Chapter 1

Introduction

Optimization, involves searching for the best solution to a given problem from the set of all possible feasible options. Mathematically, optimization refers to the minimization or maximization of an objective function or a group of objective functions over a domain of feasible solutions, subject to specified constraints. Formally written as

$$\min f(x) \;\; \text{subject to } x \in C,$$

an optimization problem consists of three fundamental components: an objective function $f(x)$, decision variables x, and constraint set C. The *objective function* quantifies the primary objective of the optimization task, such as minimizing a cost function, maximizing profit, or maximizing efficiency. The *decision variables* represent the choices available to the decision-maker. *Constraints* define the permissible values or range of the decision variables, establishing a feasible region within which the optimal solution must be found.

In practical applications, mathematical optimization finds widespread use across diverse fields that involve any form of decision-making or requires improvement in efficiency. For instance, in engineering, optimization plays a pivotal role in the design of systems, structures, and processes. In finance, optimization forms the backbone of algorithmic trading and portfolio management aimed at maximizing returns and minimizing risk. Operations research utilizes optimization to address challenges such as resource allocation, logistics planning, and production scheduling. In machine learning, optimization is fundamental to refining model parameters for optimal performance in machine tasks. Telecommunications leverages optimization for routing and network design to maximize data transmission throughput and minimize costs.

In summary, optimization significantly enhances decision-making processes across diverse industries and applications. Recent years have witnessed rapid advancements in optimization techniques, driven primarily by progress in mathematics, computational capabilities, and algorithmic innovation. For instance, there is ongoing effort to develop *global optimization* algorithms capable of identifying global optima rather than the local ones. The integration of optimization techniques with advancements in machine learning reflects another significant trend. Other significant area of focus pertains to tackling challenges characterized by inherent uncertainty, such as those in finance and logistics. *Interval-valued optimization* remains crucial in these domains for robustly handling variations or uncertainties in input variables. Additionally, the fields of *multi-objective optimization* and *set optimization* have gained increasing attention for resolving real-world challenges involving multiple conflicting objectives that necessitate simultaneous optimization, e.g., portfolio management in finance.

Within the vast field of mathematical optimization, various theories and methodologies have been developed to address traditional optimization problems typically characterized by real scalar-valued objective functions. However, in practical applications, scenarios often arise where the objective function is uncertain, primarily due to measurement errors in data collection or inherent uncertainties in modeling. Such challenges cannot be adequately modeled using conventional optimization tools. Consider, for instance, the problem of portfolio management, where uncertainties in asset returns arise due to market fluctuations over time. Suppose an investor allocates funds to two assets, denoted as Asset 1 and Asset 2, with corresponding investments represented by x_1 and x_2. Let r_1 and r_2 denote the returns on these assets. The investor's objective is to maximize the expected total return from the portfolio, given by $r_1x_1 + r_2x_2$, subject to the budget constraint $x_1 + x_2 \leq D$, where D represents the total available investment funds.

If the returns r_1 and r_2 were assumed to be fixed without any variability, the problem could be formulated and solved as a scalar optimization problem given by

$$\begin{aligned} &\max \quad r_1x_1 + r_2x_2 \\ &\text{subject to} \;\; x_1 + x_2 \leq D. \end{aligned} \tag{1.1}$$

However, this scenario rarely holds true in practice because the returns on these assets typically vary over time in a real market, driven by market fluctuations. In mathematics, several methodologies are available to represent uncertain quantities such as random variables, fuzzy numbers, intervals, and sets. Depending on the nature of r_1 and r_2, the problem (1.1) can be tackled using various optimization frameworks:

- r_1 and r_2 are random variables: stochastic optimization problem [16, 101, 179].
- r_1 and r_2 are fuzzy: fuzzy optimization problem [52].

- r_1 and r_2 are intervals: interval optimization problem [15].

- r_1 and r_2 are sets: set-valued optimization problem [138].

Among these approaches, *interval-valued optimization* and *set-valued optimization* have attracted considerable attention due to their efficacy in handling uncertainties and their broad applicability across different domains. Each method offers distinct advantages and finds numerous practical applications.

Interval optimization or interval-valued optimization focuses on problems where intervals appear in the objective functions, decision variables, or constraints. This approach is particularly suitable for scenarios involving inherent randomness or variability in input data or where robust decision-making under uncertainty is required. In such contexts, traditional scalar-valued optimization methods may not adequately address or model these uncertainties. If the portfolio management problem mentioned earlier in (1.1) is formulated as an interval optimization problem, it can be expressed as

$$\begin{aligned} &\max \quad [\underline{r}_1, \overline{r}_1] \odot x_1 + [\underline{r}_2, \overline{r}_2] \odot x_2 \\ &\text{subject to} \;\; x_1 + x_2 \leq D. \end{aligned}$$

In this context, the coefficients are intervals $[\underline{r}_1, \overline{r}_1]$, and $[\underline{r}_2, \overline{r}_2] \in I(\mathbb{R})$ are interval instead of scalar values. Interval optimization thus generalizes scalar optimization to situations where uncertainty in coefficients needs to be explicitly considered. Overall, interval-valued optimization provides a robust machinery for addressing uncertainty, variability, and randomness within mathematical optimization problems, facilitating more reliable and resilient decision-making in real-world applications.

Set optimization or set-valued optimization deals with problems where the objective function and/or constraint functions are set-valued. The sets can consist of various structures, such as intervals, polygons, and convex sets, among others. If the portfolio management problem mentioned earlier in (1.1) is formulated as a set optimization problem, it can be represented as

$$\begin{aligned} &\max \quad \left\{ r_1 x_1 + r_2 x_2 \ : \ (r_1, r_2) \in \mathbb{R}^2_+ \right\} \\ &\text{subject to} \quad x_1 + x_2 \leq D, \end{aligned}$$

Here, the set of objective functions captures all possible return values for each asset. The framework of set optimization offers a more powerful and generalized approach to handling uncertainty and variability. In the above example, instead of optimizing over a single interval, set-valued optimization allows optimization over a set of solutions.

This approach ensures that the final solution is robust and performs well under various conditions.

In this thesis, we begin by diving into the topic of interval optimization, exploring various tools and mathematical constructs used therein. Subsequently, we progress towards an advanced and generalized approach for managing variability through the framework of set optimization.

1.1 Interval-valued optimization

Returning to the portfolio management problem (1.1), if the fluctuations in market returns r_1 and r_2, the problem transforms into an interval optimization problem. This construction tends to be relatively easier to solve compared to its fuzzy or stochastic counterparts. In fuzzy optimization, fuzzy sets are characterized by families of their alpha-cuts. Each alpha-cut can be associated with a corresponding interval, allowing the fuzzy optimization problem to be reformulated as an interval optimization problem. Similarly, instead of treating r_1 and r_2 as fuzzy numbers, their confidence fuzzy intervals can be considered to convert the problem into an interval optimization format [22].

In stochastic optimization, coefficients r_1 and r_2 are considered as random variables with assumed known distributions. Researchers often resort to subjective choices regarding these distributions. Gaussian distribution is frequently chosen due to its extensive study and the abundance of mathematical tools available for its manipulation. However, these distributional assumptions may not always accurately reflect real-world conditions. Instead, when discussing intervals, there is a general agreement that assuming compact intervals to bound observed data is simpler than making assumptions about the entire distribution. Therefore, in such contexts, interval-valued optimization emerges as a potentially superior approach for addressing uncertainty.

At the core of interval optimization lies the theory of interval analysis, which traces its origins to the pioneering work of Moore [142]. Interval analysis fundamentally treats intervals as numerical entities and facilitates arithmetic operations on them. This approach is particularly valuable in computational setups prone to rounding errors, offering a systematic means to bound and manage the accumulation of these errors. In real-world applications, such errors can lead to significant costs or even catastrophic outcomes. For example, standard interval analysis could have prevented the Patriot missile incident in 1991, which resulted in 28 fatalities due to error accumulation. Similarly, the failure of the Ariane 5 rocket launch in 1996 was attributed to overflow errors. Moreover, as previously noted, many real-world problems involve inherent uncertainty or variability in their parameters. Such variations cannot always be accurately represented by single real numbers, making interval values a natural choice for capturing uncertainty within defined bounds. The primary objective of interval analysis, therefore, is to develop interval-valued algorithms capable of bounding solutions in uncertain scenarios, providing upper and lower bounds that enhance robustness and reliability in problem-solving.

We begin by defining interval-valued functions, thereby diving into the core mathematics of interval-valued optimization.

1.1.1 Interval-valued function

An interval-valued function is a function of one or more variables whose range consists of closed and bounded intervals of real numbers. This set of intervals is denoted by $I(\mathbb{R})$. Then, for the Euclidean space $\mathbb{R}^n$, a set of intervals vectors $I(\mathbb{R})^k$, and $\mathcal{X}$ being a nonempty subset of $\mathbb{R}^n$, the function $\mathbf{F}_{\mathbf{C}_v^k} : \mathcal{X} \to I(\mathbb{R})$ is called an interval-valued function (IVF) that depends on k intervals in the interval vector $\mathbf{C}_v^k = (\mathbf{C}_1,\ \mathbf{C}_2, \ldots, \mathbf{C}_k)^\top$, where $\mathbf{C}_j = [\underline{c_j},\ \overline{c_j}] \in I(\mathbb{R})$ for $j = 1, 2, \ldots, k$.

In parametric form, the vector $\mathbf{C}_v^k$ is observed by the following set

$$\Big\{c(t) \ \Big|\ c(t) = (c_1(t_1), c_2(t_2), \ldots, c_k(t_k))^T, c_j(t_j) = \underline{c_j} + t_j(\overline{c_j} - \underline{c_j}),$$
$$t = (t_1,\ t_2, \ldots, t_k)^\top,\ 0 \leq t_j \leq 1,\ j = 1, 2, \ldots, k\Big\}.$$

Hence, the function $\mathbf{F}_{\mathbf{C}_v^k}$ can be expressed as a collection of real-valued functions $f_{c(t)}$'s, i.e., for all $x \in \mathcal{X}$,

$$\mathbf{F}_{\mathbf{C}_v^k}(x) = \left\{ f_{c(t)}(x) | f_{c(t)} : \mathcal{X} \to \mathbb{R},\ c(t) \in \mathbf{C}_v^k,\ t \in [0,\ 1]^k \right\}.$$

The function $\mathbf{F}_{\mathbf{C}_v^k}$ can also be given in a different way. Consider

$$\underline{f}(x) = \min_{t \in [0,1]^k} f_{c(t)}(x) \text{ and } \overline{f}(x) = \max_{t \in [0,1]^k} f_{c(t)}(x).$$

Then, for every argument point $x \in \mathcal{X}$, $\mathbf{F}_{\mathbf{C}_v^k}$ can be given by

$$\mathbf{F}_{\mathbf{C}_v^k}(x) = \left[\underline{f}(x),\ \overline{f}(x) \right].$$

1.1.2 Formulation of interval optimization problem

Mathematically, an interval-valued optimization problem is presented as

$$\min_{x \in \mathcal{X}} \mathbf{F}(x),$$

where $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$. Depending on the nature of the interval-valued objective function $\mathbf{F}$ and whether the decision variable x is real-valued or interval-valued, we can formulate various types of optimization problems. In this thesis, our primary focus centers on the following four variants of interval optimization:

Case 1. Decision variable x is vector of real numbers and unconstrained in $\mathbb{R}^n$:

$$\min_{x \in \mathbb{R}^n} \mathbf{F}(x),$$

where $\mathbf{F} : \mathbb{R}^n \to I(\mathbb{R})$ is a non-smooth, non-convex, gH-Lipshitz interval-valued functions. This is studied in Chapter 5 of this thesis.

