Generalized-Hukuhara-Gradient Efficient-Direction Method to Solve Optimization Problems with Interval-Valued Functions and Its Application in Least-Squares Problems

Debdas Ghosh · Amit Kumar Debnath · Ram Surat Chauhan · Oscar Castillo

Received: 10 February 2021 / Revised: 28 April 2021 / Accepted: 10 August 2021 / Published online: 21 November 2021
© Taiwan Fuzzy Systems Association 2021

Abstract This article proposes a general $gH$-gradient efficient-direction method and a $WgH$-gradient efficient method for the optimization problems with interval-valued functions. The convergence analysis and the step-wise algorithms of both methods are presented. It is observed that the $WgH$-gradient efficient method converges linearly for a strongly convex interval-valued objective function. To develop the proposed methods and to study their convergence, the ideas of strong convexity and sequential criteria for $gH$-continuity of interval-valued function are illustrated. In the sequel, a new definition of $gH$-differentiability for interval-valued functions is also proposed. The new definition of $gH$-differentiability is described with the help of a newly defined concept of linear interval-valued function. It is noticed that the proposed $gH$-differentiability is superior to the existing ones. For a $gH$-differentiable interval-valued function, the relation of convexity with the $gH$-gradient of an interval-valued function and an optimality condition of an interval optimization problem is derived. For the derived optimality condition, a notion of efficient direction for interval-valued functions is introduced. The idea of efficient direction is used to develop the proposed gradient methods. As an application of the proposed methods, the least-square problem for interval-valued data by $WgH$-gradient efficient method is solved. The proposed method for least square problems is illustrated by a polynomial fitting and a logistic curve fitting.

Keywords Interval-valued functions · Linearity · $gH$-Differentiability · Interval-optimization problems · Efficient solution · Efficient direction · Least-square problems

Mathematics Subject Classification 90C30 · 65K05

1 Introduction

Each area of science, engineering, management, economics, and other practices uses optimization techniques extensively. Optimization techniques assist us in finding the best under specified circumstances. The optimization problems with interval-valued functions (IVFs), known as interval optimization problems (IOPs), have become a significant research topic over the last two decades due to inherent imprecise and uncertain events in different real-world events. In this paper, we attempt to derive a technique for IOPs to capture their solution set. The proposed method reduces to the steepest descent method for the optimization problems with real-valued functions.

1.1 Literature Survey and Motivation

The ordering and subtraction of intervals have always been an issue in pursuit of an optimal solution for IOPs [1].
order to deal with interval-valued data, Moore introduced interval arithmetic [2, 3]. However, with the interval arithmetic in [2, 3], one cannot find the additive inverse of a nondegenerate interval (whose lower and upper limits are different), i.e., for a nondegenerate interval \( A \), there does not exist an interval \( B \) such that \( A \oplus B = 0 \). Due to this reason, Wu [4–6] used a new concept of difference of intervals, known as Hukuhara difference [7] for the difference of two nonempty, closed, bounded, and convex subsets of a real linear space. In spite of the fact that the Hukuhara difference for intervals satisfies \( A \ominus_H A = 0 \), \( A \ominus_H B \) can be calculated only when the width of \( A \) is greater than equal to that of \( B \). To overcome this inefficiency of Hukuhara difference of intervals, the ‘nonstandard subtraction,’ introduced by Markov [8], has been used and named as generalized-Hukuhara difference (\( gH \)-difference) by Stefanini [9, 10]. The generalized-Hukuhara difference can be calculated for any pair of intervals and has the property that \( A \ominus_{gH} A = 0 \) [9].

In the ordering of intervals, as intervals are not linearly ordered in contrast to the real numbers, Ishibuchi and Tanaka [11] showed various partial ordering structures and solution concepts for IOPs. They suggested a method to solve a linear IOP by converting it to a bi-objective optimization problem, which is generalized by Chanas and Kuchuta [12]. For nonlinear IOPs, Ghosh studied a Newton method [13] and a quasi-Newton method [14]. Interestingly, many researchers proposed different types of algorithms to solve various types of practical IOPs, for instance, see [15–21]. Recently, Ghosh et al. [22] introduced variable ordering relations of intervals and proposed an algorithm to obtain the solutions to IOPs. However, research into the applicability of conventional optimization techniques for IOPs is still not concentrated. More surprisingly, although the interplay between geometry and calculus yields optimization techniques, the calculus for IVFs is not rigorously developed until now.

In the year 2007, with the help of a Hausdorff metric between any two intervals, Wu [4] illustrated the concept of continuity of an IVF. In the same article, [4], based on the Hukuhara difference, the concept of Hukuhara differentiability (\( H \)-differentiability) of an IVF has been proposed. Accordingly, the KKT optimality conditions for IOPs have been given in [4]. Further, applying the concept of \( H \)-differentiability, Wu [4–6] studied various duality theories of IOPs. Thereafter, showing the restrictiveness of \( H \)-differentiability, Chalco-Cano et al. [23] developed the calculus of IVFs based on the modified concept of the \( gH \)-difference, known as generalized-Hukuhara differentiability (\( gH \)-differentiability). Chalco-Cano et al. [24], and Ghosh et al. [25] also derived the KKT conditions and duality theories in the view of \( gH \)-differentiability.

In the development of interval calculus, the calculus for the fuzzy-valued functions plays an important role because intervals are particular fuzzy numbers with unit membership value of all the numbers in the interval. In fact, fuzzy systems and interval systems are often inseparably related to each other since alpha-cuts of fuzzy numbers are intervals, and any interval can be represented by a fuzzy number with membership value unity across all the numbers in the interval. In connection with fuzzy calculus, Bede and Gal [26] introduced \( gH \)-differentiability for fuzzy-valued functions, which motivated the search for \( gH \)-difference for intervals (see [9, 10, 27]) and \( gH \)-differentiability for interval-valued functions (see [28]). A recent contribution in this direction is the article by Stefanini and Arana-Jiménez [29] which contains definitions of the total, directional, and partial \( gH \)-derivatives for both interval- and fuzzy-valued functions.

In the existing literature on interval calculus, unlike the definition of differentiability of real-valued functions, none of the existing approaches used the concept of a linear IVF to define the differentiability of an IVF. Despite similar to the definition of differentiability of real-valued functions, the authors of [13] and [29] introduced the new definitions of \( gH \)-differentiability for IVFs and studied the properties \( gH \)-differentiable IVFs. However, none of them also mentioned linear IVF and used the concept of linear IVF to define \( gH \)-differentiability for IVFs.

Since the last two decades, with the development of the calculus of IVFs and theories related to IOPs, many techniques, and their algorithmic implementations to obtain the efficient solutions of various types of practical IOPs have appeared, for instance, see [14, 22, 30–34]. However, the majority of the methods are provided from the perspective of conventional bi-objective optimization. Thus, to apply those techniques, one has to explicitly express an IVF \( F \) in terms of its real-valued lower \( (f) \) and upper \( (\overline{f}) \) functions, which is quite restrictive. For example, in a general least-square problem for interval-valued data (see Sect. 6), one cannot express the interval-valued error function in terms of its lower and upper boundary functions. The authors of [1, 13, 35] developed theories and techniques to find efficient solutions to the IOPs in which objective functions can be parametrically presented. However, for the parametric representation of an IVF, one needs its explicit form, which is often practically not possible, for instance, consider the function \( E(\beta) \) in (20).

The literature on IOPs shows that there is still no emphasis on the study of conventional optimization algorithms for IOPs. Surprisingly, the basic descent method is not yet developed for IOPs. Further, to derive a technique for IOPs that is similar to the standard descent method, we need to establish the notion of the \( gH \)-differentiability.
concept for IVFs rigorously. More importantly, it must be kept in mind that the derived technique must be applicable to general IVFs regardless of whether or not

(i) the objective function can be expressed parametrically, or
(ii) the explicit forms of the lower and upper functions of the objective function can be found.

1.2 Contribution

The main contributions of this article are the following.

- After illustrating the concept of linear IVF, this paper proposes a new definition of \( gH \)-differentiability. It is shown that if an IVF is \( gH \)-differentiable, its \( gH \)-gradient exists. It is found that the proposed definition of \( gH \)-differentiability is superior to the existing ones (see Remark 8 for details).
- With the help of \( gH \)-gradient, a few characterization results for \( gH \)-differentiable convex IVFs are derived. Also, several results related to the \( gH \)-gradient of a strongly convex \( gH \)-differentiable IVF are studied.
- By the proposed \( gH \)-differentiability for IVFs, this article develops a gradient descent method for interval optimization, namely a general \( gH \)-gradient efficient-direction method for IOPs.
- Like the steepest descent method, a method is also proposed, named \( W \)-\( gH \)-gradient method, to obtain efficient solutions of IOPs. The main advantages of the proposed methods are that one needs neither the explicit forms of upper and lower functions of the objective function nor the parametric form of the IVFs of an IOP.
- It is shown that the \( W \)-\( gH \)-gradient method for IOP converges linearly in the case of strong convexity of the interval-valued objective function. To develop these methods, the notion of efficient direction for IVF and its several characteristics are studied.

1.3 Delineation

The presentation sequence of the proposed work is the following. The next section covers some basic terminologies and notions of intervals followed by the convexity, and a few topics of differential calculus of IVFs.

2 Preliminaries and Terminologies

This section provides some basic terminologies and notions on intervals followed by the convexity, and a few topics of differential calculus of IVFs.

2.1 Arithmetic of Intervals and Their Dominance Relation

At first, this section describes the generalized concept of the difference of two intervals and the ordering concepts of intervals. Along with these definitions, we use Moore’s interval addition (\( \oplus \) ) multiplication (\( \odot \) ) and division (\( \oslash \) ) throughout the paper.

Let the set of real numbers be denoted by \( \mathbb{R} \) and the set of all closed and bounded intervals be denoted by \( I(\mathbb{R}) \). Throughout the article, the elements of \( I(\mathbb{R}) \) are represented by bold capital letters \( A, B, C, \ldots \). To represent an element \( A \in I(\mathbb{R}) \) in the interval form, the corresponding small letter is used in the following way: \( A = [a, \bar{a}] \). If \( \mu = [\bar{a}, a] \), then \( A \) is called a degenerate interval.

It is to be mentioned that any singleton \( \{p\} \) of \( \mathbb{R} \) can be represented by an interval \( P = [p, p] \), where \( p = p = \bar{p} \). In particular,

\[ 0 = \{0\} = [0, 0] \text{ and } 1 = \{1\} = [1, 1] \text{.} \]

Remark 1 It is easy to check that the addition and multiplication of intervals are commutative, the addition of intervals is associative, and

\[ A \odot B = A \odot (-1) \odot B. \]

Since the property of subtraction of intervals cannot provide an additive inverse of a nondegenerate interval, in this article, we use the \( gH \)-difference of intervals, which is defined as follows.

Definition 1 \((gH\)-difference of intervals [10]) Let \( A \) and \( B \) be two elements of \( I(\mathbb{R}) \). The \( gH \)-difference between \( A \) and \( B \), denoted \( A \ominus gH B \), is defined by an interval \( C \) such that

\[ A = B \oplus C \text{ or } B = A \ominus C. \]

It is to be noted that for \( A = [a, \bar{a}] \) and \( B = [b, \bar{b}] \),
Definition 2 (Algebraic operations on $I(\mathbb{R})^n$) Let $\tilde{A} = (A_1, A_2, \ldots, A_n)^T$ and $\tilde{B} = (B_1, B_2, \ldots, B_n)^T$ be two elements of $I(\mathbb{R})^n$. An algebraic operation ‘$\star$’ between $\tilde{A}$ and $\tilde{B}$, denoted $\tilde{A} \star \tilde{B}$, is defined by

$$
\tilde{A} \star \tilde{B} = (A_1 \star B_1, A_2 \star B_2, \ldots, A_n \star B_n)^T,
$$

where $\star \in \{\oplus, \ominus, \odot, \oslash\}$.

Definition 3 (Dominance relation of interval [4]) For any two intervals $A$ and $B$ in $I(\mathbb{R})$,

(i) if $a < b$ and $\bar{a} \leq \bar{b}$, then $B$ is said to be dominated by $A$ and denoted by $A \preceq B$;

(ii) if either $a \leq \bar{b}$ and $\bar{a} < \bar{b}$ or $a < b$ and $\bar{a} \leq \bar{b}$ hold, then $B$ is said to be strictly dominated by $A$ and denoted by $A \prec B$;

(iii) if $B$ is not dominated by $A$, then $A \npreceq B$ and if $B$ is not strictly dominated by $A$, then $A \not\prec B$;

(iv) if $A \npreceq B$ and $B \npreceq A$, then it will be said that none of $A$ and $B$ dominates the other, or $A$ and $B$ are not comparable.

One can note that Wu [4, 6] used the term ‘superior than’ to describe the dominance relation between two intervals. However, in this article, we use the term ‘dominated by’ instead of ‘superior than.’

Lemma 1 For two elements $A$ and $B$ of $I(\mathbb{R})$,

(i) $A \preceq B \iff A \odot_{gH} B \preceq 0$

(ii) $A \npreceq B \iff A \odot_{gH} B \not\preceq 0$.

Proof See A.

Lemma 2 For an $A \in I(\mathbb{R})$,

(i) $0 \preceq A \iff (-1) \odot A \preceq 0$

(ii) $0 \npreceq A \iff (-1) \odot A \not\preceq 0$.

Proof As $0 \odot_{gH} A = (-1) \odot A$, replacing $A$ by 0 and $B$ by $A$ in Lemma 1, we get the required results.

Definition 4 (Norm on $I(\mathbb{R})$ [2]) For an $A = [a, \bar{a}]$ in $I(\mathbb{R})$, the function $\| \cdot \|_{l(I)} : I(\mathbb{R}) \rightarrow \mathbb{R}^+$, defined by

$$
\| A \|_{l(I)} = \max \{ |a|, |\bar{a}| \},
$$

is a norm on $I(\mathbb{R})$.

Definition 5 (Norm on $I(\mathbb{R})^n$ [3]) For an $\tilde{A} = (A_1, A_2, \ldots, A_n)^T$ in $I(\mathbb{R})^n$, the function $\| \cdot \|_{l(I)^n} : I(\mathbb{R})^n \rightarrow \mathbb{R}^+$, defined by

$$
\| \tilde{A} \|_{l(I)^n} = \sum_{i=1}^n \| A_i \|_{l(I)}
$$

is a norm on $I(\mathbb{R})^n$.

In this article, although we use the notions ‘$\| \cdot \|_{l(I)}$’ and ‘$\| \cdot \|_{l(I)^n}$’ to denote the norms on $I(\mathbb{R})$ and $I(\mathbb{R})^n$, respectively, we simply use the notion ‘$\| \cdot \|$’ to denote the usual Euclidean norm on $\mathbb{R}^n$.

2.2 Convexity and Basic Differential Calculus of Interval-Valued Functions

Let $X$ be a nonempty subset of $\mathbb{R}^n$. An IVF $F : X \rightarrow I(\mathbb{R})$, for each argument point $x \in X$, is presented by the ontic (see [36]) way:

$$
F(x) = \left[ f(x), \bar{f}(x) \right],
$$

where $f$ and $\bar{f}$ are real-valued functions on $X$. The functions $\underline{f}$ and $\bar{f}$ are called the lower and the upper functions of $F$, respectively.

Definition 6 (Convex IVF [4]) Let $X \subseteq \mathbb{R}^n$ be a convex set. An IVF $F : X \rightarrow I(\mathbb{R})$ is said to be a convex function if for any two vectors $x_1$ and $x_2$ in $X$,

$$
F(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 F(x_1) \oplus \lambda_2 F(x_2)
$$

for all $\lambda_1, \lambda_2 \in [0, 1]$ with $\lambda_1 + \lambda_2 = 1$.

