - y\|_G^2\}$$
$$= E[\max_{y \in S}\{(x - y)^T F(x, \xi) - \frac{\alpha}{2}\|x^k - y\|_G^2\}]$$
$$= \theta(x),$$

where the last inequality follows from Jensen's inequality. This completes the proof of the theorem. $\qquad\square$

5. Convergence analysis

Because there dose not exist the density function of uncertain variable, so the distribution function $\Phi(x)$ is not differentiable in uncertain theory. Then we consider the following approximation problem of (3.3) by the differentiable properties of $g(x, \xi)$. Let $T_k = \{t^i | i = 1, 2, \cdots, N_k\}$ be a set satisfying $N_k \to \infty$ as $k \to \infty$.

$$\min \theta^k(x) = g(x, t)\Phi(t)|_{t \in T} - \frac{1}{N_k}\sum_{t^i \in T_k}\Phi(t)g'(x, t^i),$$

$$\text{s.t. } x \in S.$$
(5.1)

5.1. Limiting behavior of global optimal solutions

From now on, we denote by S^* and S_k^* the sets of optimal solutions of problems (3.3) and (5.1), respectively. We first give some lemmas.

Lemma 5.1. *For any fixed $x \in S$, there holds that*

$$\theta(x) = \lim_{k \to \infty}\theta^k(x).$$

Proof. Since

$$\theta(x) = E[g(x,\xi)] = \int_T (x,t)d\Phi(t)$$

$$= g(x,t)\Phi(t)|_{t\in T} - \int_T \Phi(t)dg(x,t)$$

$$= g(x,t)\Phi(t)|_{t\in T} - \int_T \Phi(t)g^{'}(x,t)dt,$$

while

$$\theta^k(x) = g(x,t)\Phi(t)|_{t\in T} - \frac{1}{N_k}\sum_{t^i\in T_k}\Phi(t)g^{'}(x,t^i).$$

Owing to $\lim\limits_{k\to\infty}\frac{1}{N_k}\sum\limits_{t^i\in T_k}\Phi(t)g^{'}(x,t^i) = \int_T \Phi(t)g^{'}(x,t)dt$, so

$$\theta(x) = \lim_{k\to\infty}\theta^k(x).$$

$\square$

Lemma 5.2. *Assume that $\bar{M}$ is positive definite and that $M_k \in R^{n\times n}$ tends to $\bar{M}$ as $k\to\infty$. Then, there exists a number k_0 such that M_k is positive definite for every $k \geqslant k_0$.*

Proof. Suppose that $\{M_k\}$ has a subsequence $\{M_{k_j}\}$, each member of which is not positive definite. This implies that, for each j,

$$(x^j)^\mathsf{T} M_{k_j} x^j \leqslant 0, \tag{5.2}$$

holds for some $x^j \in R^n$ with $\|x^j\| = 1$. Since $\{x^j\}$ is bounded, without loss of generality, we may assume that $\lim\limits_{j\to\infty} x^j = \bar{x}$. Noting that

$$(x^j)^\mathsf{T} M_{k_j} x^j - (\bar{x})^\mathsf{T}\bar{M}\bar{x} = (x^j)^\mathsf{T}(M_{k_j}-\bar{M})x^j$$

$$+ (x^j)^\mathsf{T}\bar{M}(x^j-\bar{x}) + (x^j-\bar{x})^\mathsf{T}\bar{M}\bar{x} \xrightarrow{j\to\infty} 0,$$

we have that

$$\lim_{j\to\infty}(x^j)^\mathsf{T} M_{k_j} x^j = (\bar{x})^\mathsf{T}\bar{M}\bar{x}.$$

It follows from (5.2) that

$$(\bar{x})^\mathsf{T}\bar{M}\bar{x} \leqslant 0, \|\bar{x}\| = 1,$$

which contradicts the fact that $\bar{M}$ is a positive definite matrix. This completes the proof. $\square$

Lemma 5.3. *Suppose that $\bar{M}$ is positive definite and $\lim\limits_{k\to\infty} M_k = \bar{M}$. Let μ_k be the biggest eigenvalue of $(M_k^\mathsf{T} + M_k)/2$ for each k. Then, we have that $\lim\limits_{k\to\infty}\mu_k = \mu_{max}$.*

Proof. Lemma 5.3 can be verified easily and its proof is omitted here. $\square$

We next prove the main results of this subsection.