Case 2. Decision variable x is vector of real numbers and constrained to a non-empty *open* set $\mathcal{X} \subseteq \mathbb{R}^n$:

$$\left.\begin{array}{ll} \min & \mathbf{F}(x) \\ \text{subject to} & \mathbf{G}_j(x) \preceq \mathbf{0}, j = 1, 2, \ldots, m \\ & x \in \mathcal{X}. \end{array}\right\}, \tag{1.2}$$

where $\mathbf{F} : \mathbb{R}^n \to I(\mathbb{R})$ and $\mathbf{G}_j : \mathbb{R}^n \to I(\mathbb{R})$ for $j = 1, 2, \ldots, m$ are IVFs. This is studied in Chapter 4 of this thesis.

Case 3. Decision variable x is vector of real numbers and constrained to a non-empty set $\mathcal{X} \subseteq \mathbb{R}^n$, and objective is difference of two IVFs $\mathbf{F}_1, \mathbf{F}_2$:

$$\min_{x \in \mathcal{X}} \{\boldsymbol{F}_2(x) \ominus_{gH} \boldsymbol{F}_1(x)\}, \tag{1.3}$$

where $\boldsymbol{F}_1, \boldsymbol{F}_2 : \mathcal{X} \to I(\mathbb{R})$ are two IVFs. This is studied in Chapter 3 of this thesis.

Case 4. Decision variable $\widehat{\mathbf{X}}$ is a vector of intervals and constrained to a a non-empty subset $\boldsymbol{\mathcal{X}}$ of $I(\mathbb{R})^n$:

$$\min_{\widehat{\mathbf{X}} \in \boldsymbol{\mathcal{X}}} \mathbf{F}(\widehat{\mathbf{X}}), \tag{1.4}$$

where $\mathbf{F} : \boldsymbol{\mathcal{X}} \to I(\mathbb{R})$ be an IVF. This is studied in Chapter 1 of this thesis.

Decision space: The task of a decision-maker is to select a decision from a set of possible candidates, denoted as $\mathcal{X}$. A candidate can either be a vector of real numbers $x = (x_1, x_2, \ldots, x_n)^\top \in \mathcal{X}$ or a vector of intervals $\widehat{\mathbf{X}} = (\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n)^\top \in \mathcal{X}$, where $\mathbf{X}_i = [\underline{x}_i, \overline{x}_i]$ for all $i = 1, 2, \ldots, n$. The optimal candidate vector is referred to as a decision vector, with its components termed decision variables. The collection of decision vectors forms the decision space. Alternatively, the set $\mathcal{X}$ is also known as the decision feasible region, and a point x in $\mathcal{X}$ is referred to as a decision feasible point.

Objective space: For each point x in the decision feasible region $\mathcal{X}$, $\mathbf{F}(x)$ yields an interval in $I(\mathbb{R})$. $I(\mathbb{R})$ encompasses the range of all possible interval values of the objective function. Under the interval-valued function $\mathbf{F}$, the image of the feasible space $\mathcal{X}$ is called the feasible objective space.

Solution concept based on order relation: An interval-valued objective function can be seen as comprising infinitely many real-valued objective functions, similar to multi-objective optimization. Consequently, the optimal solution of an IOP, unlike in single-objective optimization, is not unique. This non-uniqueness arises because, in single-objective optimization, the feasible objective space forms a totally ordered subset of $\mathbb{R}$. However, in IOP, due to $I(\mathbb{R})$ not being totally ordered, the infinite-dimensional feasible objective space is not a totally ordered subset of $I(\mathbb{R})$. As a result, all IOP solutions cannot be totally ordered; instead, they are partially ordered. Moreover, due to conflicting real objectives in IOP, situations may arise where no single solution is superior to others, leading to multiple solutions. The concept of a single optimal solution does not always apply when dealing with infinitely many real objectives. This leads to the introduction of non-dominated solutions.

Non-dominated solutions: In IOP with infinitely many real objectives, a non-dominated solution refers to a feasible solution in the objective space where improving one criterion can only be achieved by worsening at least one other criterion. Therefore, within the set of non-dominated solutions, there is no solution that is optimal across all criteria simultaneously. Each non-dominated solution is considered equally acceptable for the IOP.

1.1.3 Preliminaries

The following fundamental definitions and properties of intervals are utilized extensively throughout this thesis.

1.1.3.1 Interval Arithmetic

Throughout the thesis, bold letters $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$ denote the elements of $I(\mathbb{R})$. An element $\mathbf{A}$ of $I(\mathbb{R})$ is given by the corresponding small letter: $\mathbf{A} = [\underline{a}, \overline{a}]$. In this section, we discuss Moore's interval arithmetic [144] followed by the concepts of gH-difference of two intervals and ordering of intervals [91].

Let $\mathbf{A} = [\underline{a}, \overline{a}]$ and $\mathbf{B} = [\underline{b}, \overline{b}]$. The *addition* of $\mathbf{A}$ and $\mathbf{B}$, written as $\mathbf{A} \oplus \mathbf{B}$, is given by

$$\mathbf{A} \oplus \mathbf{B} = \left[\ \underline{a} + \underline{b},\ \overline{a} + \overline{b}\ \right].$$

The *subtraction* of $\mathbf{B}$ from $\mathbf{A}$, written as $\mathbf{A} \ominus \mathbf{B}$, is given by

$$\mathbf{A} \ominus \mathbf{B} = \left[\ \underline{a} - \overline{b},\ \overline{a} - \underline{b}\ \right].$$

The *multiplication* of $\mathbf{A}$ and $\mathbf{B}$, written as $\mathbf{A} \odot \mathbf{B}$, is given by

$$\mathbf{A} \odot \mathbf{B} = \left[\min\left\{\underline{a}\ \underline{b},\ \underline{a}\overline{b},\ \overline{a}\underline{b},\ \overline{a}\overline{b}\right\},\ \max\left\{\underline{a}\ \underline{b},\ \underline{a}\overline{b},\ \overline{a}\underline{b},\ \overline{a}\overline{b}\right\}\right].$$

The *multiplication* by a real number λ to $\mathbf{A}$, written as $\lambda \odot \mathbf{A}$ or $\mathbf{A} \odot \lambda$, is given by

$$\lambda \odot \mathbf{A} = \mathbf{A} \odot \lambda = \begin{cases} [\lambda\underline{a},\ \lambda\overline{a}], & \text{if } \lambda \geq 0 \\ [\lambda\overline{a},\ \lambda\underline{a}], & \text{if } \lambda < 0. \end{cases}$$

Note that the definition of $\lambda \odot \mathbf{A}$ arises from the fact that $\lambda = [\lambda, \lambda]$ and the definition of the multiplication $\mathbf{A} \odot \mathbf{B}$.
Consider $0 \notin \mathbf{B}$. The *division* of $\mathbf{A}$ by $\mathbf{B}$, written as $\mathbf{A} \oslash \mathbf{B}$, is given by

$$\mathbf{A} \oslash \mathbf{B} = \left[\min\left\{\underline{a}/\underline{b},\ \underline{a}/\overline{b},\ \overline{a}/\underline{b},\ \overline{a}/\overline{b}\right\},\ \max\left\{\underline{a}/\underline{b},\ \underline{a}/\overline{b},\ \overline{a}/\underline{b},\ \overline{a}/\overline{b}\right\}\right].$$

Definition 1.1 (gH-difference of intervals [165]). *The gH-difference of $\boldsymbol{R}$ and $\boldsymbol{S}$, written as $\boldsymbol{R} \ominus_{gH} \boldsymbol{S}$, is given by the interval $\boldsymbol{Y}$ such that*

$$\boldsymbol{R} = \boldsymbol{S} \oplus \boldsymbol{Y} \quad \textit{or} \quad \boldsymbol{S} = \boldsymbol{R} \ominus \boldsymbol{Y}.$$

Note that for $\boldsymbol{R} = [\underline{r},\ \overline{r}]$ and $\boldsymbol{S} = [\underline{s},\ \overline{s}]$,

$$\boldsymbol{R} \ominus_{gH} \boldsymbol{S} = [\min\{\underline{r} - \underline{s}, \overline{r} - \overline{s}\},\ \max\{\underline{r} - \underline{s}, \overline{r} - \overline{s}\}] \ \textit{and}\ \boldsymbol{R} \ominus_{gH} \boldsymbol{R} = \boldsymbol{0}.$$

If $\boldsymbol{\mathcal{Y}}$ is a non-empty subset of $I(\mathbb{R})^n$ and $\widehat{\mathbf{A}} \in I(\mathbb{R})^n$, then we write $\boldsymbol{\mathcal{Y}} \ominus_{gH} \widehat{\mathbf{A}} = \{\widehat{\mathbf{Y}} \ominus_{gH} \widehat{\mathbf{A}} : \widehat{\mathbf{Y}} \in \boldsymbol{\mathcal{Y}}\}$.

Definition 1.2 [66] *For $\widehat{\boldsymbol{I}} = (\boldsymbol{I}_1, \boldsymbol{I}_2, \ldots, \boldsymbol{I}_n)^\top$ and $\widehat{\boldsymbol{J}} = (\boldsymbol{J}_1, \boldsymbol{J}_2, \ldots, \boldsymbol{J}_n)^\top$ in $I(\mathbb{R})^n$, the* *algebraic operation* $\widehat{\boldsymbol{I}} \star \widehat{\boldsymbol{J}}$ *is given as*

$$\widehat{\boldsymbol{I}} \star \widehat{\boldsymbol{J}} = (\boldsymbol{I}_1 \star \boldsymbol{J}_1, \boldsymbol{I}_2 \star \boldsymbol{J}_2, \ldots, \boldsymbol{I}_n \star \boldsymbol{J}_n)^\top,$$

where $\star \in \{\oplus, \ominus, \ominus_{gH}\}$.

Definition 1.3 (Dominance relations on intervals [185]). *Consider* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *to be two intervals in* $I(\mathbb{R})$.

(i) $\boldsymbol{D}$ *is said to be dominated by* $\boldsymbol{C}$ *if* $\underline{c} \leq \underline{d}$ *and* $\overline{c} \leq \overline{d}$, *and we write* $\boldsymbol{C} \preceq \boldsymbol{D}$;

(ii) $\boldsymbol{D}$ *is called strictly dominated by* $\boldsymbol{C}$ *if* $\boldsymbol{C} \prec \boldsymbol{D}$. *Equivalently,* $\boldsymbol{C} \prec \boldsymbol{D}$ *if and only if any of the these cases hold:*

- *Case 1.* $\underline{a} < \underline{b}$ *and* $\overline{a} \leq \overline{b}$,
- *Case 2.* $\underline{a} \leq \underline{b}$ *and* $\overline{a} < \overline{b}$,
- *Case 3.* $\underline{a} < \underline{b}$ *and* $\overline{a} < \overline{b}$;

(iii) If neither $\boldsymbol{C} \preceq \boldsymbol{D}$ *nor* $\boldsymbol{D} \preceq \boldsymbol{C}$, *then none of* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *dominates the other, or* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *are not comparable. Equivalently,* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *are not comparable if either '*$\underline{c} < \underline{d}$*' and '*$\overline{c} < \overline{d}$*' or '*$\underline{c} > \underline{d}$*' and '*$\overline{c} < \overline{d}$*';*

(iv) $\boldsymbol{C}$ *is said to be not dominated by* $\boldsymbol{D}$ *if either* $\boldsymbol{C} \preceq \boldsymbol{D}$ *or* $\boldsymbol{C}$ *and* $\boldsymbol{D}$ *are not comparable, and we write* $\boldsymbol{C} \nprec \boldsymbol{D}$.