Remark 2 (See [4]) $F$ is convex if and only if $\underline{f}$ and $\bar{f}$ are convex.

Definition 7 (Strongly convex IVF) Let $X$ be a nonempty convex subset of $\mathbb{R}^n$. An IVF $F : X \rightarrow I(\mathbb{R})$ is said to be strongly convex on $X$ if there exists a convex IVF $G : X \rightarrow I(\mathbb{R})$ and a $\sigma > 0$ such that

$$
F(x) = G(x) \oplus \frac{1}{2} \| x \|^2 \odot [\sigma, \sigma]
$$

for all $x \in X$.

Remark 3 It is to be observed that

$$
F(x) = G(x) \oplus \frac{1}{2} \| x \|^2 \odot [\sigma, \sigma]
$$

implies
Therefore,
\[ F \text{ is strongly convex } \iff G \text{ is convex } \]
\[ \iff \underline{g} \text{ and } \bar{g} \text{ are convex, by Remark 2} \]
\[ \iff f \text{ and } \bar{f} \text{ are strongly convex.} \]

**Definition 8** (gH-continuity [13]) Let \( F \) be an IVF on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \). Let \( \bar{x} \) be an interior point of \( \mathcal{X} \) and \( d \in \mathbb{R}^n \) be such that \( \bar{x} + d \in \mathcal{X} \). The function \( F \) is said to be continuous at \( \bar{x} \) if
\[
\lim_{|d| \to 0} (F(\bar{x} + d) \ominus_{gH} F(\bar{x})) = 0.
\]

**Lemma 3** An IVF \( F \) on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \) is gH-continuous if and only if \( f \) and \( \bar{f} \) are continuous.

**Proof** See B.

**Lemma 4** (Sequential criteria of gH-continuity) An IVF \( F \) on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \) is gH-continuous at a point \( \bar{x} \in \mathcal{X} \) if and only if for every sequence \( \{x_n\} \) in \( \mathcal{X} \) converging to \( \bar{x} \), the sequence \( \{F(x_n)\} \) converges to \( F(\bar{x}) \).

**Proof** See C.

**Definition 9** (gH-Lipschitz continuous IVF [37]) Let \( \mathcal{X} \subseteq \mathbb{R}^n \). An IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) is said to be gH-Lipschitz continuous on \( \mathcal{X} \) if there exists \( L > 0 \) such that
\[
\|F(x) \ominus_{gH} F(y)\|_{I(\mathbb{R})} \leq L\|x - y\| \quad \text{for all } x, y \in \mathcal{X}.
\]
The constant \( L \) is called a Lipschitz constant.

**Definition 10** (gH-derivative [27]) Let \( \mathcal{X} \subseteq \mathbb{R}^n \). The gH-derivative of an IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) at \( \bar{x} \in \mathcal{X} \) is defined by
\[
F'(\bar{x}) = \lim_{d \to 0} \frac{F(\bar{x} + d) \ominus_{gH} F(\bar{x})}{d},
\]
provided the limit exists.

**Remark 4** (See [38]) Let \( \mathcal{X} \) be a nonempty subset of \( \mathbb{R} \). The gH-derivative of an IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) at \( \bar{x} \in \mathcal{X} \) exists if the derivatives of \( f \) and \( \bar{f} \) at \( \bar{x} \) exist and
\[
F'(\bar{x}) = \left[ \min \left\{ f'(\bar{x}), \bar{f}'(\bar{x}) \right\}, \max \left\{ f'(\bar{x}), \bar{f}'(\bar{x}) \right\} \right].
\]
However, the converse is not true.

**Definition 11** (Partial gH-derivative [24]) Let \( F : \mathcal{X} \to I(\mathbb{R}) \) be an IVF, where \( \mathcal{X} \) is a nonempty subset of \( \mathbb{R}^n \). Let \( G_i \) be defined by
\[
G_i(x_i) = F(x_1, x_2, \ldots, x_{i-1}, x_i, x_{i+1}, \ldots, x_n),
\]
where \( \bar{x} = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X} \). If the gH-derivative of \( G_i \) exists at \( \bar{x}_i \), then the \( i \)-th partial gH-derivative of \( F \) at \( \bar{x} \), denoted \( D_i F(\bar{x}) \), is defined by
\[
2D_i F(\bar{x}) = G'_i(\bar{x}_i) \quad \text{for all } i = 1, 2, \ldots, n.
\]

**Definition 12** (gH-gradient [24]) Let \( \mathcal{X} \) be a nonempty subset of \( \mathbb{R}^n \). The gH-gradient of an IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) at a point \( \bar{x} \in \mathcal{X} \), denoted \( \nabla F(\bar{x}) \), is defined by
\[
\nabla F(\bar{x}) = (D_1 F(\bar{x}), D_2 F(\bar{x}), \ldots, D_n F(\bar{x}))^T.
\]

It is to be mentioned that the authors of [24] used the notations ‘\( \left( \frac{\partial F}{\partial x_i} \right)_F(\bar{x}) \)’ and ‘\( \nabla_g F(\bar{x}) \)’ for i-th partial gH-derivative and gH-gradient of \( F \) at \( \bar{x} \), respectively. However, throughout the article, we simply use the notations \( D_i F(\bar{x}) \)’ and ‘\( \nabla F(\bar{x}) \)’ for i-th partial gH-derivative and gH-gradient of \( F \) at \( \bar{x} \), respectively.

**Definition 13** (gH-Lipschitz gradient [37]) An IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) is said to have gH-Lipschitz gradient on \( \mathcal{X} \subseteq \mathbb{R}^n \) if there exists \( M > 0 \) such that
\[
\|\nabla F(x) \ominus_{gH} \nabla F(y)\|_{I(\mathbb{R})^n} \leq M\|x - y\| \quad \text{for all } x, y \in \mathcal{X}.
\]

Until now, the concepts of gH-continuity, gH-derivative, partial gH-derivative, and gH-gradient for IVFs have been discussed. In the next section, we illustrate the idea of differentiability for IVFs. This idea differentiability for IVFs is used in the rest of the paper to develop the gradient descent method of IOP.

**3 Differentiability of Interval-Valued Functions**

Behind the concept of differentiability of a function, the concept of linearity plays an important role. Thus, before exploring the concept of differentiability of an IVF, we discuss the concept of a linear IVF.

**Definition 14** (Linear IVF) Let \( \mathcal{X} \) be a linear subspace of \( \mathbb{R}^n \). A function \( F : \mathcal{X} \to I(\mathbb{R}) \) is said to be linear if
\[
F(x) = \bigoplus_{i=1}^n x_i \ominus_{gH} F(e_i) \quad \text{for all } x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X},
\]
where \( e_i \) is the \( i \)-th standard basis vector of \( \mathbb{R}^n \), \( i = 1, 2, \ldots, n \) and \( \bigoplus_{i=1}^{n} \) denotes successive addition of \( n \) number of intervals.

**Remark 5** It is noteworthy that any IVF \( F : \mathbb{R}^n \to I(\mathbb{R}) \) of the following form:

\[
F(x) = \bigoplus_{i=1}^{n} x_i \odot A_i = \bigoplus_{i=1}^{n} x_i \odot [a_i, b_i],
\]
is a linear IVF.

**Example 1** The IVF \( F(x) : \mathbb{R} \to I(\mathbb{R}) \), which is defined by

\[
F(x) = [-3, 7] \odot x
\]
is a linear IVF, which is depicted in Fig. 1 by gray-shaded region.

**Remark 6** A linear IVF \( F \) on a linear subspace \( \mathcal{X} \) of \( \mathbb{R}^n \) satisfies the following conditions:

(i) \( F(ax) = a \odot F(x) \) for all \( x \in \mathcal{X} \) and for all \( a \in \mathbb{R} \), and

(ii) for all \( x, y \in \mathcal{X} \), either

\[
F(x + y) = F(x) \oplus F(y)
\]
or none of \( F(x) \oplus F(y) \) and \( F(x + y) \) dominates the other.

For the proof, see D.

**Definition 15** \((gH\text{-differentiability})\) Let \( \mathcal{X} \) be a nonempty subset of \( \mathbb{R}^n \). An IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) is said to be \( gH \)-differentiable at a point \( \bar{x} \in \mathcal{X} \) if there exists a linear IVF \( L_{\bar{x}} : \mathbb{R}^n \to I(\mathbb{R}) \), an IVF \( E(F(\bar{x}) ; d) \) and a \( \delta > 0 \) such that

\[
(F(\bar{x} + d) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_{\bar{x}}(d) = \|d\| \odot E(F(\bar{x}) ; d)
\]
for all \( d \) such that \( \|d\| < \delta \), where \( E(F(\bar{x}) ; d) \to 0 \) as \( \|d\| \to 0 \).

If \( F \) is \( gH \)-differentiable at each point \( \bar{x} \in \mathcal{X} \), then \( F \) is said to be \( gH \)-differentiable on \( \mathcal{X} \).

**Remark 7** It is to note from Definition 15 that

\[
\lim_{\|d\| \to 0} \left[ (F(\bar{x} + d) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_{\bar{x}}(d) \right] = \lim_{\|d\| \to 0} E(F(\bar{x}) ; d)
\]
or, \( \lim_{\|d\| \to 0} (F(\bar{x} + d) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_{\bar{x}}(d) = 0 \)
or, \( \lim_{\|d\| \to 0} (F(\bar{x} + d) \ominus_{gH} F(\bar{x})) = 0 \).

Thus, every \( gH \)-differentiable IVF \( F \) is \( gH \)-continuous.

The following lemma is same as Proposition 7 in [29]. However, in [29], Proposition 7 is proved by expressing an IVF \( F \) in terms of its midpoint-radius representation, i.e., \( F = \left[ \frac{\bar{x} + \bar{r}}{2} \right] \), but in this article, to prove the following lemma, we do not use the midpoint-radius representation of an IVF.

**Lemma 5** Let \( \mathcal{X} \) be a nonempty subset of \( \mathbb{R}^n \). If an IVF \( F : \mathcal{X} \to I(\mathbb{R}) \) is \( gH \)-differentiable at \( \bar{x} \in \mathcal{X} \), then there exists a nonzero \( \lambda \) and \( \delta > 0 \) such that

\[
\lim_{\|h\| \to 0} \frac{1}{\|h\|} \odot (F(\bar{x} + \lambda h) \ominus_{gH} F(\bar{x})) = L_{\bar{x}}(h)
\]
for all \( h \in \mathbb{R}^n \) with \( |\lambda| \|h\| < \delta \), where \( L_{\bar{x}} \) is the linear IVF in Definition 15.

**Proof** See E. \( \square \)

**Theorem 1** Let an IVF \( F \) on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \) be \( gH \)-differentiable at \( \bar{x} \in \mathcal{X} \). Then, for each \( d = (d_1, d_2, \ldots, d_n)^T \in \mathbb{R}^n \), the \( gH \)-gradient of \( F \) at \( \bar{x} \) exists and the linear IVF \( L_{\bar{x}} \) in Definition 15 can be expressed by

\[
L_{\bar{x}}(d) = d^T \odot \nabla F(\bar{x}),
\]
where \( d^T \odot \nabla F(\bar{x}) = \bigoplus_{i=1}^{n} d_i \odot D_i F(\bar{x}) \).

**Proof** See F. \( \square \)

The following example shows that the converse of Theorem 1 is not true.

**Example 2** On \( \mathbb{R}^2 \), consider the following IVF:

\[
F(x) = x_1 \odot A_1 \odot x_2 \odot A_2 = x_1 \odot [-1, 1] \odot x_2 \odot [0, 2].
\]

At \( \bar{x} = (0, 1) \),

\[
F(\bar{x}) = [0, 2], D_1 F(\bar{x}) = A_1 = [-1, 1],
\]
and \( D_2 F(\bar{x}) = A_2 = [0, 2] \).

Therefore, the \( gH \)-gradient of \( F \) at \( \bar{x} = (0, 1) \) exists,

\[
\nabla F(\bar{x}) = (A_1, A_2)^T
\]
and

\[
d^T \odot \nabla F(\bar{x}) = d_1 \odot A_1 \odot d_2 \odot A_2 = F(d)
\]
for any direction $d \in \mathbb{R}^2$.

However, $F$ is not $gH$-differentiable at $\bar{x}$ because at a direction $d = (t, -t)$ with $t > 0$ and $\|d\| < \frac{1}{4}$, we obtain

$$F(\bar{x} + d) = F(t, 1 - t) = [-t, 2 - t],$$

$$F(d) = F(t, -t) = [-3t, t],$$

and

$$\lim_{t \to 0^+} \frac{1}{\|d\|} \circ (F(\bar{x} + d) \ominus_{gH} F(\bar{x})) \ominus_{gH} d^T \circ \nabla F(\bar{x})$$

$$= \lim_{t \to 0^+} \frac{1}{\sqrt{2t}} \circ \left(\left([-t, 2 - t] \ominus_{gH} [0, 2]\right) \ominus_{gH} [-3t, t]\right)$$

$$= \lim_{t \to 0^+} \frac{1}{\sqrt{2t}} \circ \left([-t, -t] \ominus_{gH} [-3t, t]\right)$$

$$= \lim_{t \to 0^+} \frac{1}{\sqrt{2t}} \circ [-2t, 2t]$$

$$= \sqrt{2} \circ [-1, 1]$$

$$\neq 0.$$
0 \preceq (x - y)^T \odot \nabla F(x) \odot_{gH} (x - y)^T \odot \nabla F(y) \text{ for all } x, y \in X.

**Proof** See I.

**Remark 11** One may think that for a $gH$-differentiable IVF $F$ on $X \subseteq \mathbb{R}^n$,

\[
(x - y)^T \odot \nabla F(x) \odot_{gH} (x - y)^T \odot \nabla F(y) = (x - y)^T \odot (\nabla F(x) \odot_{gH} \nabla F(y)) \text{ for all } x, y \in X.
\]

Unfortunately, it is not true in general even if $F$ is convex on $X$. For instance, consider the following IVF on $\mathbb{R}^2$:

\[
F(x_1, x_2) = [1, 3] \odot x_1 + [1, 4] \odot x_2 = [x_1^2 + x_1^2, 3x_1^2 + 4x_2^2].
\]

Since $f(x_1, x_2) = x_1^2 + x_1^2$ and $j(x_1, x_2) = 3x_1^2 + 4x_2^2$ are convex on $\mathbb{R}^2$, by Remark 2, $F$ is convex on $\mathbb{R}^2$.

The $gH$-gradient of $F$ is

\[
\nabla F(x) = ([2, 6] \odot x_1, [2, 8] \odot x_2)^T.
\]

Considering $x = (2, 0)$ and $y = (1, 1)$ we have

\[
(x - y)^T \odot \nabla F(x) \odot_{gH} (x - y)^T \odot \nabla F(y) = (1, -1) \odot (\odot_{gH} [2, 6] \odot, [2, 8] \odot) = [4, 12] \odot [-6, 4] = [-2, 16]
\]

and

\[
(x - y)^T \odot (\nabla F(x) \odot_{gH} \nabla F(y)) = (1, -1) \odot (([4, 12], [0, 0]) \odot_{gH} ([2, 6], [2, 8])) = (1, -1) \odot ([2, 6], [-8, -2]) = [4, 14] \neq [-2, 16].
\]

**4 Interval Optimization Problem and Its Efficient Solutions**

This section explores the connection between solutions and the $gH$-derivatives of the following IOP:

\[
\min_{x \in X \subseteq \mathbb{R}^n} F(x), \tag{2}
\]

where $F : X \to I(\mathbb{R})$ is a $gH$-differentiable function.