Theorem 5.4. *Assume that $\bar{M}$ is positive definite. We denote $\bar{k} = M^{'}(t)\Phi(t)|_{t\in T}$, $m = (\Phi(t)|_{t\in T} - 1)f(x^*)$ and $\hat{k} = \int_T \Phi(t)dt$. Let λ_{max} be the same as in (3.5) and Let μ_{max} be the same as Lemma 5.3. Let $\bar{k} > \mu_{max}$ and $\alpha \in (0, \min\{\frac{\bar{k}-\mu_{max}-m}{\hat{k}\lambda_{max}}, \frac{\bar{k}-\mu_{max}+m}{\hat{k}\lambda_{max}}\})$. Then, there exists an integer such that for each $k \geqslant k_0$, the level set $L^S_{\theta^k}(c)$ is bounded for any $c \geqslant 0$. In particular, the solution set S_k^* is nonempty and bounded for every sufficiently large k.*

Proof. We denote the solution of UVIP(f, S) by x^*, we then have (4.4). It then follows that

$$\bar{c} \geqslant \theta^k(x^j) = g(x^j, t)\Phi(t)|_{t \in T} - \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) g_t'(x^j, t^i).$$

Owing to $g(x^j, t)\Phi(t)|_{t \in T}$ is bounded, we next prove $\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) g'(x^j, t^i)$ is bounded. Let $k \geqslant k_0$. Suppose that there is a $\hat{c} \geqslant 0$ such that $\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t) g_t'(x^j, t^i)$ is unbounded. This implies that there exists a sequence $\{x^j\} \subseteq L^S_{\theta^k}(c)$ such that $\lim_{j \to \infty} \|x^j\| = +\infty$. Without loss of generality, we may assume that $\|x^j - x^*\| \geqslant 1$ for each j. Let

$$M_k := M(x^j, t)\Phi(t)|_{t \in T} - \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) M_t'(x^j, t^i).$$

It is easy to get that

$$\lim_{k \to \infty} M_k = \bar{M}.$$

It is also easy to get that

$$\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) M_t'(x^j, t^i) = M(x^j, t)\Phi(t)|_{t \in T} - M_k.$$

It then follows that

$$\hat{c} \geqslant \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) g_t'(x^j, t^i)$$

$$= \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) \max_{y \in S} \{[M_t'(t^i)x^j + q_t'(t^i)]^T (x^j - y) - \frac{\alpha}{2} \|x^j - y\|_G^2\}$$

$$\geqslant \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) \{[M_t'(t^i)x^j + q_t'(t^i)]^T (x^j - x^*) - \frac{\alpha}{2} \|x^j - x^*\|_G^2\}$$

$$= \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) \{[M_t'(t^i)(x^j - x^* + x^*) + q_t'(t^i)]^T (x^j - x^*) - \frac{\alpha}{2} \|x^j - x^*\|_G^2\}$$

$$= (x^j - x^*)^T (M_t'(t)\Phi(t)|_{t \in T} - M_k)(x^j - x^*) + [\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i))]^T (x^j - x^*)$$

$$- \frac{\alpha}{2} \|x^j - x^*\|_G^2 \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)$$

$$\geqslant (x^j - x^*)^T (M_t'(t)\Phi(t)|_{t \in T} - M_k)(x^j - x^*) - \|\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i))\| \|(x^j - x^*)\|$$

$$- \frac{\alpha}{2} \|x^j - x^*\|_G^2 \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i),$$

where we can take $\epsilon \in (0, \min\{\frac{2\bar{k} - 2\mu_{max} - \alpha\hat{k}\lambda_{max} + m}{4 + \alpha\lambda_{max}}, \frac{2\bar{k} - 2\mu_{max} - \alpha\hat{k}\lambda_{max} - m}{4 + \alpha\lambda_{max}}\})$ and

$$f(x^*) = \bar{M}x^* + \bar{q}$$

$$= (M(t)\Phi(t)|_{t \in T} - (\int_T \Phi(t)M_t'(t)dt)x^* + q(t)\Phi(t)|_{t \in T} - \int_T \Phi(t)q_t'(t)dt$$

$$= (M(t)x^* + q(t))\Phi(t)|_{t \in T} - (\int_T \Phi(t)M_t'(t)dt\, x^* + q_t'(t)dt)$$

$$= f(x^*)\Phi(t)|_{t \in T} - \lim_{k \to \infty} \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)),$$

so we have $\lim_{k \to \infty} \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) = (\Phi(t)|_{t \in T} - 1)f(x^*)$.