Definition 1.4 (Norm on $I(\mathbb{R})$ [143]). *Consider* $\boldsymbol{Y} = [\underline{y}, \overline{y}]$ *to be in* $I(\mathbb{R})$. *The function* $\|\cdot\|_{I(\mathbb{R})} : I(\mathbb{R}) \to \mathbb{R}_+$ *as given below is a norm in* $I(\mathbb{R})$*:*

$$\| \boldsymbol{Y}\|_{I(\mathbb{R})} = \max\{|\underline{y}|, |\overline{y}|\}.$$

Definition 1.5 (Norm on $I(\mathbb{R})^n$ [143]). *Consider* $\widehat{\boldsymbol{Y}} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n)^\top$ *to be in* $I(\mathbb{R})^n$. *The function* $\|\cdot\|_{I(\mathbb{R})^n} : I(\mathbb{R})^n \to \mathbb{R}_+$ *given as*

$$\| \widehat{\boldsymbol{Y}}\|_{I(\mathbb{R})^n} = \sum_{j=1}^{n} \| \boldsymbol{Y}_j \|_{I(\mathbb{R})}$$

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

1.1.3.2 Basic Properties of Intervals

The following fundamental properties of intervals, including the dominance relation of intervals, the norm of an interval, and the gH-difference of two intervals, are employed throughout this thesis.

Lemma 1.1 *Consider* $\boldsymbol{X}$, $\boldsymbol{Y}$, *and* $\boldsymbol{Z}$ *to be elements of* $I(\mathbb{R})$. *Then,*

(i) $\boldsymbol{X} \ominus_{gH} \boldsymbol{Z} = (\boldsymbol{X} \oplus \boldsymbol{Y}) \ominus_{gH} (\boldsymbol{Y} \oplus \boldsymbol{Z})$ *and*

(ii) $\boldsymbol{X} = \boldsymbol{Y} \oplus \boldsymbol{Z} \implies \boldsymbol{X} \ominus_{gH} \boldsymbol{Y} = \boldsymbol{Z}$.

Proof: See Appendix 9.1. □

Lemma 1.2 *If* $\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{Z} \in I(\mathbb{R})$ *and* $\epsilon \geq 0$, *we have*

$$\epsilon \preceq (\boldsymbol{W} \ominus_{gH} \boldsymbol{Y}) \ominus_{gH} \boldsymbol{Z} \implies \boldsymbol{Z} \oplus \epsilon \preceq \boldsymbol{W} \ominus_{gH} \boldsymbol{Y}.$$

Proof: See Appendix 10.1. □

Lemma 1.3 *If* $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}, \boldsymbol{W} \in I(\mathbb{R})$, *we have*

$$(\boldsymbol{X} \oplus \boldsymbol{Y}) \ominus_{gH} (\boldsymbol{Z} \oplus \boldsymbol{W}) \subseteq (\boldsymbol{X} \ominus_{gH} \boldsymbol{Z}) \oplus (\boldsymbol{Y} \ominus_{gH} \boldsymbol{W}).$$

Proof: See Appendix 10.2. □

Lemma 1.4 *If* $\boldsymbol{W}, \boldsymbol{Y}, \boldsymbol{Z} \in I(\mathbb{R})$, *we have*

$$\boldsymbol{0} \ominus_{gH} \{((-1 \odot \boldsymbol{W}) \ominus_{gH} (-1 \odot \boldsymbol{Y})) \ominus_{gH} (-1 \odot \boldsymbol{Z})\} = ((\boldsymbol{W} \ominus_{gH} \boldsymbol{Y}) \ominus_{gH} \boldsymbol{Z}).$$

Proof: See Appendix 10.3. □

Lemma 1.5 *For all* $\boldsymbol{X}, \boldsymbol{Y}$, *and* $\boldsymbol{Z}$ *of* $I(\mathbb{R})$,

(i) if $\boldsymbol{0} \preceq \boldsymbol{X} \ominus_{gH} \boldsymbol{Y}$, *then* $\boldsymbol{0} \ominus_{gH} \boldsymbol{Z} \preceq (\boldsymbol{X} \ominus_{gH} \boldsymbol{Y}) \ominus_{gH} \boldsymbol{Z}$,

(ii) if $\boldsymbol{Z} \preceq \boldsymbol{X} \ominus_{gH} \boldsymbol{Y}$, *then* $\boldsymbol{Z} \ominus_{gH} \boldsymbol{W} \preceq (\boldsymbol{X} \ominus_{gH} \boldsymbol{Y}) \ominus_{gH} \boldsymbol{W}$ *for all* $\boldsymbol{W} \in I(\mathbb{R})$,

(iii) if $\boldsymbol{X} \ominus_{gH} \boldsymbol{Y} \preceq [L, L]$, *then* $[-L, -L] \preceq \boldsymbol{Y} \ominus_{gH} \boldsymbol{X}$, *where* $L \in \mathbb{R}$,

(iv) if $[-\gamma, -\gamma] \preceq \boldsymbol{X} \ominus_{gH} \boldsymbol{Y}$, *then* $\boldsymbol{Y} \ominus_{gH} [\gamma, \gamma] \preceq \boldsymbol{X}$, *where* $\gamma \in \mathbb{R}$, *and*

(v) if $\boldsymbol{Z} \preceq \boldsymbol{X} \oplus \boldsymbol{Y}$, *then* $\boldsymbol{Z} \ominus_{gH} \boldsymbol{Y} \preceq \boldsymbol{X}$.

Proof: See Appendix 10.4. □

Lemma 1.6 *If* $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{C} \in I(\mathbb{R})$, *then*

$$\inf\{\boldsymbol{Q}, \boldsymbol{R} \oplus \boldsymbol{C}\} \ominus_{gH} \boldsymbol{C} \subseteq \inf\{\boldsymbol{Q} \ominus_{gH} \boldsymbol{C}, \boldsymbol{R}\}.$$

Proof: See Appendix 11.1. $\square$

Lemma 1.7 *If $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{C} \in I(\mathbb{R})$, then*

$$\inf\{\boldsymbol{Q} \ominus_{gH} \boldsymbol{C}, \boldsymbol{R}\} \ominus_{gH} \boldsymbol{R} \subseteq \inf\{(\boldsymbol{Q} \ominus_{gH} \boldsymbol{C}) \ominus_{gH} \boldsymbol{R}, \boldsymbol{0}\}.$$

Proof: See Appendix 11.2 $\square$

Lemma 1.8 *If $\boldsymbol{Q}, \boldsymbol{R}, \boldsymbol{S} \in I(\mathbb{R})$ and $\epsilon \geq 0$, we have*

$$-\epsilon \preceq (\boldsymbol{Q} \ominus_{gH} \boldsymbol{R}) \ominus_{gH} \boldsymbol{S} \Longrightarrow \boldsymbol{S} \oplus (\boldsymbol{R} \ominus_{gH} \boldsymbol{Q}) \preceq \epsilon.$$

Proof: See Appendix 11.3. $\square$

Lemma 1.9 *If $\boldsymbol{\Upsilon}_1, \boldsymbol{\Upsilon}_2 : \mathcal{Z} \to I(\mathbb{R})$ are two IVFs, then for any $\bar{z} \in \mathcal{Z}$,*

(i) $\liminf_{z \to \bar{z}} \{-1 \odot \boldsymbol{\Upsilon}_1(z)\} = -1 \odot \limsup_{z \to \bar{z}} \boldsymbol{\Upsilon}_1(z)$, *and*

(ii) if $\lim_{z \to \bar{z}} \boldsymbol{\Upsilon}_2(z)$ *exists, then*

$$\liminf_{z \to \bar{z}} \{\boldsymbol{\Upsilon}_1(z) \ominus_{gH} \boldsymbol{\Upsilon}_2(z)\} = \liminf_{z \to \bar{z}} \boldsymbol{\Upsilon}_1(z) \ominus_{gH} \lim_{z \to \bar{z}} \boldsymbol{\Upsilon}_2(z).$$

Proof: See Appendix 11.4. $\square$

1.1.3.3 Sequence of Intervals

Definition 1.6 (Sequence in $I(\mathbb{R})^n$ [66]). *An IVF $\boldsymbol{T} : \mathbb{N} \to I(\mathbb{R})^n$ is said to a sequence in $I(\mathbb{R})^n$. The image of nth element $\boldsymbol{T}(n)$ is called the nth element of the sequence $\boldsymbol{T}$. The sequence $\boldsymbol{T}$ is denoted by $\{\boldsymbol{T}(n)\}$.*

Definition 1.7 (Convergence of a sequence in $I(\mathbb{R})^n$ [66]). *A sequence $\{\widehat{\boldsymbol{G}}_k\}$ in $I(\mathbb{R})^n$ is called convergent if for each $\epsilon > 0$ there exists a $\widehat{\boldsymbol{G}} \in I(\mathbb{R})^n$ satisfying*

$$\|\widehat{\boldsymbol{G}}_k \ominus_{gH} \widehat{\boldsymbol{G}}\|_{I(\mathbb{R})^n} < \epsilon \text{ for all } k \geq m.$$

Then, we write $\lim_{k \to \infty} \widehat{\boldsymbol{G}}_k = \widehat{\boldsymbol{G}}$.

Definition 1.8 (Closed set in $I(\mathbb{R})^n$). *A non-empty subset $\boldsymbol{\mathcal{Y}} \subseteq I(\mathbb{R})^n$ is said to be* closed *if for every convergent sequence $\{\widehat{\boldsymbol{G}}_k\}$, in $\boldsymbol{\mathcal{Y}}$, converging to $\widehat{\boldsymbol{G}}$, $\widehat{\boldsymbol{G}}$ lies in $\boldsymbol{\mathcal{Y}}$.*

Definition 1.9 (Closure of a set in $I(\mathbb{R})^n$). *Consider $\boldsymbol{\mathcal{Y}} \subseteq I(\mathbb{R})^n$. Then, the intersection of all closed sets containing $\boldsymbol{\mathcal{Y}}$ is called the closure of $\boldsymbol{\mathcal{Y}}$, denoted by cl($\boldsymbol{\mathcal{Y}}$).*

Definition 1.10 (Sequentially continuous IVF). *Consider a non-empty subset* $\boldsymbol{\mathcal{Y}} \subseteq I(\mathbb{R})^n$. *Then,* $\boldsymbol{T}$ *is called* sequentially continuous IVF *at* $\widehat{\boldsymbol{C}} \in \boldsymbol{\mathcal{Y}}$ *if for every sequences* $\{\widehat{\boldsymbol{Y}}_k\}$ *in* $\boldsymbol{S}$ *with* $\lim\limits_{n\to\infty} \widehat{\boldsymbol{Y}}_k = \boldsymbol{C}$, *we have* $\lim_{n\to\infty} \boldsymbol{T}(\widehat{\boldsymbol{Y}}_k) = \boldsymbol{T}(\widehat{\boldsymbol{C}})$.