The concept of an efficient solution of the IOP (2) is defined below.

\[
\text{(Definition 16 (Efficient solution [5])) A point } \bar{x} \in X \text{ is called a global efficient solution of the IOP (2) if } F(\bar{x}) \neq F(x) \text{ for all } x \neq \bar{x} \in X.
\]

A point $\bar{x} \in X$ is called a local efficient solution of IOP (2) if there exists a $\delta$-neighborhood $N_\delta(\bar{x})$ of $\bar{x}$ such that $F(\bar{x}) \neq F(x)$ for all $x \neq \bar{x} \in N_\delta(\bar{x}) \cap X$.

It is to mention that here that Wu [5] named the efficient solution of this article as nondominated solution. However, throughout this article, we follow Definition 16 for an efficient solution, and in the rest of the article, by an efficient solution, we mean a global efficient solution.

Since an IOP is a special case of a fuzzy optimization problem [39], the following theorem can be considered as a corollary of Theorem 10 of [29]. Further, as differentiability of a function is a special case of Gâteaux differentiability, the following theorem can also be considered as a corollary of Theorem 4.2 of [37]. In this article, as we are dealing with interval optimization problems and $gH$-differentiability of interval-valued functions, we show the proof of the following theorem to enhance the readability. However, we prove the following theorem with a different approach than [37].

**Theorem 5 (Optimality condition)** Let $F$ be a $gH$-differentiable IVF on a nonempty subset $X$ of $\mathbb{R}^n$. If $\bar{x} \in X$ is an efficient solution of the IOP (2), then

\[
d^T \odot \nabla F(\bar{x}) \neq 0 \text{ for all } d \in \mathbb{R}^n.
\]

**Proof** Let $\bar{x} \in X$ be an efficient solution of the IOP (2). Therefore, for all $d \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ so that $\bar{x} + \lambda d \in X$, we have

\[
F(\bar{x} + \lambda d) \neq F(\bar{x}),
\]

or, $F(\bar{x} + \lambda d) \odot_{gH} F(\bar{x}) \neq 0$, by Lemma 1

or, $\lim_{\lambda \to 0^+} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda d) \odot_{gH} F(\bar{x})) \neq 0$

or, $\lim_{\lambda \to 0^-} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda d) \odot_{gH} F(\bar{x})) \neq 0$,

since $F$ is $gH$-differentiable on $X$.

Therefore, by Lemma 5 and Theorem 1, we obtain

\[
d^T \odot \nabla F(\bar{x}) \neq 0 \text{ for all } d \in \mathbb{R}^n.
\]

Replacing $d$ by $-d$ in (3), we obtain

\[
(-d)^T \odot \nabla F(\bar{x}) \neq 0
\]

or, $\odot ((-d)^T \odot \nabla F(\bar{x})) \neq 0$.

Thus, by Lemma 2,
\(0 \notin d^T \odot \nabla F(x). \quad (4)\)

By (3) and (4), we get \(0 \in d^T \odot \nabla F(x)\) for all \(d \in \mathbb{R}^n\). \(\square\)

**Remark 12** It is noteworthy that the converse of Theorem 5 is not always true even if \(F\) is a convex IVF. For instance, let us consider the following IOP:

\[
\min_{x \in \mathbb{R}} F(x),
\]

where

\[
F(x) = \begin{cases} 
[0, 3] \odot x^2 & \text{if } -1 \leq x \leq 1 \\
[0, 2] \odot [1, 1] \odot x^2 & \text{otherwise}
\end{cases}
\]

Thus, at \(x = -1\),

\[
\nabla F(-1) = [-2, 0] \implies 0 \in d \odot \nabla F(-1) \quad \text{for all } d \in \mathbb{R}.
\]

But it is notable that for \(0 < h < 2\),

\[
F(-1 + h) = [(-1 + h)^2, 3] \odot [1, 3] = F(-1).
\]

Therefore, although \(0 \in d \odot \nabla F(-1)\) for all \(d \in \mathbb{R}\), \(-1\) is not an efficient solution of the IOP (5).

**Corollary 1** Let \(F\) be a \(g_H\)-differentiable IVF on a nonempty subset \(X\) of \(\mathbb{R}^n\). If \(\bar{x} \in X\) is an efficient solution of the IOP (2), then

\[
0 \in D_i F(\bar{x}) \quad \text{for each } i \in \{1, 2, \ldots, n\}.
\]

**Proof** Let \(\bar{x} \in X\) be an efficient solution of the IOP (2). According to Theorem 5, for all \(d \in \mathbb{R}^n\), we have

\[
0 \in d^T \odot \nabla F(\bar{x}) = \bigoplus_{i=1}^n d_i \odot D_i F(\bar{x}).
\]

For each \(i \in \{1, 2, \ldots, n\}\), by considering \(d = e_i\), we obtain

\[
0 \in D_i F(\bar{x}). \quad \square
\]

### 5 \(g_H\)-Gradient Efficient Methods for Interval Optimization Problem

This section develops \(g_H\)-gradient efficient techniques to obtain the efficient solutions of the IOP (2). In the conventional gradient descent technique, to find a minimizer, we move sequentially along descent directions. Likewise, for IOP, to find an efficient solution, we attempt to move sequentially along efficient directions, defined below.

**Definition 17** \(\text{(Efficient direction).} \) Let \(X \subseteq \mathbb{R}^n\). A direction \(d \in \mathbb{R}^n\) is said to be an efficient direction of an IVF \(F : X \to I(\mathbb{R})\) at \(\bar{x} \in X\) if there exists a \(\delta > 0\) such that

(i) \(F(\bar{x}) \not\leq F(\bar{x} + \lambda d)\) for all \(\lambda \in (0, \delta)\),

(ii) there also exists a point \(x' = \bar{x} + \alpha d\) with \(x \in (0, \delta)\) and a positive real number \(\delta' \leq \alpha\) such that

\[
F(x' + \lambda d) \not< F(x') \quad \text{for all } \lambda \in (-\delta', \delta').
\]

The point \(x'\) is known as an efficient point of \(F\) in the direction \(d\).

In Fig. 3, the points \(\bar{x}\) and \(x'\), the direction \(d\), and the nonnegative real numbers \(\delta\) and \(\delta'\) of Definition 17 are illustrated on \(\mathbb{R}^2\) plane (\(n = 2\)).

**Remark 13** One may think that only the condition (i) of Definition 17 is sufficient to define an efficient direction. However, it is not true in general. Because, for \(A, B, C \in I(\mathbb{R})\), \(A \not\prec B\) and \(B \not\prec C\) does not imply \(A \not\prec C\) in general.

For instance, consider

\[
A = [4, 6], \quad B = [2, 10] \quad \text{and} \quad C = [5, 7].
\]

We, then, see that

\[
A \not\prec B \quad \text{and} \quad B \not\prec C \quad \text{but} \quad A \not\prec C.
\]

That is why the condition (ii) of Definition 17 is necessary to define an efficient direction.

**Remark 14** For the degenerate case of the IVF \(F\), i.e., for \(f(x) = \bar{f}(x) = f(x)\) for all \(x \in X\), Definition 17 reduces to the following. A direction \(d \in \mathbb{R}^n\) is said to be an efficient direction of \(f\) if there exists a \(\delta > 0\) such that

\[
f(\bar{x}) > f(\bar{x} + \lambda d) \quad \text{for all } \lambda \in (0, \delta).
\]
Fig. 3 Locations of \( \bar{x} \) and \( x' \) in Definition 17

Thus, an efficient direction for a degenerate IVF is a descent direction.

**Theorem 6** Let \( F \) be a \( gH \)-differentiable IVF on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \). Then, every direction \( d \in \mathbb{R}^n \) that satisfies

\[
0 \neq d^T \odot \nabla F(\bar{x}),
\]

is an efficient direction of \( F \) at \( \bar{x} \in \mathcal{X} \), where the corresponding efficient point \( x' \) is provided by

\[
x' = \bar{x} + x'd \text{ with } x' = \text{argeff}_{x \in \mathbb{R}^n} F(\bar{x} + xd),
\]

where by ‘argeff’ of \( F(\bar{x} + xd) \), we mean a point \( x' \) such that

\[
F(\bar{x} + xd) \notin F(\bar{x} + x'd) \text{ for all } x \in \mathbb{R}.
\]

**Proof** Let \( d \in \mathbb{R}^n \) be a direction that satisfies the relation (6). Since \( F \) is \( gH \)-differentiable at \( \bar{x} \in \mathcal{X} \), by Lemma 5 and Theorem 1, we have

\[
\lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda d) \ominus_{gH} F(\bar{x})) = d^T \odot \nabla F(\bar{x}),
\]

which implies

\[
\lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda d) \ominus_{gH} F(\bar{x})) = d^T \odot \nabla F(\bar{x}).
\]

Due to the relation (6), the last equation yields

\[
0 \neq \lim_{\lambda \to 0+} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda d) \ominus_{gH} F(\bar{x})).
\]

Therefore,

\[
0 \neq \lim_{\lambda \to 0+} \frac{1}{\lambda} \odot \left[ \min \left\{ f(\bar{x} + \lambda d) - f(\bar{x}), \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right\}, \max \left\{ f(\bar{x} + \lambda d) - \underline{f}(\bar{x}), \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right\} \right].
\]

Hence,

\[
[0, 0] \notin \left[ \min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\}, \max \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\} \right],
\]

which implies

\[
\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\} < 0.
\]

Thus, we have following two cases.

- **Case 1** Let

\[
\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\} = \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right).
\]

Then, by Eq. (8), we obtain

\[
\lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right) < 0.
\]

Therefore, there exists a \( \delta_1 > 0 \) such that for all \( \lambda \in (0, \delta_1) \),

\[
\frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right) < 0
\]

or; \( f(\bar{x} + \lambda d) - f(\bar{x}) < 0 \)

or; \( f(\bar{x} + \lambda d) < f(\bar{x}) \)

or; \( f(\bar{x}) \neq f(\bar{x}) \).

- **Case 2** Let

\[
\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\} = \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right).
\]

Then, by Eq. (8), we get

\[
\lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) < 0.
\]

Thus, there exists a \( \delta_2 > 0 \) such that for \( \lambda \in (0, \delta_2) \),

\[
\min \left\{ \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right), \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( \bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) \right) \right\} = \lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right).
\]

Then, by Eq. (8), we obtain

\[
\lim_{\lambda \to 0+} \frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right) < 0.
\]

Therefore, there exists a \( \delta_1 > 0 \) such that for all \( \lambda \in (0, \delta_1) \),

\[
\frac{1}{\lambda} \left( f(\bar{x} + \lambda d) - f(\bar{x}) \right) < 0
\]

or; \( f(\bar{x} + \lambda d) - f(\bar{x}) < 0 \)

or; \( f(\bar{x} + \lambda d) < f(\bar{x}) \)

or; \( f(\bar{x}) \neq f(\bar{x}) \).
where $\frac{1}{x}(\bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x})) < 0$

or, $\bar{f}(\bar{x} + \lambda d) - \bar{f}(\bar{x}) < 0$

or, $\bar{f}(\bar{x} + \lambda d) < \bar{f}(\bar{x})$

or, $[f(\bar{x}), \bar{f}(\bar{x})] \neq [f(\bar{x} + \lambda d), \bar{f}(\bar{x} + \lambda d)]$.

Choosing $\delta = \min\{\delta_1, \delta_2\}$, from Case 1 and Case 2, we see that for all $x \in (0, \delta)$,

\[
[f(\bar{x}), \bar{f}(\bar{x})] \neq [f(\bar{x} + \lambda d), \bar{f}(\bar{x} + \lambda d)]
\]

or, $F(\bar{x} + \lambda d) \neq F(\bar{x} + \lambda d)$.

Hence, $d$ satisfies the condition (i) of Definition 17 at $\bar{x}$. Further, let us choose an $\bar{x}'$ such that

\[
\bar{x}' = arg\inf_{x \in \mathbb{R}} F(\bar{x} + \lambda d).
\]

Therefore, there exists a $\bar{\delta} > 0$ such that for all $x \in (\bar{x}' - \bar{\delta}, \bar{x}' + \bar{\delta})$,

\[
F(\bar{x} + \lambda d) \neq F(\bar{x} + \lambda d).
\]

Considering $\lambda = x - x'$, we have

\[
F(\bar{x} + \lambda d + \lambda d) \neq F(\bar{x} + \lambda d)
\]

or, $F(\bar{x} + \lambda d) \neq F(\bar{x})$,

where $x' = \bar{x} + \lambda d$. Choosing $\delta' = \min\{\lambda', \bar{\delta}\}$, we have

\[
F(x' + \lambda d + \lambda d) \neq F(x')
\]

for all $x \in (\delta' - \bar{\delta}, \delta')$.

Therefore, $d$ satisfies the condition (ii) of Definition 17 at $\bar{x}$. Hence, $d$ is an efficient direction at $\bar{x}$. □

**Remark 15** A question may arise here: in the definition of efficient direction (Definition 17), whether or not the condition (i) can be replaced by

\[
F(\bar{x} + \lambda d) \prec F(\bar{x}) \quad \text{for all} \; \lambda \in (0, \delta).
\]

To answer, we note that if we choose the relation (9) in place of the condition (i) of Definition 17, then in the same way of proving Theorem 6, it can be proved that any $d \in \mathbb{R}^n$ that satisfies

\[
d^T \odot \nabla F(\bar{x}) < 0
\]

also holds the relation (9) and vice versa for a $gH$-differentiable IVF $F$. However, the relation (6) is more general than the relation (10). Because there are some directions, along which there exists an efficient solution of an IOP, satisfy the relation (6) but do not satisfy the relation (10).

For instance, consider the IOP:

\[
\min_{x \in \mathcal{X}} F(x) = [-1, 2] \otimes x_1^2 \oplus [3, 4] \otimes x_2^2,
\]

where $\mathcal{X} = [-10, 10] \times [-10, 10] \subseteq \mathbb{R}^2$.

In what follows, we show that $(\bar{x}_1, 0)^T \in \mathcal{X}$ is an efficient solution of the IOP (11) for any $\bar{x}_1 \in [-10, 10]$. On contrary, let there exist two nonzero numbers $h_1$ and $h_2$ with $(\bar{x}_1 + h_1, h_2)^T \in \mathcal{X}$ such that

\[
F(\bar{x}_1 + h_1, h_2) \prec F(\bar{x}_1, 0)
\]

or, $[-1, 2] \odot (\bar{x}_1 + h_1) \oplus [3, 4]h_2^2 \times [-1, 2] \odot \bar{x}_1^2$

or, $[-\bar{x}_1^2 - 2x_1 h_1 - h_1^2 + 3h_2^2, 2x_1^2 + 4x_1 h + 2h_1^2 + 4h_2^2]$

$\prec [-\bar{x}_1^2, 2x_1^2]$.

This implies

\[-2x_1 h_1 - h_1^2 + 3h_2^2 \leq 0 \quad \text{and} \quad 4x_1 h_1 + 2h_1^2 + 4h_2^2 \leq 0.
\]

Hence, $h_2^2 \leq 0$, which is not possible as $h_2 \neq 0$. So, there does not exist any $x \in \mathcal{X}$, which strictly dominates any $(\bar{x}_1, 0)^T \in \mathcal{X}$. Thus, $(\bar{x}_1, 0)^T$ is an efficient solution of the IOP (11).