There exists $\epsilon \geqslant 0$ such that

$$|\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) - (\Phi(t)|_{t \in T} - 1)f(x^*)| < \epsilon.$$

Then

$$-\epsilon < \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) - (\Phi(t)|_{t \in T} - 1)f(x^*) < \epsilon.$$

It implies

$$(\Phi(t)|_{t \in T} - 1)f(x^*) - \epsilon < \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) < (\Phi(t)|_{t \in T} - 1)f(x^*) + \epsilon.$$

We use the same ϵ in next two inequalities, then

$$\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(\xi) < \hat{k} + \epsilon,$$

$$\mu_k < \mu_{max} + \epsilon.$$

If $\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) > 0$, then

$$-\|\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i))\| \geqslant -\|(\Phi(t)|_{t \in T} - 1)f(x^*) + \epsilon\| = -(m + \epsilon),$$

so

$$\hat{c} \geqslant (\bar{k} - \mu_k)\|x^j - x^*\|^2 - (m + \epsilon)\|x^j - x^*\| - \frac{\alpha(\hat{k} + \epsilon)}{2}\|x^j - x^*\|_G^2$$

$$\geqslant [(\bar{k} - \mu_{max} - \epsilon - \frac{\alpha(\hat{k} + \epsilon)\lambda_{max}}{2})\|x^j - x^*\| - m - \epsilon]\|x^j - x^*\|$$

$$\geqslant \frac{1}{2}[2\bar{k} - 2\mu_{max} - \alpha\hat{k}\lambda_{max} - m - (4 + \alpha\lambda_{max})\epsilon]\|x^j - x^*\| \xrightarrow{j \to \infty} +\infty,$$

while

$$\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i)) < 0,$$

then

$$-\|\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i)(M_t'(t^i)x^* + q_t'(t^i))\| \geqslant -\|(\Phi(t)|_{t \in T} - 1)f(x^*) - \epsilon\| = m - \epsilon,$$

so

$$\hat{c} \geqslant (\bar{k} - \mu_k)\|x^j - x^*\|^2 + (m - \epsilon)\|x^j - x^*\| - \frac{\alpha(\hat{k} + \epsilon)}{2}\|x^j - x^*\|_G^2$$

$$\geqslant [(\bar{k} - \mu_{max} - \epsilon - \frac{\alpha(\hat{k} + \epsilon)\lambda_{max}}{2})\|x^j - x^*\| + m - \epsilon]\|x^j - x^*\|$$

$$\geqslant \frac{1}{2}[2\bar{k} - 2\mu_{max} - \alpha\hat{k}\lambda_{max} + m - (4 + \alpha\lambda_{max})\epsilon]\|x^j - x^*\| \xrightarrow{j \to \infty} +\infty.$$

This is a contradiction, so $\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t) g_t'(x^j, t^i)$ is bounded and hence $L_{\theta^k}^S(c)$ is bounded for any $c \geqslant 0$. $\qquad \square$

Theorem 5.5. *Assume that $\bar{M}$ is positive definite, $\alpha \in (0, 2\mu_{\min}\lambda_{\max}^{-1})$. Let $x^k \in S_k^*$ for each sufficiently large k. Then, every accumulation point of the sequence $\{x^k\}$ is contained in S^*.*