Remark 1.1.1 (See [66]). *We note that if a sequence* $\{\widehat{\boldsymbol{G}}_k\}$ *in* $I(\mathbb{R})^n$ *converges to* $\widehat{\boldsymbol{G}} \in I(\mathbb{R})^n$, *where* $\widehat{\boldsymbol{G}}_k = (\boldsymbol{G}_{k1}, \boldsymbol{G}_{k2}, \ldots, \boldsymbol{G}_{kn})^\top$ *and* $\widehat{\boldsymbol{G}} = (\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_n)^\top$, *then as per Definition 1.4 and Definition 1.5, the sequence* $\boldsymbol{G}_{kj}$ *in* $I(\mathbb{R})$ *converges to* $\boldsymbol{G}_j \in I(\mathbb{R})$ *for each* $j = 1, 2, \ldots, n$. *Further, as per Definition 1.4, the sequences* $\{\underline{g}_{kj}\}$ *and* $\{\overline{g}_{kj}\}$ *in* $\mathbb{R}$ *converge to* $\{\underline{g}_j\}$ *and* $\{\overline{g}_j\}$, *respectively, for all* j.

1.1.3.4 Some Basic Definitions and Properties of Interval-Valued Functions

Definition 1.11 (Convex IVF [69]). *Consider a convex subset* $\mathcal{Z}$ *of* $\mathbb{R}^n$. *Then, an IVF* $\boldsymbol{\Upsilon}$ *is called convex if for all* $z_1, z_2 \in \mathcal{Z}$, $\alpha \in [0, 1]$, *it holds that*

$$\boldsymbol{\Upsilon}(\alpha \odot z_1 \oplus (1-\alpha) \odot z_2) \preceq \alpha \odot \boldsymbol{\Upsilon}(z_1) \oplus (1-\alpha) \odot \boldsymbol{\Upsilon}(z_2).$$

Lemma 1.10 *If* $\boldsymbol{\Phi}$ *is a convex IVF on a convex set* $\mathcal{Y} \subseteq \mathbb{R}^n$, *then* $\underline{\phi}$ *and* $\overline{\phi}$ *are convex on* $\mathcal{Y}$ *and vice-versa.*

Definition 1.12 (Concave IVF). *If* $\mathcal{Y}$ *is convex, then an IVF* $\boldsymbol{\Phi}$ *is said to be a* concave *IVF on* $\mathcal{Y}$ *if for any* $y_1, y_2 \in \mathcal{Y}, \beta_1,\ \beta_2 \in [0,\ 1]$, *and* $\beta_1 + \beta_2 = 1$, *we have*

$$\beta_1 \odot \boldsymbol{\Phi}(y_1) \oplus \beta_2 \odot \boldsymbol{\Phi}(y_2) \preceq \boldsymbol{\Phi}(\beta_1 y_1 + \beta_2 y_2).$$

Lemma 1.11 *If* $\boldsymbol{\Phi}$ *is a concave IVF on a convex set* $\mathcal{Y} \subseteq \mathbb{R}^n$, *then* $\underline{\phi}$ *and* $\overline{\phi}$ *are concave on* $\mathcal{Y}$ *and vice-versa.*

Proof: This can be proved in a manner similar to the proof of Proposition 6.1 in [185]. □

Definition 1.13 (gH-continuity [63]). *An IVF* $\boldsymbol{\Phi}$ *is called* gH-continuous *at* $u \in \mathcal{Y}$ *if*

$$\lim_{\|d\|\to 0} (\boldsymbol{\Phi}(u+d) \ominus_{gH} \boldsymbol{\Phi}(u)) = \boldsymbol{0}.$$

If at every $u \in \mathcal{Y}$, $\boldsymbol{\Phi}$ *is* gH*-continuous, then* $\boldsymbol{\Phi}$ *is called* gH-continuous *on* $\mathcal{Y}$.

Lemma 1.12 (See [66]). *For a* gH*-continuous IVF* $\boldsymbol{\Phi}$, *its* $\underline{\phi}$ *and* $\overline{\phi}$ *are continuous and vice-versa.*

Definition 1.14 (gH-locally Lipschitz IVF [66]). *An IVF $\boldsymbol{T} : \mathcal{Y} \to I(\mathbb{R})$ on a non-empty subset $\mathcal{Y}$ of $\mathbb{R}^n$ is called gH-locally Lipschitz if there is a positive constant K and a δ neighbourhood $\mathcal{N}(z, \delta)$ such that*

$$\| \boldsymbol{T}(y) \ominus_{gH} \boldsymbol{T}(z)\|_{I(\mathbb{R})} \leq K\|y - z\| \text{ for all } y \in \mathcal{N}(z, \delta).$$

The constant K is referred to as Lipschitz constant of $\boldsymbol{T}$ at z.

Definition 1.15 (Domain of an IVF [65]) *Consider $\mathcal{Y}$ to be a non-empty subset of $\mathbb{R}^n$. For an extended IVF $\boldsymbol{T} : \mathcal{Y} \to \overline{I(\mathbb{R})}$, the domain of $\boldsymbol{T}$, denoted as dom($\boldsymbol{T}$), is expressed as*

$$dom(\boldsymbol{T}) = \{y \in \mathcal{Y} : \boldsymbol{T}(y) \prec +\infty\}.$$

Definition 1.16 (gH-lower-semicontinuous [123]) *An IVF $\boldsymbol{\Upsilon} : \mathcal{Z} \to I(\mathbb{R})$ is called gH-lower-semicontinuous (gH-lsc) at $\bar{z} \in \mathcal{Z}$ if*

$$\boldsymbol{\Upsilon}(\bar{z}) \preceq \liminf_{z \to \bar{z}} \boldsymbol{\Upsilon}(z). \tag{1.5}$$

Moreover, $\boldsymbol{\Upsilon}$ is said to be gH-lsc on $\mathcal{Z}$ if (1.5) is true for every $z \in \mathcal{Z}$.

Definition 1.17 (Infimum and supremum of an IVF [25]) *For an IVF $\boldsymbol{\Upsilon} : \mathcal{Z} \to I(\mathbb{R})$, the infimum (or, supremum) of $\boldsymbol{\Upsilon}$ over $\mathcal{Z}$ is defined as*

$$\inf_{\mathcal{Z}} \boldsymbol{\Upsilon} = \left[\inf_{\mathcal{Z}} \underline{\gamma},\ \inf_{\mathcal{Z}} \overline{\gamma}\right], \quad \left(or,\ \sup_{\mathcal{Z}} \boldsymbol{\Upsilon} = \left[\sup_{\mathcal{Z}} \underline{\gamma},\ \sup_{\mathcal{Z}} \overline{\gamma}\right]\right)$$

where $\inf_{\mathcal{Z}} \underline{\gamma} = \inf \left\{\underline{\gamma}(\varsigma) : \varsigma \in \mathcal{Z}\right\}$ *(or,* $\sup_{\mathcal{Z}} \underline{\gamma} = \sup \left\{\underline{\gamma}(\varsigma) : \varsigma \in \mathcal{Z}\right\}$*) and* $\inf_{\mathcal{Z}} \overline{\gamma} = \inf \{\overline{\gamma}(\varsigma) : \varsigma \in \mathcal{Z}\}$ *(or,* $\sup_{\mathcal{Z}} \overline{\gamma} = \sup \{\overline{\gamma}(\varsigma) : \varsigma \in \mathcal{Z}\}$*).*

Definition 1.18 (Limit infimum and limit supremum of an IVF [25]). *For an IVF $\boldsymbol{\Upsilon} : \mathcal{Z} \to I(\mathbb{R})$, the limit infimum (or, limit supremum) of $\boldsymbol{\Upsilon}$ at $\bar{z}$ in $\mathcal{Z}$ is given as*

$$\liminf_{\varsigma \to \bar{z}} \boldsymbol{\Upsilon}(\varsigma) = \left[\liminf_{\varsigma \to \bar{z}} \underline{\gamma}(\varsigma),\ \liminf_{\varsigma \to \bar{z}} \overline{\gamma}(\varsigma)\right],$$
$$\left(or,\ \limsup_{\varsigma \to \bar{z}} \boldsymbol{\Upsilon}(\varsigma) = \left[\limsup_{\varsigma \to \bar{z}} \underline{\gamma}(\varsigma),\ \limsup_{\varsigma \to \bar{z}} \overline{\gamma}(\varsigma)\right]\right)$$

where $\liminf_{\varsigma \to \bar{z}} \underline{\gamma}(\varsigma) = \lim_{\delta \to 0+} \left(\inf_{\varsigma \in \mathcal{B}(\bar{z}, \delta)} \underline{\gamma}(\varsigma)\right) \left(or,\ \limsup_{\varsigma \to \bar{z}} \underline{\gamma}(\varsigma) = \lim_{\delta \to 0+} \left(\sup_{\varsigma \in \mathcal{B}(\bar{z}, \delta)} \underline{\gamma}(\varsigma)\right)\right)$ *and* $\liminf_{\varsigma \to \bar{z}} \overline{\gamma}(\varsigma) = \lim_{\delta \to 0+} \left(\inf_{\varsigma \in \mathcal{B}(\bar{z}, \delta)} \overline{\gamma}(\varsigma)\right) \left(or,\ \limsup_{\varsigma \to \bar{z}} \overline{\gamma}(\varsigma) = \lim_{\delta \to 0+} \left(\sup_{\varsigma \in \mathcal{B}(\bar{z}, \delta)} \overline{\gamma}(\varsigma)\right)\right)$.

Definition 1.19 (Efficient point [64]). *Consider $\mathcal{Y} \subseteq \mathbb{R}^n$ and $\mathbf{\Phi} : \mathbb{R}^n \to I(\mathbb{R})$ to be an IVF. A point $u \in \mathcal{Y}$ is called an* efficient point *of the IVF $\mathbf{\Phi} : \mathcal{Y} \to I(\mathbb{R})$ if $\mathbf{\Phi}(y) \not\prec \mathbf{\Phi}(u)$ for all $y \in \mathcal{Y}$.*

Definition 1.20 (Weak efficient point [68]). *Consider $\mathcal{Y} \subseteq \mathbb{R}^n$ and $\mathbf{\Phi} : \mathbb{R}^n \to I(\mathbb{R})$ to be an IVF. A point $u \in \mathcal{Y}$ is called a* weak efficient point *of the IVF $\mathbf{\Phi} : \mathcal{Y} \to I(\mathbb{R})$ if $\mathbf{\Phi}(u) \preceq \mathbf{\Phi}(y)$ for all $y \in \mathcal{Y}$.*

Definition 1.21 (Linear IVF [64]). *Consider $\mathcal{Z}$ to be a linear subspace of $\mathbb{R}^n$. The function $\mathbf{\Upsilon} : \mathcal{Z} \to I(\mathbb{R})$ is called linear if*

(i) $\mathbf{\Upsilon}(\lambda v) = \lambda \odot \mathbf{\Upsilon}(v)$ for all $v \in \mathcal{Z}$ and $\lambda \in \mathbb{R}$, and

(ii) for all $v_1, v_2 \in \mathcal{Z}$, either $\mathbf{\Upsilon}(v_1) \oplus \mathbf{\Upsilon}(v_2) = \mathbf{\Upsilon}(v_1 + v_2)$ or none of $\mathbf{\Upsilon}(v_1) \oplus \mathbf{\Upsilon}(v_2)$ and $\mathbf{\Upsilon}(v_1 + v_2)$ dominates the other.