Now we choose a point $\bar{x} = (3, 2)^T$ and two directions $d' = (1, -2)^T$ and $d'' = (5, -2)^T$. We also choose an $\alpha = 1$. As $F(3, 0) = [-1, 2] \odot [5, 10] = F(3, 2) = F(\bar{x})$, $\bar{x}$ is not an efficient solution of the IOP (11). But the points $x' = \bar{x} + \alpha d' = (4, 0)^T$ and $x'' = \bar{x} + \alpha d'' = (8, 0)^T$ both are efficient solutions of the IOP (11). Therefore, both $d'$ and $d''$ are efficient directions of $F$ at $\bar{x}$. Further, as the $gH$-gradient of $F$ is

\[
\nabla F(x) = (D_1 F(x), D_2 F(x))^T
\]

\[
= ([-2, 4] \odot x_1, [6, 8] \odot x_2)^T,
\]

we have $\nabla F(\bar{x}) = ([6, 12], [12, 16])^T$. Therefore,

\[
d' \odot \nabla F(\bar{x}) = [-6, 12] \odot [-32, -24] = [-38, -12]
\]

$\Rightarrow d' \odot \nabla F(\bar{x}) \prec 0$ and also, $0 \neq d' \odot \nabla F(\bar{x})$.

Again,

\[
d'' \odot \nabla F(\bar{x}) = [-30, 60] \odot [-32, -24] = [-38, 36]
\]

$\Rightarrow 0 \neq d'' \odot \nabla F(\bar{x})$.

Hence, it is clear that although both $d'$ and $d''$ are efficient directions of the IVF $F$ at the point $\bar{x}$ and satisfy the relation (6) but only $d'$ satisfies the relation (10). So, the condition (i) of Definition 17 is more general than the condition (10).

### 5.1 General gH-Gradient Efficient-Direction Method for Interval Optimization Problems

To produce the efficient solutions of the IOP (2), we provide Algorithm 1. As Algorithm 1

(i) uses $gH$-gradient at every iterative step and

[2] Springer
(ii) endeavors to find an efficient solution by sequentially moving along efficient directions, we name the method as general gH-gradient efficient-direction method. The term ‘general’ is due to the reason that we do not choose a special or a particular efficient-direction $d_k$; any general $d_k$ that satisfies \( 0 \neq d_k^T \odot \nabla F(x_k) \) will lead to reaching at an efficient point.

**Algorithm 1** General gH-gradient efficient-direction method for IOP

**Require**: Given the initial point $x_0$ and the IVF $F : \mathcal{X} \subseteq \mathbb{R}^n \rightarrow I(\mathbb{R})$

1. Set $k = 0$.
2. If $0 \in d_k^T \odot \nabla F(x_k)$ for all $d \in \mathbb{R}^n$, then Return $x_k$ as an efficient solution and Stop. Otherwise go to Step 3.
3. Find a $d_k$ such that $0 \neq d_k^T \odot \nabla F(x_k)$ and an $\alpha_k$ such that
   \[
   \alpha_k = \arg\inf_{\alpha \in \mathbb{R}^n} F(x_k + \alpha d_k)
   \]
4. Calculate $x_{k+1} = x_k + \alpha_k d_k$.
5. Set $k \leftarrow k + 1$ and go to Step 2.

In the next, we give the convergence analysis of Algorithm 1. Towards the convergent analysis, we need the following notions of the algorithmic map and the closed map regarding IVFs.

**Definition 18** (Algorithmic map [40]). Let $\mathcal{X}$ be a nonempty subset of $\mathbb{R}^n$. An algorithmic map $A$ of an algorithm is a point-to-set map on its domain $\mathcal{X}$ which describes the iterating process of the algorithm such that if the sequence $\{x_k\}$ is generated by the algorithm then $x_{k+1} \in A(x_k)$.

**Remark 16** The map $A : \mathcal{X} \rightarrow \mathcal{X}$ that generates the sequence $\{x_k\}$ in Algorithm 1 with $F(x_k) \notin F(x_{k+1})$ is an algorithmic map.

**Definition 19** (Closed map [40]). A point-to-set map $A$ from a nonempty subset $\mathcal{X}$ of $\mathbb{R}^n$ to a subset $\mathcal{Y}$ of $\mathbb{R}^m$ is said to be closed at $x \in \mathcal{X}$ if for any sequences $\{x_k\}$ and $\{y_k\}$ such that $x_k \rightarrow x$ and $y_k \rightarrow y$, where $x_k \in \mathcal{X}$ and $y_k \in A(x_k)$, we have $y \in A(x)$.

**Theorem 7** Let $F : \mathbb{R}^n \rightarrow I(\mathbb{R})$ be an IVF on a nonempty open subset $\mathcal{X}$ of $\mathbb{R}^n$ and $D$ be an element of $I(\mathbb{R}^n)$. Define a point-to-set map $L : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ by $L(x, d) = \{y : y = x + \lambda d\}$ for some $\lambda \in D$. Let $F(x + \lambda d) \notin F(y)$ for each $\lambda \in D$ and $d \neq 0$. If $F$ is gH-continuous at $x$ and $d \neq 0$, then $L$ is closed at $(x, d)$.

**Proof** Let $\{(x_k, d_k)\}$ be a sequence such that $(x_k, d_k) \rightarrow (x, d)$. Let $\{y_k\}, y_k \in L(x_k, d_k)$, be a sequence such that $y_k \rightarrow y$. To prove the theorem, we have to show that $y \in L(x, d)$.

It is to note that there exists $\lambda_k \in D$ such that $y_k = x_k + \lambda_k d_k$ for $k = 1, 2, \ldots, n$. Since $d \neq 0$, for large enough $k$, we have $d_k \neq 0$. Then, $\lambda_k = \frac{|y_k - x_k|}{|d_k|}$.

Taking the limit as $k \rightarrow \infty$, we have $\lambda_k \rightarrow \frac{|y - x|}{|d|} = \lambda$, say. Hence, $y = x + \lambda d$. Furthermore, since $\lambda_k \in D$ for each $k$ and $D$ is closed, $\lambda \in D$.

Therefore, as $\lambda \in D$, $F(x_k + \lambda_k d_k) \notin F(y_k)$ for all $k$, and $F$ is gH-continuous, by Lemma 4, evidently, we have $F(x + \lambda d) \notin F(y)$. Hence, $y \in L(x, d)$ and so, $L$ is closed. $\square$

**Remark 17** As according to Theorem 7, the map $L$ is closed, the composite map $A = L \circ D$ will be closed if the direction generating map $D : \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n$, defined by $D(x) = (x, d)$, is also closed.

**Theorem 8** (Convergence of general gH-gradient efficient method for IOP). Let $F$ be a gH-differentiable IVF on a nonempty open subset $\mathcal{X}$ of $\mathbb{R}^n$. Suppose $\Omega \subseteq \mathcal{X}$ be the set of all efficient points of $F$ and $A : \mathcal{X} \rightarrow \mathcal{X}$ be an algorithmic map of Algorithm 1. Suppose that the algorithm map $A$ produces the sequence $\{x_k\}$, which converges at $\bar{x} \in \mathcal{X}$. Also, assume that there exists $M > 0$ such that $|d_k| < M$ for all $k$. If $\bar{d}$ is an accumulation point of $d_k$, then we have

\[
0 \in \bar{d}^T \odot \nabla F(\bar{x}).
\]

**Proof** Since $\{d_k\}$ is bounded, there exists an index set $K_1$ such that $\lim_{k \in K_1} d_k = \bar{d}$ and $\lim_{k \in K_1} x_k = \bar{x}$. We have the following two cases.

- **Case 1.** If $\bar{d} = 0$. Then, (12) is trivial.
- **Case 2.** If $\bar{d} \neq 0$. If possible, let us assume that the conclusion (12) is not true. Hence,

\[
0 \notin \bar{d}^T \odot \nabla F(\bar{x}) 
\]

Since, for $k \in K_1, x_k \rightarrow \bar{x}$, also $x_{k+1} \in A(x_k)$ and $x_{k+1} \rightarrow \bar{x}$, therefore, $\bar{x} \in A(\bar{x})$. As $A$ is closed at $\bar{x}$ due to Remark 17. Thus, $F(\bar{x}) \notin F(\bar{x})$, which is a contradiction. Hence, $0 \in \bar{d}^T \odot \nabla F(\bar{x})$.

$\square$

5.2 \( \mathcal{W} \)-gH-Gradient Method for Interval Optimization Problems

In this section, we develop a particular type of efficient direction. Towards this, for two given numbers $w, w' \in [0, 1]$ with $w + w' = 1$, we define a mapping $\mathcal{W} : I(\mathbb{R}^n) \rightarrow \mathbb{R}^n$ by
\[ W(A_1, A_2, \ldots, A_n) = (w_{a_1} + w'a_1, w_{a_2} + w'a_2, \ldots, w_{a_n} + w'a_n)^T. \]

**Remark 18** It is to observe that for any two elements \( \bar{A}, \bar{B} \) in \( I(\mathbb{R})^n \),
\[ W(\bar{A} \oplus \bar{B}) = W(\bar{A}) + W(\bar{B}). \]

**Lemma 6** For an interval \( A = [a, \bar{a}] \), \((w_{a} + w'a) \odot [a, \bar{a}] \neq 0\), where \( w, w' \in [0, 1] \) with \( w + w' = 1 \).

**Proof** For an interval \( A = [a, \bar{a}] \), we have the following three cases.

- **Case 1** If \( a \geq 0 \) and \( \bar{a} \geq 0 \), then \((w_{a} + w'a) \odot [a, \bar{a}] \neq 0\) and \( \bar{a}(w_{a} + w'a) \geq 0 \). Hence,
  \[ 0 \leq (w_{a} + w'a) \odot [a, \bar{a}] \implies (w_{a} + w'a) \odot [a, \bar{a}] \neq 0. \]

- **Case 2** If \( a < 0 \) and \( \bar{a} < 0 \), then \((w_{a} + w'a) \odot [a, \bar{a}] \neq 0\) and \( \bar{a}(w_{a} + w'a) \geq 0 \). Hence,
  \[ 0 \leq (w_{a} + w'a) \odot [a, \bar{a}] \implies (w_{a} + w'a) \odot [a, \bar{a}] \neq 0. \]

- **Case 3** If \( a \leq 0 \) and \( \bar{a} > 0 \), then either \((w_{a} + w'a) = 0\) or \((w_{a} + w'a) \neq 0\).
  If \((w_{a} + w'a) = 0\), we have \((w_{a} + w'a) \odot [a, \bar{a}] = 0 \implies (w_{a} + w'a) \odot [a, \bar{a}] \neq 0. \)
  Further, if \((w_{a} + w'a) \neq 0\) then, the terms \((w_{a} + w'a)\) and \( \bar{a}(w_{a} + w'a) \) are alternative in sign. So,
  \[ (w_{a} + w'a) \odot [a, \bar{a}] \neq 0. \]

**Theorem 9** Let \( \mathcal{X} \subseteq \mathbb{R}^n \) be a nonempty set and \( F(x) : \mathcal{X} \to I(\mathbb{R}) \) be \( gh \)-differentiable at a point \( x \in \mathcal{X} \). Then, the direction \( -W(\nabla F(x)) \) is an efficient direction of \( F \) at \( x \), provided \( 0 \not\in D_i F(x) \) for at least one \( i \in \{1, 2, \ldots, n\} \).

**Proof** Let \( D_i F(x) = A_i \) for each \( i \in \{1, 2, \ldots, n\} \) and \( 0 \not\in A_i \) for at least one \( i \). Thus, \( (W(\nabla F(x)))^T \odot \nabla F(x) = \bigoplus_{i=1}^n (w_{a_i} + w'a_i) \odot [a_i, \bar{a}_i]. \)

Let \( 0 \not\in A_i \) for \( i \neq j \). Therefore, \( a_j \) and \( \bar{a}_j \) both are either positive or negative. Thus, \((w_{a_j} + w'a_j) \neq 0\), and hence, \((w_{a_j} + w'a_j) \odot [a_j, \bar{a}_j] \neq 0. \)

Therefore, according to the property of interval addition, we get
\[ (W(\nabla F(x)))^T \odot \nabla F(x) \neq 0. \quad (13) \]

If possible, let \((W(\nabla F(x)))^T \odot \nabla F(x) = 0. \) Hence, \((w_{a_i} + w'a_i) \odot [a_i, \bar{a}_i] = 0 \) for at least one \( i \in \{1, 2, \ldots, n\} \), which is not possible due to Lemma 6. Therefore,
\[ (W(\nabla F(x)))^T \odot \nabla F(x) \neq 0. \quad (14) \]

By relations (13) and (14), we obtain \( 0 \not\in -W(\nabla F(x))^T \odot \nabla F(x). \)

Hence, by Theorem 6, \( -W(\nabla F(x)) \) is an efficient direction of \( F \) at \( x \), where the corresponding efficient point \( x' \), given by
\[ x' = x - \lambda W(\nabla F(x)) \]

with \( \lambda = \text{argmin}_{\lambda \in \mathbb{R}} F(x - \lambda W(\nabla F(x))) \), satisfies the condition (i) of Definition 17 at \( x \).

Based on Theorem 9 and Corollary 2, Algorithm 1 is reduced to the following Algorithm 2. As Algorithm 2 is a particular case of Algorithm 1, we name the method as \( W-gH \)-gradient efficient-direction method. The letter ‘\( W \)’ is due to the reason that we use the mapping \( W : I(\mathbb{R})^n \to \mathbb{R}^n \) to generate efficient direction at each iteration in Algorithm 2.

**Algorithm 2** \( W-gH \)-Gradient Method for IOP

**Require:** \( w \in [0, 1] \), the initial point \( x_0 \) and the IVF \( F : \mathcal{X} \subseteq \mathbb{R}^n \to I(\mathbb{R}). \)
1. Set \( k = 0 \)
2. If \( 0 \in D_i F(x_k) \) for all \( i \in \{1, 2, \ldots, n\} \), then return \( x_k \) as an efficient solution and Stop. Otherwise go to Step 3.
3. Set \( d_k = -W(\nabla F(x_k)) \) and find an \( \alpha_k \), where
\[ \alpha_k = \text{argmin}_{\alpha \in \mathbb{R}^+} F(x_k - \alpha d_k). \]
4. Calculate \( x_{k+1} = x_k + \alpha_k d_k. \)
5. Set \( k \leftarrow k + 1 \) and go to Step 2.

**Remark 19** It is to be mentioned that in Algorithm 2, for the degenerate case of the IVF \( F \), i.e., \( f(x) = f(x) = f(x) \) for all \( x \in \mathcal{X} \), the direction \( d_k \) will be \( -\nabla f(x_k) \) and step length \( z_k \) will be \( \text{argmin}(x_k + z d_k) \) at each iteration \( k \), for any values of \( w, w' \geq 0 \) with \( w + w' = 1 \) in the mapping \( W \). Thus, in that case, \( W-gH \)-gradient efficient method is same as steepest descent method.

**Remark 20** At each iteration \( k \) in Algorithm 2, one may choose \( z_k \) such that
\[ z_k \in [z|w_{a_k} + w'a_k = 0, \text{ where } [a, \bar{a}] = F'(x_k + z_k) \}, \]
(15)
where \( F'(x_k + z d_k) \) is the \( gh \)-derivative of \( F(x_k + z d_k) \) with respect to \( z \). But it is to be kept in mind that (15) is a necessary condition to be \( z_k \) an argeff of \( F(x_k - z d_k) \), not sufficient.
Remark 21 One question may arise: for a given pair of nonnegative $w$ and $w'$ with $w + w' = 1$, if we consider the real-valued function $w f(x) + w' f(x)$ corresponding to the IOP (2) and apply the conventional steepest descent method, whether the obtained direction and step length in each iteration are identical with those obtained in each iteration of $W$-gradient efficient method?

The answer is yes only when

$$D_i F(x) = \left[ \frac{\partial f(x)}{\partial x_i}, \frac{\partial f(x)}{\partial x_i} \right] \text{ for all } i \in \{1, 2, \ldots, n\}$$

ateach $x \in X$, which is not true in general (see [23] for details).