Proof. Let x^* be an accumulation point of $\{x^k\}$. Without loss of generality, let $\{x^k\}$ converge to x^*. It is obvious that $x^* \in S$, we first show that

$$\lim_{k \to \infty} (\theta^k(x^k) - \theta^k(x^*)) = 0. \tag{5.3}$$

In fact, since $x^k \to x^*$ as $k \to +\infty$, there exists a constant $C > 0$ such that $\|x^*\| \leqslant C$ and $\|x^k\| \leqslant C$. It follows from the mean-value theorem and Theorem 4.2 that, for any $t \in T$ and any k,

$$|g(x^k, t) - g(x^*, t)| = |\nabla_x g(y^k, t)(x^k - x^*)|$$
$$\leqslant C'((\|M(t)\| + \|q(t)\| + 1)^2)\|x^k - x^*\|,$$

so $|g_t'(x^k, t) - g_t'(x^*, t)| \leqslant C' \nabla_t ((\|M(t)\| + \|q(t)\| + 1)^2)\|x^k - x^*\|$, where y^k is between $\{x^k\}$ and x^*,

$$C' = (C + 1) \max\{\frac{2}{\alpha\lambda_{\min}}, 1 + \frac{2\|G\|}{\lambda_{\min}}\}.$$

Then, we have

$$|\theta^k(x^k) - \theta^k(x^*)| = |g(x^k, t)\Phi(t)|_{t \in T} - \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t) g_t'(x^k, t^i)$$

$$- g(x^*, t)\Phi(t)|_{t \in T} - \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) g_t'(x^*, t^i)|$$

$$\leqslant |g(x^k, t)\Phi(t)|_{t \in T} - g(x^*, t)\Phi(t^i)|_{t \in T}|$$

$$+ |\frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t)(g_t'(x^k, t^i) - g_t'(x^*, t^i))|$$

$$\leqslant C'\|x^k - x^*\| \frac{1}{N_k} \sum_{t^i \in T_k} \Phi(t^i) \nabla_t ((\|M(t)\| + \|q(t)\| + 1)^2)$$

$$\leqslant C'\|x^k - x^*\| \frac{1}{N_k} \sum_{t^i \in T_k} \nabla_t ((\|M(t)\| + \|q(t)\| + 1)^2) \xrightarrow{k \to \infty} 0,$$

where the limit follows from the facts that the sequence $\{x^k\}$ converges to x^*. By (3.4), we can get

$$\lim_{k \to \infty} \frac{1}{N_k} \sum_{t^i \in \Xi_k} \nabla_t ((\|M(t)\| + \|q(t)\| + 1)^2) = (\|M(t)\| + \|q(t)\| + 1)^2 < +\infty.$$

Therefore, (5.3) is valid.

On the other hand, noting that

$$|\theta^k(x^k) - \theta(x^*)| \leqslant |\theta^k(x^k) - \theta^k(x^*)| + |\theta^k(x^*) - \theta(x^*)|.$$

From Lemma 5.1 and (5.3), we have

$$\lim_{k \to \infty} \theta^k(x^k) = \theta^k(x^*). \tag{5.4}$$

For each sufficiently large k, $x^k \in S_k^*$, then

$$\theta^k(x^k) \leqslant \theta^k(x), \tag{5.5}$$

holds for any $x \in S$. Letting $k \to +\infty$ in (5.5) and taking (5.4) and Lemma 5.1 into account, we get $\theta^k(x^k) \leqslant \theta(x)$, which means $x^* \in S^*$. $\qquad \square$

5.2. Limiting behavior of stationary points

Suppose that $S = \{x \in \mathbb{R}^n | c(x) \leqslant 0\}$, where $c_i : \mathbb{R}^n \to \mathbb{R}$ are all continuously differentiable convex functions for $i = 1, 2, \cdots, m$. We next consider the limiting behavior of the stationary points of problems (5.2). We will use the standard definition of stationarity, i.e., a point x^k is said to be stationary to (5.2) if there exists a Lagrange multiplier vector $\mu^k \in \mathbb{R}^m$ such that

$$\nabla\theta^k(x^k) + \sum_{i=1}^{m} \mu_i^k \nabla c_i(x^k) = 0, \tag{5.6}$$

$$0 \leqslant \mu^k, c(x^k) \leqslant 0, \quad \text{and} \quad (\mu^k)^T c(x^k) = 0. \tag{5.7}$$

For problem (3.3), x^* is said to be a stationary point if there exists a Lagrange multiplier vector $\mu^* \in \mathbb{R}^m$ such that

$$\nabla\theta(x^*) + \sum_{i=1}^{m} \mu_i^* \nabla c_i(x^*) = 0, \tag{5.8}$$

$$0 \leqslant \mu^*, c(x^*) \leqslant 0, \quad \text{and} \quad (\mu^*)^T c(x^*) = 0. \tag{5.9}$$

In addition, we say that the Slater's constraint qualification holds, if there exists a vector $y \in \mathbb{R}^n$ such that $c_i(y) < 0$ for each $i = 1, 2, \cdots, m$.