Definition 1.22 (Proper IVF [123]). *An IVF $\mathbf{\Upsilon} : \mathcal{Z} \to I(\mathbb{R}) \cup \{-\infty, +\infty\}$ is called a proper IVF if there is a $\bar{z} \in \mathcal{Z}$ so that*

$$\mathbf{\Upsilon}(\bar{z}) \prec +\boldsymbol{\infty}, \text{ and } -\boldsymbol{\infty} \prec \mathbf{\Upsilon}(z) \text{ for all } z \in \mathcal{Z}.$$

Definition 1.23 (gH-derivative [23]). *Consider $\mathcal{Y} \subseteq \mathbb{R}^n$. The gH-*derivative *of an IVF $\mathbf{\Phi} : \mathcal{Y} \to I(\mathbb{R})$ at $u \in \mathcal{Y}$ is defined as the limit*

$$\mathbf{\Phi}'(u) := \lim_{d \to 0} \tfrac{1}{d} \odot \{\mathbf{\Phi}(u + d) \ominus_{gH} \mathbf{\Phi}(u)\}.$$

Definition 1.24 (gH-Gáteaux derivative [64]). *Consider an IVF $\mathbf{\Phi}$ to be defined on a non-empty open subset $\mathcal{Y}$ of $\mathbb{R}^n$. Then, $\mathbf{\Phi}$ is called gH-*Gáteaux differentiable *with* gH-Gáteaux derivative $\mathbf{\Phi}_{\mathscr{G}}(u)$ *at $u \in \mathcal{Y}$ if the limit below*

$$\mathbf{\Phi}_{\mathscr{G}}(u)(h) := \lim_{\beta \to 0+} \frac{1}{\beta} \odot (\mathbf{\Phi}(u + \beta h) \ominus_{gH} \mathbf{\Phi}(u))$$

is finite for all $h \in \mathbb{R}^n$ and $\mathbf{\Phi}_{\mathscr{G}}(u)$ is a gH-continuous and linear IVF from $\mathbb{R}^n$ to $I(\mathbb{R})$.

Definition 1.25 (gH-Fréchet derivative [64]). *Consider an IVF $\mathbf{\Phi}$ on a non-empty open subset $\mathcal{Y}$ of $\mathbb{R}^n$. Then, $\mathbf{\Phi}$ is called gH-*Fréchet differentiable *at $u \in \mathcal{Y}$ if there is a gH-continuous and linear mapping $\boldsymbol{G} : \mathcal{Y} \to I(\mathbb{R})$ so that*

$$\lim_{\|h\| \to 0} \tfrac{1}{\|h\|} \odot \left(\|\mathbf{\Phi}(u + h) \ominus_{gH} \mathbf{\Phi}(u) \ominus_{gH} \boldsymbol{G}(h)\|_{I(\mathbb{R})}\right) = 0,$$

where $\boldsymbol{G}$ is referred to as $\mathbf{\Phi}_{\mathscr{F}}(u)$.

Definition 1.26 (gH-Hadamard derivative [25]). *Consider an IVF* $\boldsymbol{\Upsilon} : \mathcal{Z} \to I(\mathbb{R})$*. If for* $\bar{z} \in \mathcal{Z}$ *and* $v \in \mathbb{R}^n$*, the following limit exists*

$$\boldsymbol{\Upsilon}_{\mathscr{H}}(\bar{z})(v) = \lim_{\substack{t \to 0+ \\ h \to v}} \tfrac{1}{t} \odot (\boldsymbol{\Upsilon}(\bar{z} + th) \ominus_{gH} \boldsymbol{\Upsilon}(\bar{z})),$$

then we write $\boldsymbol{\Upsilon}_{\mathscr{H}}(\bar{z})(v)$*. If* $\boldsymbol{\Upsilon}_{\mathscr{H}}(\bar{z})$ *is a linear IVF from* $\mathcal{Z}$ *to* $I(\mathbb{R})$*, then* $\boldsymbol{\Upsilon}_{\mathscr{H}}(\bar{z})(v)$ *is said to be gH-Hadamard derivative of* $\boldsymbol{\Upsilon}$ *along the direction* v *at* $\bar{z}$*. If this limit exists for all* $v \in \mathbb{R}^n$*, then* $\boldsymbol{\Upsilon}$ *is said to be gH-Hadamard differentiable at* $\bar{z}$*.*

1.1.4 Literature survey

1.1.4.1 Interval analysis

The goal of interval analysis is to confine the impacts of uncertainties and errors in a quantity within a specified interval. To achieve this, various researchers have independently developed arithmetic methods for interval analysis. Notable contributions include those by Dwyer in 1951 [46], Warmus in 1956 [183], Sunaga in 1958 [172], and Moore in 1966 [142]. Although Warmus and Sunaga were among the first to develop interval arithmetic, Moore is widely recognized as the true father of interval analysis due to his pioneering contributions, particularly those presented in his 1966 book [142]. His work transformed the concept of intervals into a practical tool for error analysis. Among the subsequent developments, [142] established interval arithmetic for intervals with finite endpoints. Kahan [99] and Hanson [80] extended interval arithmetic to accommodate intervals with infinite endpoints. Notably, Hanson [81] developed a practical Newton interval algorithm that enabled division by an interval containing zero.

Conventional interval arithmetic is insufficient for finding the additive inverse of a non-degenerate interval $\mathbf{A}$; there is no interval $\mathbf{B}$ such that $\mathbf{A} \oplus \mathbf{B} = \mathbf{0}$. Furthermore, the conventional definition of the difference between two compact intervals has several deficiencies. For example,

(i) subtraction of $\mathbf{A}$ and $\mathbf{B}$ does not result in $\{0\}$,

(ii) if $\mathbf{A}$ and $\mathbf{C}$ are two compact intervals, then $(\mathbf{A} \ominus_{gH} \mathbf{C}) \oplus \mathbf{C} \neq \mathbf{A}$. For example, $\mathbf{A} = [-1, 2], \mathbf{C} = [-5, 6]$, then $(\mathbf{A} \ominus_{gH} \mathbf{C}) \oplus \mathbf{C} = [-9, 10] \neq \mathbf{A}$.

(iii) For $\mathbf{C} = \mathbf{A} \ominus \mathbf{B}$, $\mathbf{B} \oplus \mathbf{C}$ may not be equal $\mathbf{A}$. For example, for $\mathbf{A} = [1, 3]$ and $\mathbf{B} = [4, 6]$, if $\mathbf{C} = \mathbf{A} \ominus \mathbf{B}$, then $\mathbf{B} \oplus \mathbf{C} = [-1, 5] \neq \mathbf{A}$.

To address these issues, Hukuhara [89] introduced a workaround known as the Hukuhara-difference ($\ominus_H$) between intervals. However, the Hukuhara-difference also had significant limitations. For instance, the operation $\mathbf{A} \ominus_{gH} \mathbf{B}$ is defined only when the interval width of $\mathbf{A}$ is less than or equal to the interval width of $\mathbf{B}$. Hukuhara-difference between $\mathbf{A}$ and $\mathbf{B}$ does not exist [24]. Markov addressed the limitations of the Hukuhara difference in 1979 by introducing the non-standard difference of intervals [136]. Similarly, Stefanini and Bede extended the concept further in 2009 to introduce the generalized Hukuhara difference ($\ominus_{gH}$) [168]. The concept of generalized difference enabled the operation of finding the difference between any two intervals. Next, Stefanini combined Moore's interval arithmetic with gH-difference, demonstrating the law of cancellation for interval addition and the distributive law for interval subtraction by a scalar. Furthermore, [134] utilized Moore's interval norm and gH-difference to demonstrate that the set of compact intervals forms a quasi-normed linear space. Later, in 2016, Tao [174] presented findings on semi-linear interval differential equations utilizing the gH-difference. Among recent studies, Ghosh in 2019 [64] and Kumar et al. in 2020 [122] extended the concepts of smooth and non-smooth analysis to interval-valued functions. They also formulated interval variational inequalities using dominance relations and the gH-difference.

1.1.4.2 Calculus of interval-valued functions

To analyze the characteristics of interval-valued functions, calculus serves as a fundamental tool. Moreover, it plays a critical role in devising techniques aimed at solving optimization problems that incorporate interval-valued functions. The pioneering work in the calculus of interval-valued functions dates back to Hukuhara's introduction of H-differentiability for IVFs using H-difference in 1967 [89]. However, H-differentiability is recognized to have significant limitations, as discussed in [24]. To address the shortcomings of H-differentiability, Bede and Gal [14] introduced the concept of strongly generalized derivative (G-derivative) for IVFs. They also formulated a Newton-Leibniz-type equation in their work. Markov in 1979 [136] investigated the mean-value theorem for IVFs by introducing the concept of interval difference and defining the differentiability of IVFs based on this notion. In 2009, Stefanini and Bede [167] expanded the concept of generalized Hukuhara difference to introduce generalized Hukuhara differentiability (gH-differentiability) for IVFs. Various essential concepts like gH-derivative, gH-gradient, gH-partial derivative, and gH-differentiability have been formally defined in studies such as [63, 166, 168].

Ghosh et al. [64] introduced the notions of gH-directional derivative, gH-Gâteaux derivative, and gH-Fréchet derivative for IVFs, aimed at analyzing optimality conditions in interval optimization problems. Kumar and Ghosh [123] have introduced Ekeland's variational principle tailored for IVFs in their research. Ghosh et al. [71]

have extended the definitions of gH-differentiability, gradient, partial derivative, and directional derivative to accommodate interval-valued functions with interval-valued variables. Ghosh et al. [65] have recently explored the concepts of gH-subgradient and gH-subdifferential in the context of non-smooth and non-differentiable convex IVFs. In their study, the authors demonstrated that the gH-directional derivative equals the maximum of all products of directions and gH-subgradients. In Chauhan et al. [26], they introduced the concept of gH-Clarke derivative for non-smooth IVFs. Anshika et al. [68] investigated the gH-subdifferential of interval-valued functions. Upadhyay et al. [173] explored generalized subdifferential for non-smooth non-convex IVFs, focusing on a class of generalized approximate LU-convex functions using Mordukhovich's subdifferential. Recently, Ghosh et al. [70] introduced the concept of gH-weak subdifferential and examined its properties, particularly the inclusion for sum rule of this concept for general IVFs.

1.1.4.3 Interval Optimization Problem

In recent decades, the study of interval optimization problems has significantly gained popularity, emerging as a prominent research area in applied mathematics. In early research, Wu [185] introduced two solution concepts for interval optimization problems in 2007. These concepts involved employing partial ordering on the set of closed intervals and applying the H-derivative to derive KKT optimality conditions. Later, in 2009, Wu [187] extended the study to KKT optimality conditions for multi-objective IOPs. In 2012, Bhujree and Panda [15] investigated a general optimization problem by representing IVFs in parametric form and developed a method to find the existence of efficient solutions. In 2013, Chalco-Cano et al. [23] introduced KKT optimality criteria for interval optimization problems using the gH-derivative, demonstrating its advantages over conventional H-derivative methods. Additionally, Singh et al. [163] introduced the concept of Pareto optimal solutions for interval-valued multi-objective optimization in 2016. Furthermore, Ghosh [63] proposed a Newton method along with an enhanced version to tackle IOPs. Several authors have made significant contributions to the study of optimality conditions and duality in IOPs. For instance, Wu [186] explored Wolfe duality for non-linear IOPs and derived duality theorems using concepts of non-dominated solutions. Wu [184] also introduced Lagrangian duality for non-linear IOPs. Zhang [193] developed KKT conditions that are necessary and sufficient for IOPs under invexity assumptions. Furthermore, they proposed a strong duality theorem that does not rely on duality gap assumptions in both strong and weak senses.