Lemma 7 For two elements $\mathbf{A}$ and $\mathbf{B}$ in $I(\mathbb{R})^n$ and a vector $h \in \mathbb{R}^n$,

$$0 \preceq h^T \otimes (\mathbf{A} \otimes gH \mathbf{B}) \implies (W(\mathbf{A}) - W(\mathbf{B}))^T h \geq 0.$$  

Proof

$$h^T \otimes (\mathbf{A} \otimes gH \mathbf{B})$$

$$= h^T \otimes (\{A_1, A_2, \ldots, A_n\}^T \otimes gH(\{B_1, B_2, \ldots, B_n\}^T)$$

$$= h^T \otimes (\{A_1 \otimes gH B_1, A_2 \otimes gH B_2, \ldots, A_n \otimes gH B_n\})^T$$

$$= \sum_{i=1}^n h_i \otimes (A_i \otimes gH B_i).$$

According to Remark 1, without loss of generality, let us assume

$$\sum_{i=1}^n h_i \otimes (A_i \otimes gH B_i)$$

$$= \sum_{k=1}^p h_k \otimes (A_k \otimes gH B_k) \oplus \sum_{l=p+1}^n h_l \otimes (A_l \otimes gH B_l)$$

$$= \sum_{k=1}^p c_k \oplus \sum_{l=p+1}^n c_l,$$

where

$$[c_k, \bar{c}_k] = h_k \otimes (A_k \otimes gH B_k) = [(a_k - b_k)h_k, (\bar{a}_k - \bar{b}_k)h_k] \text{ for all } k \text{ and } [(\bar{c}_l, \bar{c}_l) = h_l \otimes (A_l \otimes gH B_l) = [(\bar{a}_l - \bar{b}_l)h_l, (\bar{a}_l - \bar{b}_l)h_l] \text{ for all } l \text{'s. Thus,}$$

$$0 \preceq h^T \otimes (\mathbf{A} \otimes gH \mathbf{B})$$

$$\implies 0 \preceq \sum_{k=1}^p c_k \oplus \sum_{l=p+1}^n c_l$$

$$\implies 0 \preceq \sum_{k=1}^p c_k + \sum_{l=p+1}^n c_l$$

Therefore,

$$\sum_{k=1}^p c_k + \sum_{l=p+1}^n c_l \geq 0 \text{ and } 0 \leq \sum_{k=1}^p c_k + \sum_{l=p+1}^n c_l.$$

Lemma 8 If $F$ is $gH$-differentiable IVF, then

$$\|W(\nabla F(x)) - W(\nabla F(y))\|_X \leq \|\nabla F(x) \otimes gH \nabla F(y)\|_{I(\mathbb{R})^n}$$

for all $x, y \in X$.

Proof Let

$$\nabla F(x) = \mathbf{A} = (A_1, A_2, \ldots, A_n)^T$$

$$= ([a_1, \bar{a}_1], [a_2, \bar{a}_2], \ldots, [a_n, \bar{a}_n])^T$$

and

$$\nabla F(y) = \mathbf{B} = (B_1, B_2, \ldots, B_n)^T$$

$$= ([b_1, \bar{b}_1], [b_2, \bar{b}_2], \ldots, [b_n, \bar{b}_n])^T.$$

Therefore,

$$\|\nabla F(x) \otimes gH \nabla F(y)\|_{I(\mathbb{R})^n}$$

$$\|\nabla F(x) \otimes gH \nabla F(y)\|_{I(\mathbb{R})^n}$$

$$\|\nabla F(x) \otimes gH \nabla F(y)\|_{I(\mathbb{R})^n}$$

We note that

$$W(\nabla F(x)) = (w a_1 + w' a_1, w a_2 + w' a_2, \ldots, w a_n + w' a_n)^T$$

and

$$W(\nabla F(y)) = (w b_1 + w' b_1, w b_2 + w' b_2, \ldots, w b_n + w' b_n)^T.$$

So,
\[ \|W(\nabla F(x)) - W(\nabla F(y))\|_X \]
\[ = \|\left( w(a_1 - b_1) + w(\bar{a}_1 - \bar{b}_1), w(a_2 - b_2) + w(\bar{a}_2 - \bar{b}_2), \ldots, w(a_n - b_n) + w(\bar{a}_n - \bar{b}_n) \right)\|_X \]
\[ \leq \|w(a_1 - b_1) + w(\bar{a}_1 - \bar{b}_1)\| + \|w(a_2 - b_2) + w(\bar{a}_2 - \bar{b}_2)\| + \ldots + \|w(a_n - b_n) + w(\bar{a}_n - \bar{b}_n)\| \]
\[ \leq \max\{\|a_1 - b_1\|, \|\bar{a}_1 - \bar{b}_1\|\} + \max\{\|a_2 - b_2\|, \|\bar{a}_2 - \bar{b}_2\|\} + \ldots + \max\{\|a_n - b_n\|, \|\bar{a}_n - \bar{b}_n\|\} \]
\[ = \|A_1 \otimes_{GH} B_1\|_{I(\mathbb{R}^n)} + \|A_2 \otimes_{GH} B_2\|_{I(\mathbb{R}^n)} + \ldots + \|A_n \otimes_{GH} B_n\|_{I(\mathbb{R}^n)} \]
\[ = \|\nabla F(x) \otimes_{GH} \nabla F(y)\|_{I(\mathbb{R}^n)}. \]

**Lemma 9** Let \( F \) be \( gH \)-differentiable and a strongly convex IVF on a nonempty convex subset \( \mathcal{X} \) of \( \mathbb{R}^n \). Then, there exists a \( \sigma > 0 \) such that
\[ (W(\nabla F(x)) - W(\nabla F(y)))^T(x - y) \geq \sigma\|x - y\|^2 \]
for all \( x, y \in \mathcal{X} \).

**Proof** Since \( F \) is strong convex on \( \mathcal{X} \), then there exists a convex IVF \( G \) on \( \mathcal{X} \) such that
\[ G(x) + \frac{1}{2}\|x\|^2 + [\sigma, \sigma] = F(x) \]
for all \( x \in \mathcal{X} \).

where \( \sigma > 0 \). Therefore,
\[ \nabla G(x) \otimes ([\sigma x_1, \sigma x_1], [\sigma x_2, \sigma x_2], \ldots, [\sigma x_n, \sigma x_n]) = \nabla F(x) \]
\[ \iff W(\nabla G(x)) + \sigma x = W(\nabla F(x)), \text{ by Remark 18.} \]

This implies
\[ W(\nabla G(x)) = W(\nabla F(x)) - \sigma x. \] (17)

Further, since \( F \) is \( gH \)-differentiable, \( G \) is \( gH \)-differentiable; as \( G \) is also convex IVF, by Lemma 7 and Theorem 4, we have
\[ (W(\nabla G(x)) - W(\nabla G(y)))^T(x - y) \geq 0 \]
for all \( x, y \in \mathcal{X} \).

Then, from (17), we obtain
\[ (W(\nabla F(x)) - W(\nabla F(y)) - \sigma(x - y))^T(x - y) \geq 0 \]
\[ \iff (W(\nabla F(x)) - W(\nabla F(y)))^T(x - y) - \sigma\|x - y\|^2 \geq 0 \]
\[ \iff (W(\nabla F(x)) - W(\nabla F(y)))^T(x - y) \geq \sigma\|x - y\|^2. \]

\[ \square \]

**Lemma 10** If a \( gH \)-differentiable IVF \( F \) on a nonempty subset \( \mathcal{X} \) of \( \mathbb{R}^n \) has \( gH \)-Lipschitz gradient, then for some \( L > 0 \), we have
\[ \|W(\nabla F(x)) - W(\nabla F(y))\| \leq L\|x - y\| \]
for all \( x, y \in \mathcal{X} \).

**Proof** Since \( F \) has \( gH \)-Lipschitz gradient, for some \( L > 0 \), we have
\[ \|\nabla F(x) \otimes_{GH} \nabla F(y)\|_{I(\mathbb{R}^n)} \leq L\|x - y\| \]
for all \( x, y \in \mathcal{X} \).

By Lemma 8 and the last relation, we obtain
\[ \|W(\nabla F(x)) - W(\nabla F(y))\| \leq L\|x - y\| \]
for all \( x, y \in \mathcal{X} \).

\[ \square \]

**Lemma 11** If \( F \) is a strong convex and \( gH \)-differentiable IVF on a nonempty convex subset \( \mathcal{X} \) of \( \mathbb{R}^n \) with \( gH \)-Lipschitz gradient, then for all \( x, y \in \mathcal{X} \), there exists a \( \sigma > 0 \) and an \( L > 0 \) such that
\[ (W(\nabla F(x)) - W(\nabla F(y)))^T(y - x) \geq \frac{\sigma}{L^2}\|W(\nabla F(x)) - W(\nabla F(y))\|^2 \]
for all \( x, y \in \mathcal{X} \).

**Proof** Since \( F \) has \( gH \)-Lipschitz gradient, due to Lemma 10, we get
\[ \|W(\nabla F(x)) - W(\nabla F(y))\| \leq L\|x - y\| \]
for all \( x, y \in \mathcal{X} \).

By Lemma 9 and the last inequality, we obtain
\[ (W(\nabla F(x)) - W(\nabla F(y)))^T(y - x) \geq \frac{\sigma}{L^2}\|W(\nabla F(x)) - W(\nabla F(y))\|^2 \]
for all \( x, y \in \mathcal{X} \).

\[ \square \]

**Theorem 10** (Linear convergence under strong convexity). Let \( \mathcal{X} \) be a nonempty convex subset of \( \mathbb{R}^n \). If \( F \) is a strong convex and \( gH \)-differentiable IVF with \( gH \)-Lipschitz gradient on \( \mathcal{X} \). Then, there exists a \( \sigma > 0 \) and an \( L > 0 \) such that if \( \bar{x} \in \mathcal{X} \) be an efficient solution of the IOP (2), the mapping \( H_\sigma(x) = x - \sigma W(\nabla F(x)) \) with constant step size \( x \in [0, \frac{L^2}{2\sigma}] \) satisfies
\[ \|H_\sigma(x) - H_\sigma(\bar{x})\| \leq \|x - \bar{x}\| \]
for all \( x \in \mathcal{X} \).

**Proof** Let \( F \) be strong convex and \( gH \)-differentiable IVF with \( gH \)-Lipschitz gradient on \( \mathcal{X} \). Therefore, for all \( x, y \in \mathcal{X} \), we have
\[ \|H_x(x) - H_x(y)\|^2 \]
\[ = \| (x - zW(\nabla F(x))) - (y - zW(\nabla F(y))) \|^2 \]
\[ = \| x - y \|^2 - 2zW(\nabla F(x)) - W(\nabla F(y))(x - y) \]
\[ + x^2 \| W(\nabla F(x)) - W(\nabla F(y)) \|^2 \]
\[ \leq \| x - y \|^2 - 2\| x - y \|^2 + x^2L^2\| x - y \|^2 \]
by Lemma 9 and Lemma 11
\[ \leq \| x - y \|^2, \text{since } x \in [0,2,\frac{2}{L^2}]. \]

Taking \( y = \bar{x} \), we have \( \| H_x(x) - H_x(\bar{x}) \| \leq \| x - \bar{x} \| \) for all \( x \in C \). Hence, the \( W\)-\( gH \)-gradient efficient method converges linearly.

5.3 Numerical Examples

Here we apply the proposed algorithm of \( W\)-\( gH \)-gradient efficient method on the IOPs of the following two examples and capture the efficient solutions of the IOPs.

Example 4 Consider the following IOP:
\[
\min_{x \in [-3,7]} F(x) = [1,2] \odot x^2 \odot [-8,0] \odot x \odot [3,25].
\]

The \( gH \)-gradient of \( F \) is
\[
\nabla F(x) = [2,4] \odot x \odot [-8,0] \text{ for all } x \in [-3,7].
\]
The graph of the IVF \( \mathbf{F} \) is illustrated in Fig. 4 by the gray-shaded region, and the region of the efficient solutions is marked by bold black line on \( x \)-axis. From Fig. 4, it is clear that each \( x \in [0,4] \) is an efficient solution of the IOP (18).

Since the IOP (18) has infinite number of efficient solutions, applying \( W\)-\( gH \)-gradient efficient method for various \( w \) and \( w' \), we obtain few of them, which are presented in Table 1.

As the IOP (18) in Example 4 has infinite number of efficient solutions, Table 1 shows that the algorithm of \( W\)-\( gH \)-gradient efficient method has stopped with different efficient solutions for different combinations of \( w, w' \) and initial points.

In the next example, we consider an IOP which has only one efficient solution and we show that the algorithm of \( W\)-\( gH \)-gradient efficient method will stop at the efficient solution for different combinations of \( w, w' \) and initial points.

Example 5 Consider the following IOP:
\[
\min_{x \in A \subseteq \mathbb{R}^2} F(x) = [2,6] \odot (x_1 - 2)^2 \odot [5,7] \odot (x_2 - 3)^2 \odot [5,12],
\]
where \( A = [0,6] \times [0,6] \). The \( gH \)-gradient of \( F \) is

\[
\nabla F(x) = (D_1 F(x), D_2 F(x))^T
\]
\[ = ([4,12] \odot (x_1 - 2), [10,14] \odot (x_2 - 3))^T. \]

We show that \( x = (2,3)^T \) is an efficient solution to IOP (19). On contrary, if possible, let there exist two nonzero numbers \( h_1 \) and \( h_2 \) with \( 0 \leq 2 + h_1 \leq 6 \) and \( 0 \leq 3 + h_2 \leq 6 \) such that
\[
F(2 + h_1, 3 + h_2) \prec F(2,3)
\]
or,\([2,6] \odot h_1^2 \odot [5,7] \odot h_2^2 \odot [5,12] \prec [5,12]
\]
or,\([2,6] \odot h_1^2 \odot [5,7] \odot h_2^2 \odot [5,12] \odot gH [5,12] \prec [0,0]
\]
or,\([2h_1^2 + 5h_2^2, 6h_1^2 + 7h_2^2] \prec [0,0].
\]

| \( w \) | \( w' \) | Initial point | Number of iterations | Efficient solution |
|---|---|---|---|---|
| 0 | 1 | -2 | 22 | 0 |
| 0 | 0.5 | 20 | 0 |
| 6 | 1 | 3 |
| 0.2 | 0.8 | -0.5 | 3 | 0.24 |
| 0.5 | 0.5 | 2 | 0.2719 |
| 6 | 1 | 1.9778 |
| 0.4 | 0.6 | -2 | 2 | 0.2621 |
| 0.5 | 0.5 | 1.549 |
| 6 | 1 | 0.6875 |
| 0.5 | 0.5 | -2 | 1 | 0.1333 |
| 0.5 | 0.5 | 1 | 0.2424 |
| 6 | 2 | 0.3111 |
| 0.7 | 0.3 | -2 | 3 | 0.7891 |
| 0.5 | 0.5 | 2 | 0.4489 |
| 6 | 1 | 0.8051 |
| 0.9 | 0.1 | -2 | 1 | 2.25 |
| 0.5 | 0.5 | 1 | 0.7233 |
| 6 | 2 | 3.6968 |
which is not possible. Thus, there does not exist any \(x(x) \neq \bar{x}) \in X\) such that \(F(x) \prec F(\bar{x})\). Hence, \(\bar{x}\) is an efficient solution of the IOP (19). Also, one can easily check that \(F(\bar{x}) \prec F(x)\). Thus, \(\bar{x}\) is the only efficient solution of the IOP (19).