Theorem 5.6. *Suppose that* $\lim_{k\to\infty} x^k = x^*$. *Then, we have*

$$\lim_{k\to\infty} \nabla\theta^k(x^k) = \nabla\theta(x^*).$$

Proof. From the nonexpansive property of $\text{Proj}_{S,G}$ and (3.5), we obtain

$$\begin{aligned}
\|H(x^*, t^i) - H(x^k, t^i)\|_G &\leqslant \|(x^* - \alpha^{-1}G^{-1}F(x^*, t^i)) - (x^k - \alpha^{-1}G^{-1}F(x^k, t^i))\|_G \\
&\leqslant \|x^* - x^k\|_G + \alpha^{-1}\|G^{-1}M(t^i)(x^k - x^*)\|_G \\
&\leqslant \lambda_{max}^{\frac{1}{2}}(1 + \alpha^{-1}\|G^{-1}\|\|M(t^i)\|)\|(x^k - x^*)\|.
\end{aligned}$$

Using (3.5) again, we have

$$\begin{aligned}
\|H(x^*, t^i) - H(x^k, t^i)\| &\leqslant \lambda_{min}^{-\frac{1}{2}}\|H(x^*, t^i) - H(x^k, t^i)\|_G \\
&\leqslant \lambda_{min}^{-\frac{1}{2}}\lambda_{max}^{\frac{1}{2}}(1 + \alpha^{-1}\|G^{-1}\|\|M(t^i)\|)\|(x^k - x^*)\| \xrightarrow{k\to\infty} 0.
\end{aligned}$$

Here anything that has nothing to do with x can be considered a constant.

From the nonexpansive property of $\text{Proj}_{S,G}$ and (3.5) again, it implies that

$$\begin{aligned}
\|H'(x^*, t^i) - H'(x^k, t^i)\|_G &\leqslant \|(x^* - \alpha^{-1}G^{-1}F'(x^*, t^i)) - (x^k - \alpha^{-1}G^{-1}F'(x^k, t^i))\|_G \\
&\leqslant \|x^* - x^k\|_G + \alpha^{-1}\|G^{-1}M'(t^i)(x^k - x^*)\|_G \\
&\leqslant \lambda_{max}^{\frac{1}{2}}(1 + \alpha^{-1}\|G^{-1}\|\|M'(t^i)\|)\|(x^k - x^*)\|.
\end{aligned}$$

Though (3.5) again, we obtain

$$\begin{aligned}
\|H'(x^*, t^i) - H'(x^k, t^i)\| &\leqslant \lambda_{min}^{-\frac{1}{2}}\|H'(x^*, t^i) - H'(x^k, t^i)\|_G \\
&\leqslant \lambda_{min}^{-\frac{1}{2}}\lambda_{max}^{\frac{1}{2}}(1 + \alpha^{-1}\|G^{-1}\|\|M'(t^i)\|)\|(x^k - x^*)\| \xrightarrow{k\to\infty} 0.
\end{aligned}$$

Here only x is the variable, others can be considered a constant. It then follows from (5.1)

$$\lim_{k\to\infty} \nabla\theta^k(x^k) = \lim_{k\to\infty} \nabla_x g(x^k,t)\Phi(t)|_{t\in T} - \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)g''_{xt}(x^k,t^i)$$

$$= \nabla_x g(x^*,t)\Phi(t)|_{t\in T} - \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)g''_{xt}(x^k,t^i),$$

while

$$\lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)g''_{xt}(x^k,t^i)$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)[F'_t(x^k,t^i) - F''_{xt}(x^k,t^i)(H(x^k,t^i) - x^k) - (F'_x(x^k,t^i) - \alpha G)H'_t(x^k,t^i)]$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)[M'_t(t^i)x^k + q'_t(t^i) + M'_t(t^i)(x^k - H(x^k,t^i)) - (M(t^i) - \alpha G)H'_t(x^k,t^i)]$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i) \cdot \lim_{k\to\infty} x^k + \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)q'_t(t^i)$$