With respect to literature on non-smooth interval optimization, Antczak [8] demonstrated the existence of weakly LU-efficient solutions using Fritz John and KKT optimality conditions. They also established Mond-Weir type duality for a set of non-smooth non-differentiable interval-valued multiobjective optimization problems under convexity assumptions. Kummari and Ahmad [126] introduced KKT type sufficient optimality conditions for IOPs under L-invex-infine functions defined with the limiting subdifferential of locally Lipschitz functions and obtained appropriate duality theorems for a Wolfe type dual model. Kumar et al. [125] developed a dual characterization for the set of weak sharp minima of interval-valued objective functions in convex unconstrained and constrained IOPs, utilizing their newly introduced gH-subdifferential calculus. Initially, Karaman [105] introduced concepts related to subdifferentials and weak subdifferentials. Subsequently, Karaman [102] explored interval scalar optimization problems through subdifferential approaches for IVFs. Finally, Ghosh et al. [71] defined non-smooth normal and tangent cone concepts for sets of intervals and applied them to characterize efficient solutions in IOPs using Lagrange multipliers.

Researchers have also made significant progress in non-smooth and non-convex interval optimization. Upadhyay et al. [178] established connections between a class of generalized Stampacchia vector variational inequalities and non-smooth interval-valued multiobjective programming problems (NIVMPP). They identified KKT vector critical points for NIVMPP. Anshika et al. [68] applied gH-subdifferential calculus for convex IVFs to a non-convex composite model of IOPs. Kumar et al. [124] introduced the concept of gH-Fréchet subdifferential for general IVFs and derived necessary optimality conditions for unconstrained non-convex IOPs, as well as necessary conditions for unconstrained weak sharp minima for IVFs. Ghosh et al. [67] proposed the gH-Clarke subdifferential for a class of IVFs and applied it to establish sufficient optimality conditions for a specific non-smooth IOP where the objective IVFs are support functions of given convex and compact sets. In their recent work, Ghosh et al. [70] utilized the gH-weak subdifferential to derive various optimality conditions for non-smooth IOPs.

1.2 Set-valued optimization

Set optimization is a branch of optimization that focuses on minimizing or maximizing objectives that are set-valued rather than single values. This framework is more general, allowing the objective set to consist of various types of objects, including scalars, vectors, intervals, and others. It can be viewed as an extension of multi-objective optimization, where the objectives are not restricted to scalar values but can encompass diverse types and structures.

To differentiate set optimization from traditional scalar, vector, or interval optimization, let's examine the following hypothetical example problems, adapted from [7], which involve identifying the optimal player/bowler or team in the sport of cricket.

- Among all cricket players (the decision space), identifying the fastest bowler based solely on their bowling speed can be formulated as a *single-valued scalar optimization problem*.

- The task of identifying the bowler(s) based on their multiple qualities, such as speed, in-swing/out-swing, bounce, etc., can be formulated as a *vector optimization problem*. In this context, the objective function maps each player to a vector that represents their qualities as individual components.

- Note that the speed, swing angle, and bounce of a bowler are not fixed and can vary over a range during a cricket match, series, or throughout their career. This variability introduces inherent uncertainty in these objective quantities. For instance, the speed of a fast bowler typically varies between $110 - 160$ km/h, while a spin bowler's speed may range from $65 - 90$ km/h. The swing angle can vary from approximately -40 degrees (in-swing) to $+40$ degrees (out-swing) from the axis of the ball's initial trajectory, and the bounce of a delivery can range from around 1 foot to less than 7 feet. Due to this variability, these variables cannot be treated as fixed scalar quantities but should instead be considered as interval-valued variables. Using a single quality such as speed, the task of finding the optimal player becomes an *interval-valued single-objective optimization problem*. However, when considering all these qualities together, we face an *interval-valued multi-objective optimization problem*.

- Finally, the task of identifying good teams from all the teams in a cricket league, where each team is composed of multiple players, each player assigned values for their qualities (speed, in-swing, out-swing, bounce), can be formulated as a set optimization problem. Here, each set represents a team, with the elements of the set being vectors that represent players. Each vector's components correspond to the four different qualities.

In recent years, set optimization has garnered significant attention from researchers due to its mathematical advantages and practical applications. Its robust and versatile structure allows for modeling various decision-making scenarios characterized by uncertainty, ambiguity, and multi-valued outputs. Set optimization finds practical utility across diverse fields including game theory [77], economics [113], mathematical finance [138], duality theory [76], optimal control [79], uncertain optimization [118], radiotherapy treatment [128], socio-economics [148], welfare economics [12], and others.

In the field of mathematical finance, set-valued formulations of risk measures in portfolio optimization are utilized to focus on individual components of portfolios, thereby minimizing the risk of high losses [138]. For a comprehensive survey on the applications of set optimization, please refer to [111].

The main reason for increasing popularity of set optimization is that it offers several mathematical advantages over other classical optimization approaches, for example:

(1) *Set-valued approach in convex analysis:* In the field of convex analysis, whether dealing with scalar-valued, vector-valued, or set-valued functions, fundamental concepts like duality, subdifferentials, and conjugates play a crucial role. However, in the context of vector-valued functions, there has been limited exploration of a canonical convex analysis that encompasses these concepts comprehensively. Many existing definitions of conjugates for vector-valued functions impose restrictive assumptions, hindering their applicability. In contrast, the set-valued approach in duality theory offers a definition of conjugates that avoids such limitations, making it suitable for vector-valued functions as well.

(2) *Set-order relation for defining infimum and supremum:* In majority of the literature on vector-valued optimization problems, the notions of infimum and supremum do not have a meaning. This is mostly because, for a preorder set with vector relation, some of the subsets may not have an infimum or supremum if the ordering cone does not make the preorder set a complete lattice structure. As a result, the infimum or supremum of the whole set can also not be defined. This, therefore, is unable to offer the useful concept which could be used to find the non-dominated (i.e. the minimal or maximal) solutions. At this point, defining the set order relation on the power set of the preorder set (instead of defining the relation on the preorder set itself) provides a way of filling the absence of infimum and supremum in vector optmization. [132] showed this by extending the set valued approach to complete lattice approach. Additionally, this approach leads to solution concepts which now is not restricted to finding only the maximal or minimal points but can further look for solution points from the perspective of infimum or supremum.

(3) *Set-valued duality in terms of scalarization:* Scalarization techniques play an important part in multi-objective optimization. A convex vector-valued or set-valued function can be equivalently described as a collection of extended real-valued functions, for example as the collection of support functions of the images. In terms of duality of set-valued functions, Fenchel conjugate of set valued-function can be defined in terms of Fenchel conjugate of the corresponding scalar-valued function.

Therefore, scalarization helps to develop a strong duality theory for set-valued functions. This set-valued duality theory can now be used to develop effective algorithms for vector and set optimization problems. Moreover, this offers a new geometric duality that when applied to polyhedral and linear vector optimization problems can be used to develop algorithms that find the solutions from the set-valued viewpoint [160].

(4) *Set-valued tools for mathematical finance:* Set-valued approaches in finance are experiencing an increase in popularity due to their effectiveness in addressing market risks affected by frictions like transaction costs and bid-ask price [78]. Recently developed tools such as set-based superhedging portfolios and set-valued risk measures have proven instrumental in analyzing risk in such complex market environments. Notably, established results in finance, such as Kabanov's superhedging theorem [100], have been shown to be specific instances of more generalized set-valued duality theorems. Optimization problems in finance, such as risk minimization, optimal risk allocation, and constrained hedging, are now being approached from the viewpoint of set-valued risk measures. This approach is proving valuable across disciplines, including mathematical finance, statistics, and insurance mathematics, and it offers robust tools for studying multivariate risks.

We begin our study of set optimization by defining set-valued functions, thereby diving into the core mathematics of set-valued optimization.

1.2.1 Set-valued function

At the center of set optimization is the set-valued function or map. A set-valued map $F : M \to Y$, for any two linear spaces M and Y, is a correspondence $x \to F(x) \subseteq Y$. It can also be regarded as a function to the power set of Y, i.e. $F : M \to 2^Y$. Formally, the domain of F can be defined as

$$\mathrm{Dom}(F) = \{x \in M : F(x) \neq \emptyset\}.$$

The image of F can be defined as

$$\mathrm{Im}(F) = \bigcup_{x \in \mathrm{Dom}(F)} F(x).$$

In contrast to traditional scalar-valued, interval-valued, or vector-valued maps, which are limited to scalar functions, interval functions, or sets of scalar functions respectively, set-valued maps are not confined to any specific type of function and can take any structure. The image of a set-valued map can consist of a variety of forms, including collections of scalars, intervals, lines, triangular regions, rectangular regions, circular regions, vectors, and combinations of these. Following are a few examples:

1. Collection of circular regions: $F : \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined by

$$F(x) = \{(a, b) \in \mathbb{R}^2 : a^2 + b^2 \leq x^2 + 1\}, \quad \text{for all } x \in \mathbb{R}.$$

2. Collection of square regions: $F : [0, \infty) \rightrightarrows \mathbb{R}^2$ defined by

$$F(x) = [-x, x] \times [-x, x], \quad \text{for all } x \in [0, \infty).$$

3. Collection of triangular regions: $F : \mathbb{R} \rightrightarrows \mathbb{R}^2$ defined by

$$F(x) = conv\{(x, x), (x + 1, x), (x, x + 1)\}, \quad \text{for all } x \in \mathbb{R}.$$

In this thesis, we focus on the following particular form of set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$

$$F(x) = \left\{f^1(x), f^2(x), \ldots, f^p(x)\right\}, \ x \in \mathbb{R}^n, \tag{1.6}$$

where $f^1, f^2, \ldots, f^p : \mathbb{R}^n \to \mathbb{R}^m$ are twice continuously differentiable vector-valued functions.

1.2.2 Set optimization problem formulation and its solution concept

Mathematically, an unconstrained SOP is given by

$$\text{(SOP)} \begin{cases} \min F(x) \\ \text{subject to } x \in \mathbb{R}^n, \end{cases} \tag{1.7}$$

where $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is a set-valued map with $F(x) \neq \emptyset$ for all $x \in \mathbb{R}^m$. There are primarily two approaches in the literature to define the solution concept of the SOP (1.7): the *vector approach* and the *set approach*.

Vector approach compares the vectors y's in the image set

$$F(\mathbb{R}^n) = \{y \in \mathbb{R}^m : y \in F(x) \text{ for some } x \in \mathbb{R}^n\}.$$

An element $x_0 \in \mathbb{R}^n$ is referred to as a minimal solution of the SOP if there exists $y_0 \in F(x_0)$ such that y_0 is a minimal element of the collection of y's that are in $F(\mathbb{R}^n)$. In other words, the optimal solution x_0 and the corresponding optimal set $F(x_0)$ are identified based solely on one of its elements, $y_0 \in F(x_0)$, being the minimal element in $F(\mathbb{R}^n)$, while disregarding all other elements of $F(x_0)$. However, this formulation of the solution concept has a significant drawback from a modeling perspective. For instance, consider the cricket example discussed in Section 1.2. When identifying an optimal team using the vector approach, if a team $F(x_0)$ consists of only one exceptional player $y_0 \in F(x_0)$ who is the best in the league (i.e., y_0 best in $F(\mathbb{R}^n)$), this team might still be considered the best team, even if its other members are below average. Therefore, for practical purposes, this approach may be inadequate and misleading.