The IVF \(F\) is depicted in Fig. 5 by the gray-shaded surface, and the efficient solution is pointed by black dot on \(xy\)-plane. Figure 5 shows that \(\bar{x}\) is the only efficient solution of \(F\) on \(X\).

From Table 2, we see that for every combination of \(w\), \(w'\) and initial point, the \(W\)-\(gH\)-gradient efficient method has stopped at the efficient solution \(x = (2, 3)^T\).

### 6 Application

This section applies the \(W\)-\(gH\)-gradient efficient method in solving the least-square problems for interval-valued data. Suppose a set of \(n\) pairs of data \((X_k, Y_k)\) is given, where \(Y_k \in I(\mathbb{R})\) is the corresponding interval-valued output of \(X_k \in I(\mathbb{R})\) for all \(k \in \{1, 2, \ldots, n\}\). We attempt to fit a function \(\mathcal{H}(\cdot; \beta) : I(\mathbb{R}) \rightarrow I(\mathbb{R})\), where \(\beta \in \mathbb{R}^l\) is a parameter vector such that \(\mathcal{H}(X_k; \beta)\) will be one of the best approximations of \(Y_k\) for all \(k \in \{1, 2, \ldots, n\}\). By ‘one of the best approximations,’ we mean that \(\mathcal{H}(X; \beta)\) gives a sum square error that is nondominated. Evidently, if \(\hat{\beta}\) is an efficient solution of the following IOP:

\[
\min_{\beta \in \mathbb{R}^l} \mathcal{E}(\beta) = \sum_{k=1}^n \left( \mathcal{H}(X_k; \beta) \otimes_{gH} Y_k \right) \odot \left( \mathcal{H}(X_k; \beta) \otimes_{gH} Y_k \right),
\]

(20)

then \(\mathcal{H}(X; \hat{\beta})\) can be considered as an efficient choice of the approximating function \(\mathcal{H}(X; \beta)\).

It is noteworthy that the error function \(\mathcal{E}\) and the function \(\mathcal{H}(X_k; \cdot)\) are IVFs from \(\mathbb{R}^l\) to \(I(\mathbb{R})\) for all \(X_k\). The partial \(gH\)-derivative of \(\mathcal{E}\) with respect to \(\beta_i\) is

\[
D_i \mathcal{E}(\beta) = 2 \odot \sum_{k=1}^n \left( \mathcal{H}(X_k; \beta) \otimes_{gH} Y_k \right) \odot D_i \mathcal{H}(X_k; \beta)
\]

for all \(i \in \{1, 2, \ldots, l\}\). Hence, by applying the \(W\)-\(gH\)-gradient efficient method on the IOP (20), one can easily obtain an efficient parameter vector \(\hat{\beta}\) for the function \(\mathcal{H}(\cdot; \beta)\). For examples, we consider the following two types of fitting with interval-valued data.

#### 6.1 Polynomial Fitting

Let us consider a set of 21 pairs of interval-valued data that are displayed in Table 3. We attempt fit a polynomial function \(\mathcal{H}^l(\cdot; \beta) : I(\mathbb{R}) \rightarrow I(\mathbb{R})\), defined by

\[
\mathcal{H}^l(X; \beta) = \beta_1 \odot C \odot \beta_2 \odot X \odot \beta_3 \odot X^2,
\]

where \(C\) is a constant interval and \(\beta = (\beta_1, \beta_2, \beta_3) \in \mathbb{R}^3\). Therefore, for each \(k = 1, 2, \ldots, 21\), the partial \(gH\)-derivatives of \(\mathcal{H}^l(X_k; \cdot)\) with respect to \(\beta_1, \beta_2, \) and \(\beta_3\) are

| Table 2: Output of Algorithm 2 to find efficient solutions of IOP (19) |
|-----------------|-----------------|-----------------|-----------------|
| \(w\)           | \(w'\)          | Initial point   | Number of iterations | Efficient solution |
| 0.1             | 0.9             | \((0.6)^T\)     | 30                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 21                | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 13                | \((2, 3)^T\)        |
| 0.3             | 0.7             | \((0.6)^T\)     | 16                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 15                | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 13                | \((2, 3)^T\)        |
| 0.4             | 0.6             | \((0.6)^T\)     | 12                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 10                | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 9                 | \((2, 3)^T\)        |
| 0.6             | 0.4             | \((0.6)^T\)     | 10                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 9                 | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 11                | \((2, 3)^T\)        |
| 0.9             | 0.1             | \((0.6)^T\)     | 20                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 19                | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 18                | \((2, 3)^T\)        |
| 1.0             | 0.0             | \((0.6)^T\)     | 25                | \((2, 3)^T\)        |
|                 |                 | \((5.2)^T\)     | 21                | \((2, 3)^T\)        |
|                 |                 | \((2.5, 2.5)^T\)| 22                | \((2, 3)^T\)        |

Fig. 5 Interval-valued function and efficient solution of the IOP (19) in Example 5
Fig. 6

\[ D_1 \mathcal{H}^1(X_k; \beta) = C, \]
\[ D_2 \mathcal{H}^1(X_k; \beta) = X_k, \]
and \( D_3 \mathcal{H}^1(X_k; \beta) = X_k, \) respectively.

Considering \( C = [1.70, 12.00] \) and an initial \( \beta = (6, -8, 9)^T, \) and applying Algorithm 2 in Matlab R2015a platform with Intel Core i5-2430M, 2.40 GHz CPU, 3 GB RAM, 32-bit Windows 7 environment on the IOP (20) corresponding to the function \( \mathcal{H}^1(\cdot; \beta) \) with \( w = 0.3, \) we obtain the value of \( \hat{\beta} \) equal to \((-0.0876, -0.2974, 0.5458)^T\) in 24 iterations. With \( w = 0.5, \) we obtain the value of \( \hat{\beta} \) equal to \((-0.0896, -0.2777, 0.5352)^T\) in 10 iterations.

Both the figures of Fig. 6 show the comparison of the actual interval-valued output \( Y_k \) with the estimated output \( \mathcal{H}^1(X_k; \hat{\beta}) \) of the interval-valued data \( X_k \) in polynomial fitting for the values of \( w \) equal to 0.3 and 0.5, respectively, for \( k = 1, 2, \ldots, 21. \) The common portions of \( Y_k \) with \( \mathcal{H}^1(X_k; \hat{\beta}) \) are depicted by orange regions, whereas the extended portions of \( Y_k \) and \( \mathcal{H}^1(X_k; \hat{\beta}) \) are illustrated by red and yellow regions, respectively.

### 6.2 Logistic Fitting

Let us consider a set of 15 pairs of interval-valued data that are displayed in Table 4, and we fit a logistic function \( \mathcal{H}^2(\cdot; \beta) : I(\mathbb{R}) \rightarrow I(\mathbb{R}) \) defined by

\[ \mathcal{H}^2(X_k; \beta) = \frac{1}{1 + e^{-(\beta_1 \circ \odot_1 \circ_0 \beta_2 \circ_0 X_k)}}, \]

where \( C \) is a constant interval and \( \beta = (\beta_1, \beta_2) \in \mathbb{R}^2. \) Thus, the partial \( gH \)-derivatives of \( \mathcal{H}^2(X_k; \cdot) \) with respect to \( \beta_1 \) and \( \beta_2 \) are

\[ D_1 \mathcal{H}^2(X_k; \beta) = \frac{1}{(1 + e^{-(\beta_1 \circ \odot_1 \circ_0 \beta_2 \circ_0 X_k)})^2} \odot C, \]
and
\[ D_2 \mathcal{H}^2(X_k; \beta) = \frac{1}{(1 + e^{-(\beta_1 \circ \odot_1 \circ_0 \beta_2 \circ_0 X_k)})^2} \odot X_k, \]
respectively, for all \( k = 1, 2, \ldots, 15. \)

Taking \( C = [1.30, 3.40] \) and the initial value of \( \beta \) as \((7, -4)^T, \) and applying Algorithm 2 in Matlab R2015a platform with Intel Core i5-2430M, 2.40 GHz CPU, 3 GB RAM, 32-bit Windows 7 environment with \( w = 0.7 \) on the IOP (20) corresponding to the function \( \mathcal{H}^2(\cdot; \beta), \) we obtain the value of \( \beta \) as \((3.3940, 8.5835)^T\) in 10 iterations.

![Fig. 6](image)

(a) For \( w = 0.3 \)

![Fig. 6](image)

(b) For \( w = 0.5 \)

**Fig. 6** Comparison of the actual \( Y_k \) with the estimated output \( \mathcal{H}^1(X_k; \hat{\beta}) \), by the polynomial fitting, of the interval-valued data \( X_k \) in Table 3, \( k = 1, 2, \ldots, 21. \)

| \( X_k \) | \( Y_k \) | \( X_k \) | \( Y_k \) |
|---|---|---|---|
| [−1.99, −1.86] | [1.21, 2.60] | [−1.80, −1.66] | [0.72, 2.00] |
| [−1.61, −1.54] | [0.43, 1.67] | [−1.48, −1.33] | [0.20, 1.39] |
| [−1.29, −1.14] | [0.01, 1.22] | [−1.08, −0.96] | [−0.28, 1.00] |
| [−0.89, −0.73] | [−0.60, 0.60] | [−0.68, −0.51] | [−0.8, 0.34] |
| [−0.45, −0.29] | [−0.95, 0.12] | [−0.23, −0.03] | [−1.12, −0.10] |
| [0.01, 0.15] | [−1.24, 0.00] | [0.19, 0.33] | [−1.3, 0.01] |
| [0.39, 0.54] | [−1.32, 0.02] | [0.60, 0.74] | [−1.25, 0.10] |
| [0.79, 0.93] | [−1.05, 0.13] | [0.98, 1.13] | [−0.96, 0.20] |
| [1.19, 1.33] | [−0.69, 0.40] | [1.39, 1.54] | [−0.34, 0.78] |
| [1.60, 1.74] | [−0.03, 1.09] | [1.79, 1.95] | [0.24, 1.47] |
| [2.00, 2.15] | [0.55, 1.80] |
Figure 7 shows the comparison of the actual interval-valued output $Y_k$ with the estimated output $\mathcal{H}^2(X_k; \hat{\beta})$ of the interval-valued data $X_k$ in logistic fitting for the value of $w = 0.7$, $k = 1, 2, \ldots, 12$. The common portions of $Y_k$ with $\mathcal{H}^2(X_k; \hat{\beta})$ are illustrated by orange regions. The extended portions of $Y_k$ and $\mathcal{H}^2(X_k; \hat{\beta})$ are depicted by red and yellow regions, respectively.

### 6.3 Real-Life Application

In this section, we apply the proposed $\mathcal{V}$-$gH$-gradient efficient method to fit a linear model for the cardiological data [41] illustrated in Table 5, where $X, X'$, and $Y$ denote systolic blood pressure, diastolic blood pressure, and pulse rate of patients (suffering different types of illness). It is to mention that, instead of taking pair of intervals, we have taken 3 tuples of intervals. Our aim is to fit a linear function $\mathcal{H}^2 : I(\mathbb{R}) \rightarrow I(\mathbb{R})$ defined by

$$\mathcal{H}^2(X_k; X'_k; \beta) = \beta_1 \odot C \oplus \beta_2 \odot X \oplus \beta_3 \odot X'$$

where $C$ is a constant interval and $\beta = (\beta_1, \beta_2, \beta_3)^T \in \mathbb{R}^3$. Therefore, for each $k = 1, 2, \ldots, 21$, the partial $gH$-derivatives of $\mathcal{H}^2(X_k; X'_k; \cdot)$ with respect to $\beta_1, \beta_2$ and $\beta_3$ are

$$D_1 \mathcal{H}^2(X_k; X'_k; \beta) = C,$n$$D_2 \mathcal{H}^2(X_k; X'_k; \beta) = X_k,$

and $D_3 \mathcal{H}^2(X_k; X'_k; \beta) = X'_k$, respectively.

#### Table 4 Data for logistic fitting

| $X_k$             | $Y_k$            |
|-------------------|------------------|
| $[-2.70, -2.55]$  | $[10^{-8}, 13 \times 10^{-5}]$ |
| $[-2.51, -2.35]$  | $[6 \times 10^{-8}, 10^{-3}]$  |
| $[-2.32, -2.23]$  | $[3 \times 10^{-7}, 0.003]$   |
| $[-2.19, -2.02]$  | $[10^{-6}, 0.02]$    |
| $[-2.00, -1.83]$  | $[6 \times 10^{-6}, 0.078]$  |
| $[-1.79, -1.65]$  | $[3 \times 10^{-5}, 0.301]$  |
| $[-1.60, -1.42]$  | $[25 \times 10^{-5}, 0.760]$ |
| $[-1.39, -1.20]$  | $[15 \times 10^{-4}, 0.930]$ |
| $[-1.16, -0.98]$  | $[0.012, 0.980]$    |
| $[-0.94, -0.72]$  | $[0.080, 0.990]$   |
| $[-0.69, -0.50]$  | $[0.440, 0.998]$   |
| $[-0.46, -0.28]$  | $[0.880, 0.999]$   |
| $[-0.24, -0.02]$  | $[0.981, 1.000]$   |
| $[0.02, 0.21]$    | $[0.997, 0.999]$   |
| $[0.26, 0.44]$    | $[0.998, 1.000]$   |

#### Table 5 Cardiological data

| $X_k$ (systolic BP) | $X'_k$ (diastolic BP) | $Y_k$ (pulse rate) |
|---------------------|-----------------------|--------------------|
| $[90, 100]$         | $[50, 70]$            | $[44, 68]$         |
| $[90, 130]$         | $[70, 90]$            | $[60, 72]$         |
| $[140, 180]$        | $[90, 100]$           | $[56, 90]$         |
| $[110, 142]$        | $[80, 108]$           | $[70, 112]$        |
| $[90, 100]$         | $[50, 70]$            | $[54, 72]$         |
| $[130, 160]$        | $[80, 110]$           | $[70, 100]$        |
| $[130, 160]$        | $[76, 90]$            | $[72, 100]$        |
| $[110, 190]$        | $[70, 110]$           | $[76, 98]$         |
| $[138, 180]$        | $[90, 110]$           | $[86, 96]$         |
| $[110, 150]$        | $[78, 100]$           | $[86, 100]$        |
| $[60, 100]$         | $[140, 150]$          | $[63, 75]$         |
7 Comparison

Although as an application, we have applied the proposed $\mathcal{W}gH$-gradient efficient method to solve the least-square problems with interval-valued data, the main objective of this study was to develop the basic gradient method to obtain efficient solutions to interval optimization problems. No emphasis is laid in this article to provide a better method for regression analysis with interval-valued data. From the literature, it is to observe that all the existing methods to obtain the efficient solutions to an interval optimization problem have been developed either from the perspective of conventional bi-objective optimization, i.e., using the explicit forms of the lower and upper functions of the interval-valued objective function or by using the parametric representation of the interval-valued objective function. However, those two techniques are quite restrictive (please see Sect. 1.1 for details). The proposed methods are superior in the sense that these are applicable to all the differentiable IOPs regardless of whether or not

(i) the interval-valued objective function can be expressed parametrically, or
(ii) the explicit forms of the lower and upper functions of the interval-valued objective function can be found.

8 Conclusion and Future Directions

In this article, a general $gH$-gradient efficient-direction method and a $\mathcal{W}gH$-gradient efficient method for IOPs have been developed. The convergence analysis and the algorithmic implementations of both methods have been presented. It has been shown that the $\mathcal{W}gH$-gradient efficient method converges linearly for a strongly convex interval-valued objective function. In the sequel, a few topics of calculus and convexity of IVFs have been proposed, which were needed for the development of the methods. For a $gH$-differentiable IVFs, the relation between convexity and the gradient of a $gH$-differentiable IVF and an optimality condition of an IOP have been derived. Further, a notion of efficient direction for IVFs has been introduced, which is used to develop the general $gH$-gradient efficient and $\mathcal{W}gH$-gradient efficient methods.