$$+ \lim_{k\to\infty} [\frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i)]^T \cdot \lim_{k\to\infty} x^k$$

$$- \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i)H(x^*,t^i)$$

$$+ \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i)(H(x^*,t^i) - H(x^k,t^i))$$

$$- \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)(M(t^i) - \alpha G)H'_t(x^*,t^i)$$

$$+ \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)(M(t^i) - \alpha G)(H'_t(x^*,t^i) - H'_t(x^k,t^i))$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i) \cdot \lim_{k\to\infty} x^* + \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)q'_t(t^i)$$

$$+ \lim_{k\to\infty} [\frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i)]^T \cdot \lim_{k\to\infty} x^*$$

$$- \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)M'_t(t^i)H(x^*,t^i) - \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)(M(t^i) - \alpha G)H'_t(x^*,t^i)$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)[M'_t(t^i)x^* + q'_t(t^i) + M'_t(t^i)(x^* - H(x^*,t^i)) - (M(t^i) - \alpha G)H'_t(x^*,t^i)]$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)[F'_t(x^*,t^i) - F''_{xt}(x^*,t^i)(H(x^*,t^i) - x^*) - (F'_x(x^*,t^i) - \alpha G)H'_t(x^*,t^i)]$$

$$= \lim_{k\to\infty} \frac{1}{N_k}\sum_{t^i\in T_k} \Phi(t^i)g''_{xt}(x^*,t^i)$$

$$= \int_T \Phi(t)g''_{xt}(x^*,t)dt.$$

We know that

$$\nabla\theta(x^*) = E[\nabla_x g(x^*, t)] = \int_T \nabla_x g(x^*, t) d\Phi(t)$$

$$= \nabla_x g(x^*, t)\Phi(t)|_{t\in T} - \int_T \Phi(t) d\nabla_x g(x^*, t)$$

$$= \nabla_x g(x^*, t)\Phi(t)|_{t\in T} - \int_T \Phi(t) g''_{xt}(x^*, t) dt.$$

So, we have

$$\lim_{k\to\infty} \nabla\theta^k(x^k) = \nabla\theta^k(x^*).$$

$\square$

Theorem 5.7. *Let x^k be stationary to (5.2) for each k and let $\lim_{k\to\infty} x^k = x^*$. Suppose that the Slater's constraint qualification holds. Then, $\{\mu^k\}$ is bounded.*

Proof. We denote

$$\upsilon_k := \sum_{i=1}^m \mu_i^k. \tag{5.10}$$

Suppose that $\{\mu^k\}$ is unbounded, which means $\lim_{k\to\infty} \upsilon_k = +\infty$. Taking a subsequence, we may assume that the limits $\bar{\mu}_i := \lim_{k\to\infty} \frac{\mu_i^k}{\upsilon_k}$ $(i = 1, 2, \cdots, m)$ exist. It is easy to see from (5.7) that $\bar{\mu}_i = 0$ for every $i \notin \Upsilon(x^*) := \{i \mid c_i(x^*) = 0, 1 \leqslant i \leqslant m\}$. Then we have from (5.10) that

$$\sum_{i\in\Upsilon(x^*)} \bar{\mu}_i = \sum_{i=1}^m \bar{\mu}_i = 1. \tag{5.11}$$

Note that ∇c_i is continuous for each i, by Theorem 5.6, $\{\nabla\theta^k(x^k)\}$ is convergent. Then, dividing (5.6) by υ_k and taking a limit, we obtain

$$\sum_{i\in\Upsilon(x^*)} \bar{\mu}_i \nabla c_i(x^*) = \sum_{i=1}^m \bar{\mu}_i \nabla c_i(x^*) = 0. \tag{5.12}$$

Recall that the Slater's constraint qualification holds. So, there exists a vector $y \in \mathbb{R}^n$ such that $c_i(y) < 0$ for each $i = 1, 2, \cdots, m$. From c_i is convex, we have

$$(y - x^*)^\mathsf{T} \nabla c_i(x^*) \leqslant c_i(y) - c_i(x^*) = c_i(y) < 0, \quad \forall i \in \Upsilon(x^*). \tag{5.13}$$