In contrast, the *set approach* does not encounter this issue. This approach first establishes a preference or set order relation (generated by a preorder $\preceq$) on the power set constructed from the image space of set-valued objective functions. It then looks for the minimal element from the collection of set-valued objectives. In other words, the set approach considers the entire set to determine the optimal solution rather than focusing on a single point in the image space.

In terms of set order relations, there are numerous options discussed in the literature. Some well-known examples include upper and lower set order relations, weighted set order relations, min-max order relations, Karaman's m_1 order relation using Minkowski difference, and set-less or KNY order relation, among others. For a comprehensive review of set order relations, please see Section 1.2.5.1. Among these, we specifically utilize the lower set order relation, which is elaborated in detail in the following subsection, Section 1.2.3.

1.2.3 Basic definitions

For the last two chapters on set optimization, we recall the following definitions:

To compare any two vectors or sets, an ordered relation needs to be defined. To establish this ordered relation, we first recall the concept of a cone. A non-empty $K \in \mathscr{P}(\mathbb{R}^m)$ is said to be a cone in $\mathbb{R}^m$ if $y \in K$ means that $ty \in K$, for all $t \geq 0$. A cone K is convex if $K + K = K$, pointed if $K \cap (-K) = \{0\}$, and solid if $\text{int}(K) \neq \emptyset$. If K is a convex and pointed cone in $\mathbb{R}^m$, it generates a partial ordering on $\mathbb{R}^m$ given by

$$y \preceq_K z \Longleftrightarrow z - y \in K,$$

and moreover, if K is solid, it gives a strict ordering on $\mathbb{R}^m$ defined by

$$y \prec_K z \Longleftrightarrow z - y \in \operatorname{int}(K).$$

For the last two chapters, we will use the notation K to always mean a closed, pointed, convex and solid cone in $\mathbb{R}^m$.

Next, we define the crucial concept of the minimal element of a set.

Definition 1.27 [111] *Let $\mathcal{A} \in \mathscr{P}(\mathbb{R}^m)$ and $y_0 \in \mathcal{A}$.*

(i) The element y_0 is said to be a minimal element of $\mathcal{A}$ with respect to K if $(y_0 - K) \cap \mathcal{A} = \{y_0\}$. The collection of all minimal elements of $\mathcal{A}$ is referred to as $Min(\mathcal{A}, K)$.

(ii) The element y_0 is said to be a weakly-minimal element of $\mathcal{A}$ with respect to K if $(y_0 - int(K)) \cap \mathcal{A} = \emptyset$. The collection of all weakly-minimal elements of $\mathcal{A}$ is referred to as $WMin(\mathcal{A}, K)$.

Next, we discuss the lower set less relation $\preceq^l_K$ on $\mathscr{P}(\mathbb{R}^m)$ with respect to a given K for comparing pairs of sets.

Consider that $\mathcal{A}, \mathcal{B} \in \mathscr{P}(\mathbb{R}^m)$. The relation $\preceq^l_K$ on $P(\mathbb{R}^m)$ is given as

$$\mathcal{A} \preceq^l_K \mathcal{B} \Longleftrightarrow \mathcal{B} \subseteq \mathcal{A} + K.$$

Similarly, the strict lower set less relation $\prec^l_K$ on $\mathscr{P}(\mathbb{R}^m)$ is defined as

$$\mathcal{A} \prec^l_K \mathcal{B} \Longleftrightarrow \mathcal{B} \subseteq \mathcal{A} + \operatorname{int}(K).$$

Next, we use the relation $\preceq^l_K$ to define the concept of optimality for the SOP having the set-valued map (1.6). In this map, j-th component function of the function $f^i : \mathbb{R}^n \to \mathbb{R}^m$ is denoted by $f^{i,j}$, i.e.,

$$f^i(x) = \left(f^{i,1}(x), f^{i,2}(x), \ldots, f^{i,m}(x)\right)^\top, \ x \in \mathbb{R}^n.$$

We can also denote the function F in (1.6) as $F = \{f^i\}_{i \in [p]}$. Then, under the relation $\preceq^l_K$, we reformulate the SOP (1.7) as the following unconstrained set optimization problem

$$(\preceq^l_K) \quad \min_{x \in \mathbb{R}^n} F(x), \qquad (\mathcal{SOP}^l_K)$$

where the solution concept in term of $\preceq^l_K$ is given as follows: we call $\bar{x} \in \mathbb{R}^n$ a (resp., weakly) minimal solution of ($\mathcal{SOP}^l_K$) if there does not exist any $x \in \mathbb{R}^n$ such that (resp., $F(x) \prec^l_K F(\bar{x})$) $F(x) \preceq^l_K F(\bar{x})$.

Next, we revisit several auxiliary functions involving active indices. These functions aid in identifying weakly-minimal points of an SOP through the solution of a family of vector optimization problems.

Definition 1.28 [18]

(i) The function $I : \mathbb{R}^n \rightrightarrows [p]$, expressed as

$$I(x) = \{i \in [p] : f^i(x) \in Min(F(x), K)\},\ x \in \mathbb{R}^n,$$

is termed the set of active indices of minimal elements of the set $F(x)$;

(ii) the function $I_0 : \mathbb{R}^n \rightrightarrows [p]$, expressed as

$$I_0(x) = \{i \in [p] : f^i(x) \in WMin(F(x), K)\},\ x \in \mathbb{R}^n,$$

is termed the set of active indices of weakly-minimal elements of the set $F(x)$;

(iii) for a vector $v \in \mathbb{R}^m$, the function $I_v : \mathbb{R}^n \rightrightarrows [p]$ is given by

$$I_v(x) = \{i \in I(x) : f^i(x) = v\}.$$

(iv) The map $\omega : \mathbb{R}^n \to \mathbb{R}$ is given as $\omega(x) = |Min(F(x), K)|,\ x \in \mathbb{R}^n$.

Definition 1.29 [18] *For a $\bar{x} \in \mathbb{R}^n$, let $\{v_1^{\bar{x}}, v_2^{\bar{x}}, \ldots, v_{w(\bar{x})}^{\bar{x}}\}$ be an enumeration of the set $WMin(F(\bar{x}), K)$. Then, the* partition set *at $\bar{x}$ is given by*

$$P_{\bar{x}} = I_{v_1^{\bar{x}}}(\bar{x}) \times I_{v_2^{\bar{x}}}(\bar{x}) \times \cdots \times I_{v_{w(\bar{x})}^{\bar{x}}}(\bar{x}).$$

Note that $\bigcup_{j=1}^{[w(\bar{x})]} I_{v_j^{\bar{x}}}(\bar{x}) = I_0(\bar{x})$ and for any $j_1, j_2 \in [\omega(\bar{x})]$ with $j_1 \neq j_2$, we have $I_{v_{j_1}^{\bar{x}}}(\bar{x}) \cap I_{v_{j_2}^{\bar{x}}}(\bar{x}) = \emptyset$.

Definition 1.30 (Regular point for ($\mathcal{SOP}^l_K$) [18]). *A point $\bar{x}$ is called a regular point of the set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ if the following conditions hold:*

(i) $Min(F(\bar{x}), K) = WMin(F(\bar{x}), K)$, and

(ii) the cardinality function ω as per Definition 1.28 is constant in a neighbourhood of $\bar{x}$.

Lemma 1.2.1 [18] *Let $\bar{x}$ is a regular point of F, then there is a neighborhood $\mathcal{N}$ of $\bar{x}$ such that*

$$\omega(x) = \omega(\bar{x}) \text{ and } P_x \subseteq P_{\bar{x}} \text{ for all } x \in \mathcal{N}.$$

1.2.4 Methods of solving set optimization problems

To solve the aforementioned problem ($\mathcal{SOP}^l_K$), we need the help of numerical methods and algorithms. In the literature of set approach-based methods, seven main types of numerical algorithms are identified: *SetOpt* (for polyhedral and convex set optimization problems), *Newton methods*, *Derivative-free methods*, *Sorting-based methods*, *Branch-and-Bound method*, and *Steepest-descent method*. These methods can be further categorized into three categories based on their approach to solving an SOP:

1. Methods which transform the set optimization problem into a scalar optimization problem using a technique known as *scalarization*. Subsequently, solve the resulting scalar optimization problem using established numerical algorithms. Newton method, derivative-free methods, sorting-based methods, and the branch-and-bound method fall into this category.

2. Methods that transform the set optimization problem into a vector optimization problem using a *vectorization* technique, and subsequently solve this vector optimization problem using established methods found in the literature. This include method like SetOpt.

3. Methods that convert the set optimization problem into a scalar optimization problem by first transforming it through an intermediate vectorization step, then applying a scalarization technique. Finally, solve the resulting scalar optimization problem using established numerical algorithms. Steepest-descent method belong to this category.

Scalarization of an SOP involves converting a set optimization problem into a scalar form that can be solved using existing algorithms from the literature. This approach works by defining the minimal solutions of the original set optimization problem in terms of the optimal solutions of the transformed scalar problem. Scalarization techniques utilize a scalarization functional, which can be linear or non-linear. Among the non-linear functionals, oriented distance, Gerstewitz, and Drummond-Svaiter functionals are the most popular. For further information on scalarization functionals and their applications, please refer to Section 1.2.5.2.

Vectorization of an SOP involves applying vectorial relaxations to the original set optimization problem, thereby transforming it into a family of multi-objective optimization problems, which can be solved relatively easily.

The aforementioned methods, however, each have their own drawbacks. For example, SetOpt is limited to polyhedral, convex set-valued maps. The Newton method is constrained by metric regularity assumptions. Derivative-free methods do not utilize first and second-order derivative information. For the form of ($\mathcal{SOP}^l_K$) considered in our paper, sorting-based and branch-and-bound methods are not applicable. Lastly, the steepest-descent method suffers from a slow rate of convergence and does not guarantee global convergence.

Therefore, in this thesis, we propose three new methods for solving a set optimization problem, following the approach in the third category (mentioned above). Each of our proposed methods consists of three steps:

Step 1. Convert the SOP into an intermediate vector optimization problem (VOP) using a vectorization technique.

Step 2. Convert the intermediate VOP into a scalar optimization problem using a scalarization technique.

Step 3. Solve the final scalar optimization problem using an appropriate numerical algorithm.

In Step 1., our proposed vectorization technique is inspired by the strategy used in the steepest-descent method, converting the discrete set optimization problem into a family of vector optimization problems (VOPs) with respect to a vector order relation equivalent to the lower set order relation. These VOPs then require a scalarization technique for solving. Therefore, in Step 2., we use the oriented distance scalarization technique to convert the VOP into a scalar optimization problem (whereas the steepest-descent method uses the Gerstewitz functional). We choose the oriented distance functional due to its computational ease and its flexibility in handling optimization problems regardless of convexity and solid cone assumptions. In Step 3., we employ trust region-based methods to solve the scalar optimization problem.