The proposed $\mathcal{W}gH$-gradient efficient method has been used to solve the least-square problems with interval-valued data. The application has been exemplified by a polynomial fitting and a logistic curve fitting. Also, as a practical example, the proposed $\mathcal{W}gH$-gradient efficient method is applied to fit a linear regression model for the practical field cardiological interval-valued data.

After the appearance of this theoretical work, our future research will shortly appear with the rigorous simulation aspect of the proposed method. The application of the proposed methods to different physical real-world problems will be shown in detail. Mainly, a rigorous error analysis of the red and yellow portions of Figs. 6 and 7 will be performed.

In connection with the proposed research, few more future directions are as follows.

- One may attempt to develop an appropriate approach to find the step lengths $z_k$ in Algorithms 1 and 2. Towards this, one can also concentrate on developing exact or inexact line search techniques for IOPs.
- It is notable that in the definition of strongly convex IVFs (Definition 7), we have taken a degenerate interval $[\sigma, \sigma]$ for some $\sigma > 0$. One can attempt to make a generalization of the used $[\sigma, \sigma]$ to $[\sigma_1, \sigma_2]$.
- Analysis of the proposed method for more flexible IVFs, especially for nonconvex IVFs can be performed in the future.
- One may attempt to apply the proposed logistic regression on the classification problems with interval-valued data.
- Applications of least-square technique in practical problems can be performed.
- Also, in the future, one can try to develop the proposed methods of this article based on constrained interval analysis [42].
- Future research can be done on applying the proposed methods in constrained IOPs.
- As IVFs are special cases of fuzzy-valued functions and IOPs are special cases of fuzzy optimization problems, similar to the proposed results, one can derive results

| Actual pulse rate ($Y_i$) | Estimated pulse rate $\left( \mathcal{H}^2(\mathbf{x}_i; \mathbf{x}_j' ; \beta) \right)$ |
|---------------------------|-----------------------------------------------|
|                           | For $w = 0.2$       | For $w = 0.5$       | For $w = 0.8$       |
| [44, 68]                  | [48, 58]            | [51, 61]            | [56, 67]            |
| [60, 72]                  | [54, 75]            | [57, 79]            | [62, 87]            |
| [56, 90]                  | [78, 96]            | [83, 102]           | [91, 113]           |
| [70, 112]                 | [64, 84]            | [68, 89]            | [74, 97]            |
| [54, 72]                  | [48, 58]            | [51, 61]            | [56, 67]            |
| [70, 100]                 | [72, 91]            | [76, 97]            | [83, 106]           |
| [72, 100]                 | [71, 86]            | [75, 91]            | [82, 100]           |
| [76, 98]                  | [61, 103]           | [65, 109]           | [71, 120]           |
| [86, 96]                  | [78, 99]            | [82, 105]           | [90, 115]           |
| [86, 100]                 | [63, 85]            | [67, 90]            | [73, 99]            |
| [63, 75]                  | [61, 79]            | [64, 84]            | [68, 90]            |
for fuzzy-valued functions and fuzzy optimization problems.

Appendix 1: A Proof of Lemma 1

Proof

(i) \[A \preceq B \iff [a, \bar{a}] \preceq [b, \bar{b}]\]

\[\iff a \leq b \text{ and } \bar{a} \leq \bar{b}\]

\[\iff a - b \leq 0 \text{ and } \bar{a} - \bar{b} \leq 0\]

\[\iff \min\{a - b, \bar{a} - \bar{b}\} \leq 0\]

\[\iff A \odot gH B \preceq 0.\]

(ii) According to the Definition 15, \(A = B\) \iff \(A \odot gH B = 0.\) If \(A \neq B,\) then

\[A \nless B \iff [a, \bar{a}] \nless [b, \bar{b}]\]

\[\iff \max\{a - b, \bar{a} - \bar{b}\} > 0\]

\[\iff A \odot gH B \nless 0.\]

Appendix 2: Proof of Lemma 3

Proof Let \(F\) be \(gH\)-continuous at a point \(\bar{x}\) of the set \(\mathcal{X}\). Thus, for any \(d \in \mathbb{R}^n\) such that \(\bar{x} + d \in \mathcal{X}\),

\[\lim_{|d| \to 0} (F(\bar{x} + d) \odot_{gH} F(\bar{x})) = 0,\]

which implies

\[\lim_{|d| \to 0} \left(\left[\begin{array}{c} f(\bar{x} + d) \\ \bar{f}(\bar{x} + d) \end{array}\right] \odot_{gH} \left[\begin{array}{c} f(\bar{x}) \\ \bar{f}(\bar{x}) \end{array}\right]\right) = [0, 0].\]

Hence, by the definition of \(gH\)-difference, we have

\[\lim_{|d| \to 0} (f(\bar{x} + d) - f(\bar{x})) \to 0 \text{ and } \lim_{|d| \to 0} (\bar{f}(\bar{x} + d) - \bar{f}(\bar{x})) \to 0,\]

i.e., \(f\) and \(\bar{f}\) are continuous at \(\bar{x} \in \mathcal{X}\).

Conversely, let the functions \(f\) and \(\bar{f}\) be continuous at \(\bar{x} \in \mathcal{X}\). If possible, let \(F\) be not \(gH\)-continuous at \(\bar{x}\). Then, as \(|d| \to 0, (F(\bar{x} + d) \odot_{gH} F(\bar{x})) \not\to 0.\) Therefore, as \(|d| \to 0\) at least one of the functions \((f(\bar{x} + d) - f(\bar{x}))\) and \((\bar{f}(\bar{x} + d) - \bar{f}(\bar{x}))\) does not tend to 0. So it is clear that at least one of the functions \(f\) and \(\bar{f}\) is not continuous at \(\bar{x}\). This contradicts the assumption that the functions \(f\) and \(\bar{f}\) both are continuous at \(\bar{x}\). Hence, \(F\) is \(gH\)-continuous at \(\bar{x}.\)

Appendix 3: Proof of Lemma 4

Proof The result is followed by Lemma 3 and the fact that the real-valued functions \(f\) and \(\bar{f}\) are continuous at the point \(\bar{x} \in \mathcal{X}\) if and only if for every sequence \(\{n_{\alpha}\}\) in \(\mathcal{X}\) converging to \(\bar{x}\), the sequences \(\{f(n_{\alpha})\}\) and \(\{\bar{f}(n_{\alpha})\}\) converge to \(f(\bar{x})\) and \(\bar{f}(\bar{x})\), respectively.

Appendix 4: Proof for the Problem in Remark 6

Proof Due to Definition 14, \(F(y) = \bigoplus_{i=1}^n y_i \odot F(e_i)\) for all \(y = (y_1, y_2, \ldots, y_n)^T \in \mathcal{X}.\)

Replacing \(y\) by \(\lambda x,\) where \(\lambda \in \mathbb{R}\) and \(x = (x_1, x_2, \ldots, x_n)^T \in \mathcal{X},\) we have

\(F(\lambda x) = \bigoplus_{i=1}^n \lambda x_i \odot F(e_i) = \lambda \odot \bigoplus_{i=1}^n x_i \odot F(e_i) = \lambda \odot F(x)\)

Let us consider a pair of elements \(x = (x_1, x_2, \ldots, x_n)^T\) and \(y = (y_1, y_2, \ldots, y_n)^T\) of \(\mathcal{X}\). If the corresponding \(x_i\) and \(y_i\) are of same sign for all \(i \in \{1, 2, \ldots, n\},\) then

\(F(x) \odot F(y) = \bigoplus_{i=1}^n x_i \odot F(e_i) \odot \bigoplus_{i=1}^n y_i \odot F(e_i)\)

\[= \bigoplus_{i=1}^n (x_i + y_i) \odot F(e_i),\]

by Remark 1

\(= F(x + y).\)

Let a few of the \(x_i\)s and the corresponding \(y_i\)s are of different signs. Without loss of generality, let the first \(p\) number of \(x_i\)s are of the same signs with their corresponding \(y_i\)s, thereafter consecutive \(q - p\) numbers of \(x_i\)s are nonnegative but corresponding \(y_i\)s are nonpositive, and the last \(n - q\) numbers of \(x_i\)s are nonpositive but corresponding \(y_i\)s are nonnegative, where \(p \leq q \leq n.\) Also, let

\(F(e_i) = A_i = [a_i, \bar{a}_i],\) for all \(e_i, i \in \{1, 2, \ldots, n\}\)

Then, we have
\( F(x) \oplus F(y) \)
\[
= \left( \oplus_{k=1}^q (x_k + y_k) \circ F(e_k) \oplus \bigoplus_{l=p+1}^q (x_l + y_l) \circ F(e_l) \right) \\
\oplus \bigoplus_{m=q+1}^n (x_m + y_m) \circ F(e_m) \\
= \left( \oplus_{k=1}^q (x_k + y_k) \circ F(e_k) \oplus \bigoplus_{l=p+1}^q (x_l + y_l) \circ F(e_l) \right) \\
\oplus \bigoplus_{m=q+1}^n (y_m \circ F(e_m) \oplus u_m \circ F(e_m))
\]
where \( B = \bigoplus_{k=1}^q (x_k + y_k) \circ F(e_k), \forall y \) and \( u_m = -x_m \)
for all \( l = p + 1, p + 2, \ldots, q \) and \( m = q + 1, q + 2, \ldots, r \)
\[
= B \oplus \left[ \sum_{l=p+1}^q (\alpha_l \circ \gamma_l) + \sum_{m=q+1}^n (\alpha_m \circ \gamma_m) \right]
\]
where \( \alpha = \sum_{l=p+1}^q (\alpha_l \circ \gamma_l) + \sum_{m=q+1}^n (\alpha_m \circ \gamma_m) \)
and \( \gamma = \sum_{m=q+1}^n (\alpha_m \circ \gamma_m) + \sum_{m=q+1}^n (\alpha_m \circ \gamma_m) \)

Further, we note that
\[
F(x + y) = \left( \oplus_{k=1}^q (x_k + y_k) \circ F(e_k) \oplus \bigoplus_{l=p+1}^q (x_l + y_l) \circ F(e_l) \right) \\
\oplus \bigoplus_{m=q+1}^n (x_m + y_m) \circ F(e_m)
\]
\[
= B \oplus \bigoplus_{l=p+1}^q (x_l + y_l) \circ F(e_l) + \bigoplus_{m=q+1}^n (y_m - u_m) \circ F(e_m)
\]
\[
= B \oplus \bigoplus_{l=p+1}^q (x_l + y_l) \circ F(e_l) + \bigoplus_{m=q+1}^n (y_m - u_m) \circ F(e_m)
\]
Again, without loss of generality, we let that among \( q-p \) numbers of \( x_l - y_l \) the first \( p-r \) elements are nonnegative and the rest are nonpositive, where \( p \leq r \).
Similarly, also we let that among \( n-q \) numbers of \( y_m - u_m \) the first \( s-q \) elements are nonnegative and rest are nonpositive, where \( q \leq s \leq n \).
Then, we have \( F(x + y) = B \oplus D \), where
\[
\delta = \sum_{l=p+1}^q \alpha_l (x_l - y_l) + \sum_{m=q+1}^n \alpha_m (y_m - u_m)
\]
and
\[
\mathcal{I} \delta = \sum_{l=p+1}^r \alpha_l (x_l - v_l) + \sum_{l=r+1}^q \alpha_l (x_l - v_l) + \sum_{m=q+1}^s \alpha_m (y_m - u_m) + \sum_{m=s+1}^n \alpha_m (y_m - u_m)
\]

**Appendix 5: Proof of Lemma 5**

**Proof** Let \( F \) be \( gH \)-differentiable at \( \bar{x} \in \mathcal{X} \). By Definition 15, there exists a \( \delta > 0 \) such that
\[
(F(\bar{x} + d) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_\delta(d) = ||d|| \circ E(F(\bar{x}); d)
\]
for all \( ||d|| < \delta \), where \( E(F(\bar{x}); d) \rightarrow 0 \) as \( ||d|| \rightarrow 0 \).
Considering \( d = \lambda h \) for \( \lambda \neq 0 \) and \( h \in \mathbb{R}^n \) with \( ||\lambda|| ||h|| < \delta \), from Eq. (21), we obtain
\[
\frac{1}{\lambda} \odot \left[ (F(\bar{x} + \lambda h) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_\lambda(\lambda h) \right]
\]
\[
= \frac{1}{\lambda} \odot E(F(\bar{x}); h)
\]
and
\[
\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda h) \ominus_{gH} F(\bar{x})) \ominus_{gH} L_\lambda(\lambda h) = 0,
\]

Hence,
\[
\lim_{\lambda \rightarrow 0} \frac{1}{\lambda} \odot (F(\bar{x} + \lambda h) \ominus_{gH} F(\bar{x})) = L_\lambda(h).
\]
Proof of Theorem 1

Proof If \(d = (0,0,\ldots,0)^T \in \mathbb{R}^n\), both the sides of the equation (1) become 0. Hence, Eq. (1) is trivially true for \(d = (0,0,\ldots,0)^T\).

Let us assume that \(d \neq (0,0,\ldots,0)^T\). Since \(F\) is \(gH\)-differentiable at \(\bar{x}\), by Lemma 5 there exists a \(\delta > 0\) such that

\[
\lim_{\lambda \to 0} \frac{1}{\lambda} \left( \nabla F(\bar{x}) \otimes gH F(\bar{x}) \right) = L_\lambda(h),
\]

where \(d = \lambda h\) with \(\lambda \neq 0, h \in \mathbb{R}^n\) and \(|\lambda||h| < \delta\).

Taking \(h = e_i\), the \(i\)-th unit vector in the standard basis of \(\mathbb{R}^n\), from Eq. (23) we obtain

\[
\lim_{\lambda \to 0} \frac{1}{\lambda} \left( \nabla F(\bar{x}) \otimes gH F(\bar{x}) \right) = L_\lambda(e_i)
\]
or, \(D_iF(\bar{x}) = L_\lambda(e_i)\).

Therefore, all the \(i\)-th partial \(gH\)-derivative \(D_iF(\bar{x})\) of \(F\) at \(\bar{x}\) exist. Hence, the gradient of \(F\) at \(\bar{x}\) exists.

According to Definition 14 of linear IVF, we get

\[
L_\lambda(d) = L_\lambda(d_1, d_2, \ldots, d_n) = d_1 \otimes L_\lambda(e_1) \oplus d_2 \otimes L_\lambda(e_2) \oplus \cdots \oplus d_n \otimes L_\lambda(e_n)
\]

\[
= \bigoplus_{i=1}^{n} d_i \otimes D_iF(\bar{x}) = d^T \otimes \nabla F(\bar{x}).
\]

\(\square\)

Appendix 6: Proof of Theorem 2

Proof Consider an arbitrary \(\bar{x} \in X\). Let for any \(d \in N_\delta(\bar{x}) \cap X\),

\[
F(\bar{x} + d) = F(\bar{x}) \oplus F(d).
\]

Since

\[
\nabla F(\bar{x}) = (F(e_1), F(e_2), \ldots, F(e_n))^T,
\]

for any \(d \in N_\delta(\bar{x}) \cap X\), we obtain

\[
d^T \otimes \nabla F(\bar{x}) = \bigoplus_{i=1}^{n} d_i \otimes F(e_i) = F(d).
\]

Therefore,

\[
\lim_{\lambda \to 0} \frac{1}{\lambda} \left( (\nabla F(\bar{x}) \otimes gH F(\bar{x})) \right) \otimes gH d^T \otimes \nabla F(\bar{x})
\]

\[
= \lim_{\lambda \to 0} \frac{1}{\lambda} \left( (\nabla F(\bar{x}) \oplus F(d) \otimes gH F(\bar{x})) \otimes gH F(d)\right)
\]

\[
= \lim_{\lambda \to 0} \frac{1}{\lambda} \left( [\nabla F(\bar{x}) + f(d) - f(\bar{x})] \otimes gH F(d) \right)
\]

\[
= 0.
\]

Hence, due to Definition 15 and Theorem 1, the linear IVF \(F\) is differentiable at \(\bar{x} \in X\). \(\square\)

Appendix 7: Proof of Theorem 3

Proof Let the function \(F\) be convex on \(X\). Then, for any \(x, y \in X\) and \(\lambda \in (0, 1]\), we get

\[
F(x + \lambda(y - x)) = F(\lambda y + (1 - \lambda)x) \leq \lambda \circ F(y) \oplus (1 - \lambda) \circ F(x),
\]

where \(\lambda' = 1 - \lambda\).