Since $\bar{\mu}_i \geqslant 0$ for each i, it follows from (5.12) and (5.13) that $\bar{\mu}_i = 0$ for each $i \in \Upsilon(x^*)$. This contradicts (5.11) and hence $\{\mu^k\}$ is bounded. $\square$

Theorem 5.8. *Let x^k be stationary to (5.2) for each k and let x^* be an accumulation point of $\{x^k\}$. If the Slater's constraint qualification holds, then x^* is a stationary point of problem (3.3).*

Proof. For simplicity, let x^k converge to x^*. By Theorem 5.7, the sequence $\{\mu^k\}$ is bounded. There exists a subsequence, still denoted by $\{\mu^k\}$ such that the limit $\mu^* := \lim_{k\to\infty} \mu^k$. Note that both c_i and ∇c_i are continuous for each i, by Theorem 5.6, it holds $\lim_{k\to\infty} \nabla\theta^k(x^k) = \nabla\theta(x^*)$. Taking a limit in (5.6) and (5.7), we obtain (5.8) and (5.9) immediately. That is, x^* is stationary to problem (3.3). $\square$

6. Conclusions

We have proposed an ERM model for the uncertain variational inequality problem which provide a new optimization problem formulation of the variational inequality problem. Properties of the model have been investigated. We have also completed convergence analysis about the function θ based on uncertain theory. In addition, there have been presented a number of gap functions for the classical variational inequality problems, see [4, 6] for details. It may be possible to extend the proposed approach instead of the regularized gap function by other gap functions.

Acknowledgment

References

[1] X.-J. Chen, M. Fukushima, *Expected residual minimization method for stochastic linear complementarity problems*, Math. Oper. Res., **30** (2005), 1022–1038. 1, 3

[2] X.-W. Chen, D. A. Ralescu, *Liu process and uncertain calculas*, J. Uncert. Anal. Appl., **2013** (2013), 6 pages. 2

[3] Q.-Q. Chen, Y.-G. Zhu, *A class of uncertain variational inequality problems*, J. Inequal Appl., **2015** (2015), 13 pages. 1

[4] F. Facchinei, J.-S. Pang, *Finite-Dimensional Variational Inequalities and Complementarity Problems*, Springer, New York, (2003). 1, 6

[5] M. Fukushima, *Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems*, Math. Programming, **53** (1992), 99–110. 1, 1, 3, 4

[6] M. Fukushima, *Merit functions for variational inequality and complementarity problems*, Nonlinear Optim. Appl., **68** (1996), 155–170. 1, 1, 3, 6

[7] P. T. Harker, J.-S. Pang, *Finite-dimensional variational inequality and nonlinear complementarity problems: A survey of theory, algorithms and applications*, Math. Programming, **48** (1990), 161–220. 4

[8] B. Liu, *Uncertainty Theory*, Springer-Verlag, Berlin, (2007). 1, 2.1, 2.2, 2.4, 2.5, 2, 2.9, 2.10

[9] B. Liu, *Uncertainty Theory*, Springer-Verlag, Berlin, (2015). 2.6, 2.7, 2.8

[10] Y.-H. Yao, M. Postolache, Y.-C. Liou, Z.-S. Yao, *Construction algorithms for a class of monotone variational inequalities*, Optim. Lett., **10** (2016), 1519–1528. 1

[11] Y.-H. Yao, X. Qin, J.-C. Yao, *An improved algorithm based on Korpelevich's method for variational inequalities in Banach spaces*, J. Nonlinear Convex Anal., (in press).

[12] Y.-H. Yao, N. Shahzad, *Strong convergence of a proximal point algorithm with general errors*, Optim. Lett., **6** (2012), 621–628.

[13] Y.-H. Yao, N. Shahzad, *An algorithmic approach to the split variational inequality and fixed point problem*, J. Nonlinear Convex Anal., **18** (2017), 977–991.

[14] H. Zegeye, N. Shahzad, Y.-H. Yao, *Minimum-norm solution of variational inequality and fixed point problem in Banach spaces*, Optimization, **64** (2015), 453–471. 1