Hence,

\[
F(x + \lambda(y - x)) \otimes_{gH} F(x)
\]

\[
\leq (\lambda \circ F(y) \oplus \lambda' \circ F(x)) \otimes_{gH} F(x)
\]

\[
= [\lambda f(y) + \lambda' f(x), \lambda' f(y) + \lambda f(x)] \otimes_{gH} [f(x), f(y)]
\]

\[
= [\min\{\lambda f(y) + \lambda' f(x) - f(x), \lambda' f(y) + \lambda f(x) - f(x)\}, \max\{\lambda f(y) + \lambda' f(x) - f(x), \lambda' f(y) + \lambda f(x) - f(x)\}]
\]

\[
= \lambda \circ \left[ \min\{\lambda f(y) - \lambda f(x), \lambda' f(y) - \lambda' f(x)\}, \max\{\lambda f(y) - \lambda f(x), \lambda' f(y) - \lambda' f(x)\} \right],
\]

\[
= \lambda \circ \left( F(y) \otimes_{gH} F(x) \right)
\]

which implies

\[
\frac{1}{\lambda} \otimes (F(x + \lambda(y - x)) \otimes_{gH} F(x)) \leq F(y) \otimes_{gH} F(x).
\]

Since \(F\) is \(gH\)-differentiable at \(x \in X\), taking \(\lambda \to 0^+\), by Theorem 1, we have

\[
(y - x)^T \otimes \nabla F(x) \leq F(y) \otimes_{gH} F(x) \quad \text{for all } x, y \in X.
\]

\(\square\)
Appendix 8: Proof of Theorem 4

Proof Let the function $F$ be convex on $\mathcal{X}$. As $F$ is also $gH$-differentiable on $\mathcal{X}$, by Theorem 3, for all $x, y \in \mathcal{X}$ we obtain

$$(y-x)^T \circ \nabla F(x) \preceq F(y) \circ_{gH} F(x)$$

(24)

and

$$(x-y)^T \circ \nabla F(y) \preceq F(x) \circ_{gH} F(y).$$

(25)

For a given pair of points $x, y \in \mathcal{X}$, let us suppose

$$F(x) \circ_{gH} F(y) = [\tilde{a}, \tilde{b}], (x-y)^T \circ \nabla F(x) = [\bar{b}, \bar{b}],$$

and $(x-y)^T \circ \nabla F(y) = [\bar{c}, \bar{c}]$.

Thus, from (24) and (25), respectively, we have

$$[-\bar{b}, -\bar{b}] \preceq [-\bar{a}, -\bar{a}] \preceq [\tilde{a}, \tilde{a}]$$

$$\Rightarrow [\tilde{a}, \tilde{a}] \preceq [\bar{b}, \bar{b}]$$

and $[\tilde{a}, \tilde{a}] \preceq [-\bar{a}, -\bar{a}] \preceq [-\bar{c}, -\bar{c}]$

$$\Rightarrow a \leq b, \quad \bar{a} \leq \bar{b}, \quad -a \leq -c \quad \text{and} \quad -\bar{a} \leq -\bar{c}$$

$$\Rightarrow 0 \preceq b - c \quad \text{and} \quad 0 \preceq \bar{b} - \bar{c}$$

$$\Rightarrow 0 \preceq \min\{b - c, \bar{b} - \bar{c}\}, \max\{b - c, \bar{b} - \bar{c}\}$$

$$\Rightarrow 0 \preceq (x - y)^T \circ \nabla F(x) \circ_{gH} (x - y)^T \circ \nabla F(y) \quad \text{for all} \quad x, y \in \mathcal{X}.$$ 

Acknowledgements The authors put a sincere thanks to the anonymous reviewers and editors for their valuable comments to enhance the paper.

References

1. Ghosh, D., Ghosh, D., Bhuiya, S.K., Patra, L.K.: A saddle point characterization of efficient solutions for interval optimization problems. J. Appl. Math. Comput. 58(1–2), 193–217 (2018)
2. Moore, R.E.: Interval Analysis. Prentice-Hall, Englewood Cliffs (1966)
3. Moore, R.E.: Method and Applications of Interval Analysis. Society for Industrial and Applied Mathematics, Philadelphia (1987)
4. Wu, H.C.: The Karush-Kuhn-Tucker optimality conditions in an optimization problem with interval-valued objective function. Eur. J. Oper. Res. 176, 46–59 (2007)
5. Wu, H.C.: On interval-valued nonlinear programming problems. J. Math. Anal. Appl. 338(1), 299–316 (2008)
6. Wu, H.C.: Duality theory for optimization problems with interval-valued objective function. J. Optim. Theory Appl. 144(3), 615–628 (2009)
7. Hukuhara, M.: Intégration des applications measurable dont la valeur est un compact convex. Funkcialaj Ekvacioj 10, 205–223 (1967)
8. Markov, S.: Calculus for interval functions of a real variable. Computing 22(4), 325–337 (1979)
9. Stefanini, L.: A generalization of Hukuhara difference—soft methods for handling variability and imprecision. Adv. Soft Comput. 48, 203–210 (2009)
10. Stefanini, L.: A generalization of Hukuhara difference and division for interval and fuzzy arithmetic. Fuzzy Sets Syst. 161, 1564–1584 (2010)
11. Ishibuchi, H., Tanaka, H.: Multiobjective programming in optimization of the interval objective function. Eur. J. Oper. Res. 48(2), 219–225 (1990)
12. Chanas, S., Kuchta, D.: Multiobjective programming in optimization of interval objective functions—a generalized approach. Eur. J. Oper. Res. 94(3), 594–598 (1996)
13. Ghosh, D.: Newton method to obtain efficient solutions of the optimization problems with interval-valued objective functions. J. Appl. Math. Comput. 53, 709–731 (2017)
14. Ghosh, D.: A quasi-newton method with rank-two update to solve interval optimization problems. Int. J. Appl. Comput. Math. 3(3), 1719–1738 (2017)
15. Chen, S.H., Wu, J., Chen, Y.D.: Interval optimization for uncertain structures. Finite Elem. Anal. Des. 40, 1379–1398 (2004)
16. Chen, S.H., Wu, J.: Interval optimization of dynamic response for uncertain structures with natural frequency constraints. Eng. Struct. 26(2), 221–232 (2004)
17. Chinneck, J.W., Ramadan, K.: Linear programming with interval coefficients. J. Oper. Res. Soc. 51(2), 209–220 (2000)
18. Csendes, T.: New subinterval selection criteria for interval global optimization. J. Glob. Optim. 19, 307–327 (2001)
19. Limbourg, P., Aponte, D.E.S.: An optimization algorithm for imprecise multiobjective problem functions. In: IEEE Congress on Evolutionary Computation, Edinburgh, Scotland, UK, 2–5 September 2005, vol. 1, pp. 459–466 (2005)
20. Wolfe, M.A.: Interval mathematics, algebraic equations and optimization. J. Comput. Appl. Math. 124, 263–280 (2000)
21. Wu, X.Y., Huang, G.H., Liu, L., Li, J.B.: An interval nonlinear program for the planning of waste management systems with economies-of-scale effects—a case study for the region of Hamilton, Ontario, Canada. Eur. J. Oper. Res. 171(2), 349–372 (2006)
22. Ghosh, D., Deb Nath, A.K., Pedrycz, W.: A variable and a fixed ordering of intervals and their application in optimization with interval-valued functions. Int. J. Approx. Reason. 121, 187–205 (2020)
23. Chalco-Cano, Y., Rufian-Lizana, A., Roman-Flores, H., Jimenez-Gamero, M.D.: Calculus for interval-valued functions using generalized Hukuhara derivative and applications. Fuzzy Sets Syst. 219, 49–67 (2013)
24. Chalco-Cano, Y., Lodwick, W.A., Rufian-Lizana, A.: Optimality conditions of type KKT for optimization problem with interval-valued objective function via generalized derivative. Fuzzy Optim. Decis. Mak. 12, 305–322 (2013)
25. Ghosh, D., Singh, A., Shukla, K.K., Manchanda, K.: Extended Karush-Kuhn-Tucker condition for constrained interval optimization problems and its application in support vector machines. Inf. Sci. 504, 276–292 (2019)
26. Bede, B., Gil, S.G.: Generalizations of the differentiability of fuzzy-number-valued functions with applications to fuzzy differential equations. Fuzzy Sets Syst. 151(3), 581–599 (2005)
27. Stefanini, L., Bede, B.: Generalized Hukuhara differentiability of interval-valued functions and interval differential equations. Nonlinear Anal. Theory Methods Appl. 71(3–4), 1311–1328 (2009)
28. Bede, B., Stefanini, L.: Generalized differentiability of fuzzy-valued functions. Fuzzy Sets Syst. 230(1), 119–141 (2013)
29. Stefanini, L., Arana-Jiménez, M.: Karush-Kuhn-Tucker conditions for interval and fuzzy optimization in several variables
under total and directional generalized differentiability. Fuzzy Sets Syst. 362, 1–34 (2019)
30. Antczak, T.: Optimality conditions and duality results for non-smooth vector optimization problems with the multiple interval-valued objective function. Acta Math. Sci. 37B(4), 1133–1150 (2017)
31. Bao, Y., Zao, B., Bai, E.: Directional differentiability of interval-valued functions. J. Math. Comput. Sci. 16(4), 507–515 (2016)
32. Effati, S., Pakdaman, M.: Solving the interval-valued linear fractional programming problem. Am. J. Comput. Math. 2, 51–55 (2012)
33. Gong, D., Sun, J., Miao, Z.: A set-based genetic algorithm for interval many-objective optimization problems. IEEE Trans. Evol. Comput. 22(1), 47–60 (2018)
34. Jiang, C., Xiea, H.C., Zhang, Z.G., Hana, X.: A new interval optimization method considering tolerance design. Eng. Optim. 47(12), 1637–1650 (2015)
35. Bhurjee, A.K., Panda, G.: Efficient solution of interval optimization problem. Math. Methods Oper. Res. 76(3), 273–288 (2012)
36. Cousso, I., Dubois, D.: Statistical reasoning with set-valued information: ontic vs. epistemic views. Int. J. Approx. Reason. 55(7), 1502–1518 (2014)
37. Ghosh, D., Chauhan, R.S., Mesiar, R., Deb Nath, A.K.: Generalized Hukuhara Gâteaux and Fréchet derivatives of interval-valued functions and their application in optimization with interval-valued functions. Inf. Sci. 510, 317–340 (2020)
38. Chalco-Cano, Y., Román-Flores, H., Jiménez-Gamero, M.D.: Generalized derivative and σ-derivative for set-valued functions. Inf. Sci. 181(11), 2177–2188 (2011)
39. Ghosh, D. and Chakraborty, D.: An Introduction to Analytical Fuzzy Plane Geometry, Springer International Publishing (2019)
40. Bazaraa, M.S., Sherali, H.D., Shetty C.M.: Nonlinear Programming, 3rd edn. Wiley, New York (2006)
41. Neto, E.D.A.L., de Carvalho, F.D.A.: Constrained linear regression models for symbolic interval-valued variables. Comput. Stat. Data Anal. 54(2), 333–347 (2010)
42. Lodwick, W.A., Jamison, K.D.: A constraint fuzzy interval analysis approach to fuzzy optimization. Inf. Sci. 426, 38–49 (2018)

Debadas Ghosh is currently an Assistant Professor in the Department of Mathematical Sciences, Indian Institute of Technology (BHU) Varanasi, India. In 2014, he obtained Ph.D. degree from IIT Kharagpur in Mathematics. His research interest includes Optimization Theory. He is a recipient of Prof. J. C. Bose Memorial Gold Medal from IIT Kharagpur (2009). He has been awarded Outstanding Potential for Excellence in Research and Academics Award (2014) from his earlier employer, BITS-Pilani. So far, he has published twenty seven international journal papers and twelve conference papers. He has written a research monograph entitled An Introduction to Analytical Fuzzy Plane Geometry which is published by Springer in the book series of Studies in fuzziness and Soft Computing. He has edited two conference proceedings volumes on Mathematics and Computing which are published by Springer.

Amit Kumar Deb Nath received his M.Sc. degree in Mathematics from Indian Institute of Technology Kharagpur, Kharagpur, West Bengal, India. Presently, he is a research scholar of Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India. His research area is optimization with uncertainties.

Ram Surat Chauhan received his M.Sc. degree in Mathematics from Dr. Ram Manohar Lohia Avadh University, Ayodhya, Uttar Pradesh, India. Presently, he is a research scholar of Indian Institute of Technology (BHU), Varanasi, Uttar Pradesh, India. His research area is interval analysis and interval optimization.

Oscar Castillo holds the Doctor in Science degree (Doctor Habilitatus) in Computer Science from the Polish Academy of Sciences (with the Dissertation ‘Soft Computing and Fractional Theory for Intelligent Manufacturing’). He is a Professor of Computer Science in the Graduate Division, Tijuana Institute of Technology, Tijuana, Mexico. In addition, he is serving as a Research Director of Computer Science and head of the research group on Hybrid Fuzzy Intelligent Systems. Currently, he is the President of HAFSA (Hispanic American Fuzzy Systems Association) and Past President of IFSAS (International Fuzzy Systems Association). Prof. Castillo is also Chair of the Mexican Chapter of the Computational Intelligence Society (IEEE). He also belongs to the Technical Committee on Fuzzy Systems of IEEE and to the Task Force on ‘Extensions to Type-1 Fuzzy Systems.’ He is also a member of NAFIPS, IFSA, and IEEE. He belongs to the Mexican Research System (SNI Level 3). His research interests are in Type-2 Fuzzy Logic, Fuzzy Control, Neuro-Fuzzy, and Genetic-Fuzzy hybrid approaches. He has published over 300 journal papers, 10 authored books, 50 edited books, 300 papers in conference proceedings, and more than 300 chapters in edited books, in total more than 970 publications with h index of 79 according to Google Scholar. He has been Guest Editor of several successful Special Issues in the past, like in the following journals: Applied Soft Computing, Intelligent Systems, Information Sciences, Soft Computing, Non-Linear Studies, Fuzzy Sets and Systems, and JAMRIS and Engineering Letters. He is currently an
Associate Editor of the Information Sciences Journal, Journal of Engineering Applications on Artificial Intelligence, International Journal of Fuzzy Systems, Journal of Complex and Intelligent Systems, Granular Computing Journal and Intelligent Systems Journal (Wiley). He was an Associate Editor of Journal of Applied Soft Computing and IEEE Transactions on Fuzzy Systems. He has been elected IFSA Fellow in 2015 and MICAI Fellow in 2016. Finally, he recently received the Recognition as Highly Cited Researcher in 2017 and 2018 by Clarivate Analytics and Web of Science.