As a further contribution, we highlight that the trust region method incorporates a monotonic step criterion within its algorithm. This strict criterion, requiring a decrease in function values at each iteration, may not only slow down the convergence speed but also hinder convergence, particularly for ill-conditioned and highly non-linear optimization problems. With this in mind, non-monotonic trust-region schemes, which relax the requirement for monotonically decreasing function values, can be promising in such scenarios. As observed in the fields of single-objective and multi-objective optimization, these schemes can enhance the possibility and speed of convergence, and they hold potential for improving set optimization as well.

Motivated by this, we propose two new non-monotonic trust-region methods (NTRM) for set optimization, based on two popular variants: the Max-type NTRM and the Average-type NTRM for set optimization in Chapter 7. NTRM differs from TRM primarily by adopting a non-monotonic step acceptance criterion instead of a monotonic one. For these methods, the Max-type NTRM considers the maximum of function values from the last M iterations, while the Avg-type NTRM uses an exponentially weighted moving average of function values up to the current iteration.

1.2.5 Literature survey

In this thesis, the proposed solution concept for the set optimization problem is based on set order relation. This relation primarily helps in identifying the minimal solutions of the SOP by offering a method to compare the function's values in the image space. We start with a brief literature review on set order relations in the following section.

1.2.5.1 Set order relation

The idea of set approach was first introduced by Kuroiwa [127] where six types of set relations were studied. Kuroiwa's set approach was generalized by Karaman et al [103] and Jahn and Ha. [97]. Since then, set order relations have extensively studied with respect to duality theory, existence of solution, optimality criteria, and for numerous numerical algorithms for solving set optimization problems. Many types and variations of set relations have been studied in the literature. For example, Kuroiwa's upper and lower set order relation [127], weighted set order relation [28], minmax order relation [97], Karaman's m_1 order relation using Minkowski difference [104], set-less or KNY order relation [4], and many more. For more details see [3, 27, 85, 86, 114, 115].

Especially for studying optimality conditions of SOP, set relations with generalized differentiation concepts (belonging to the primal as well as dual space) have been extensively studied. For example, [96] considered the set optimization problem where the set-valued map is described by functional constraints. Using a vectorization scheme derived by Jahn [95], this set-valued problem was converted to a vector-valued problem on which the classical KKT optimality condition was employed to find the solution. [119, 120] considered the idea where, under certain assumptions, the well-established optimality criteria for vector approach can be used in the sense of optimality for set approach. Karaman et al. [105], using generalized directional derivative, studied a necessary and sufficient condition for finding strictly minimal solution of an m_1-set optimization problem. For non-smooth set optimization problem, the first use of set relation to define the optimality criteria can be found in [84]. All these aforementioned works use an ordered cone to define the set relation. [48] showed instead the use of variable domination structure to study some optimality criteria for set optimization.

1.2.5.2 Theoretical work on solving SOP

For solving set optimization problems, the literature primarily consist of three frameworks: vectorization, direct scalarization, and indirect scalarization (involving an intermediate vectorization step).

The vectorization strategy was first introduced by Küçük et al. [121]. Löhne and Schrage [160] utilized the vectorization framework to develop the SetOpt algorithm, specifically for problems where the graph of the set-valued objective is polyhedral and convex, with the solution concept based on the lower set-less relation. Additionally, Jahn [95] proposed two vectorization methods for set optimization problems with respect to the set-less order relation. [50] removed the restriction that the set-valued objective must be polyhedral and employed a discretized vectorization strategy to establish an equivalence between the set optimization problem with a corresponding infinite-dimensional multi-objective optimization problem. While this approach relaxes the polyhedral assumption, it still imposes convexity on the set-valued objective. To eliminate this restriction, [49] introduced a vectorizing scheme for solving general non-convex set-optimization problems with the lower set-less order relation.

Compared to vectorization methods, direct scalarization techniques for set optimization have been more extensively studied in the literature for determining optimality conditions. [84] proposed an extension of the Gerstewitz function over sets and derived characterizations of both minimal and weakly minimal solutions with respect to the lower set-less order relation. [188] proposed a new generalized non-linear scalarization method using the oriented distance function for determining minimal and weakly minimal solutions. Karaman et al. [104] introduced an extension of the Gerstewitz function by utilizing set order relations on the family of non-empty bounded sets with the help of the Minkowski difference. [27] characterized set relations using oriented distance functions to examine optimality conditions.

Kobis et al. [117], Ansari et al. [6] introduced non-linear scalarizing methods that use set order relations with variable domination structures to characterize minimal solutions. [35] introduced a generalized form of the oriented distance function to characterize vector-criterion based efficient and weakly efficient solutions.

On indirect scalarization, the only work currently present in the literature is the steepest descent method [18]. This method employs an intermediate vectorization scheme that first converts the set optimization problem into a family of vector optimization problems, which are then ultimately converted into scalar optimization problem.

1.2.5.3 Algorithms for solving set optimization problem

Theoretical investigations have extensively explored optimality conditions for set optimization problems using the set approach. However, there have been relatively few algorithmic and numerical contributions in this area. To the best of our knowledge, there are primarily seven types of numerical algorithms documented in the literature. These are namely *SetOpt* (for polyhedral, convex set optimization problems), *Newton methods*, *derivative-free methods*, *sorting-based methods*, *branch-and-bound method*, and *steepest-descent method*. These can be divided into the three categories of vectorization, direct scalarization, and indirect scalarization (with intermediate vectorization) as follows: These algorithms can be categorized into the earlier mentioned three groups: vectorization, direct scalarization, and indirect scalarization as follows:

1. Vectorization: SetOpt
2. Direct Scalarization: Newton method, Derivative-free methods, Sorting-based methods, and Branch-and-Bound method.
3. Indirect Scalarization: Steepest-descent method.

For the SetOpt algorithm, Lohne, Schrage [160] applied vectorial relaxation to derive an algorithm for a special class of set optimization problems where the objectives are polyhedral and convex.

The Newton method was introduced by Dias and Smirav et al. [40] for solving metrically regular set-valued maps under geometric constraints. Among Derivative-free methods, which depend on the direction of descent and an appropriate step size, Jahn et al. [93] derived these parameters from a finite set of points on a subset of the unit sphere. Kobis et al. [116] adopted a comparable approach but relaxed the assumption of convexity. Jahn et al. [94] expanded these methods to incorporate multiple descent directions at each iteration, thereby forming a tree structure with the initial point as the root and solutions as leaves. Next, Sorting-based methods involve an initial sorting step to filter out non-minimal elements, thereby avoiding pairwise comparisons between every element in the set-valued image space. Kobis et al. [116] extended the forward-backward reduction algorithm developed by Jahn et al. to sort out non-minimal elements of set for vector optimization. Gunther et al. [73, 74] initially established an enumeration of images of the set-valued map whose values are $\preceq$-increasing through scalarization using a strongly monotone functional. Next, the Branch and Bound method described in [47] begins with an initial box containing the feasible set. It then iteratively divides the current active box into two sub-boxes and determines whether to discard or retain these sub-boxes for further iterations. The algorithm terminates when the width of a sub-box falls below a specified precision threshold.

Finally, the Steepest Descent method [18] transforms the set optimization problem into a family of vector optimization problems. It then applies Gerstewitz scalarization to convert these vector optimization problems into scalar form, which is subsequently solved using the Steepest Descent algorithm.

1.2.5.4 Trust-Region method

In the field of numerical optimization methods, Trust-Region algorithms have been extensively researched. Early contributions to Trust-Region methods can be traced back to [51], where Levenberg introduced a modified Gauss-Newton method for solving nonlinear least squares problems. Later, [137] adopted Levenberg's method, contributing to the development of what is now widely recognized as the Levenberg-Marquardt method. The Levenberg-Marquardt method modifies the Gauss-Newton method by incorporating a damping term that limits the size of the Gauss-Newton step. This adjustment prevents the algorithm from making excessively large steps when the Jacobian of the function at the current point is close to singular matrix. From this perspective, the trust region method is also referred to as the restricted step method, a term first introduced by [54]. While trust region methods have not yet been developed for set optimization, there exists a vast body of literature on trust-region methods for both single-objective and multi-objective optimization.

In the context of single-objective optimization, Sun et al. [170] introduced a variant of the trust region method tailored to handle bounded errors in problem solving. Powell [152] employed the trust region algorithm for unconstrained minimization problems without derivatives, constructing linear or quadratic models through interpolation of objective function values. Hoseini et al. [88] introduced a non-smooth trust region algorithm using the Goldstein-subdifferential for non-smooth, non-convex optimization problems. Friedlander et al. [60] introduced a new kind of trust-region algorithm for minimizing a differentiable function of large variables with box constraints. Wang et al. [182] developed a practical trust-region algorithm that utilizes a linear model for solving unconstrained optimization problems. Shi et al. [161] introduced a trust region method for unconstrained optimization problems by adapting the trust-region radius of the associated subproblem at each iteration.

In the context of multi-objective optimization, Qu et al. [153] devised a trust-region algorithm for unconstrained vector optimization problems. Thomann et al. [175] formulated a trust region method tailored for multi-objective heterogeneous optimization problems. Villacorta et al. [180] expanded the conventional scalar trust-region method to identify Pareto critical points in unconstrained multi-objective optimization problems. Carrizo et al. [20] explored the concept of decreasing conditions and introduced the notion of predicted reduction in the context of multi-objective optimization. Finally, to approximate the set of Pareto critical points in multi-objective optimization, Mohammadi et al. introduced a trust region algorithm in [141].

1.2.5.5 Non-Monotone trust-region method

Since trust-region methods have not yet been adapted for set optimization, non-monoton e trust-region (NMTR) methods are also notably absent in the literature. However, given the extensive research on NMTR methods for single-objective and multi-objective problems, these algorithms are highly important. Over the years, there has been extensive research into modifying trust-region methods to incorporate non-monotonic variations. Deng et al. [38] were the first to introduce such modifications for single-objective optimization. Since then, numerous subsequent studies have further explored and evaluated the performance of NMTR methods. For instance, Sun [171], Ahookhosh et al. [2], Mo et al. [140], Maciel et al. [135], Chen et al. [29], among others, have contributed to this area of research.

In the context of multi-objective optimization, non-monotone methods have been successfully applied through their two primary variants: Max-type NMTR and Average-type NMTR. For example, Ramirez et al. [154] modified the trust region method presented in Carrizo et al. [21] to relax the constraint of monotonic decrease in function values. They demonstrated improvements in computation time, number of iterations, and the quantity of subproblems solved. Next, Ding et al. [42] further refined Ramirez's method by introducing an adaptive non-monotone trust-region approach tailored for a specific class of multi-objective non-linear bi-level optimization (MNBLO) problems. Their primary adaptation involved using a convex combination of the current function value and the largest function value from several previous iterations to determine the next step. Advancing further, Ghalavand et al. [61] extended the max-type adaptive non-monotone trust region method in [42] to a broader class of multi-objective optimization problems. They also explored the average-type version of the adaptive trust region method.

1.3 Motivation, objective and contribution of the thesis

The thesis contributes broadly in two categories: theoretical and numerical aspects, addressing both interval and set optimization problems. Under theoretical contributions to interval optimization, we introduce essential concepts designed to handle non-smooth functions: the normal cone, tangent cone, gH-weak subdifferential, gH-Dini-Hadamard subdifferential, and gH-Clarke subdifferential. Under numerical contributions to interval optimization, we introduce a computationally efficient algorithm for solving IOP that builds upon the proposed theoretical concept of the gH-weak subdifferential. In terms of theoretical contributions to set optimization, we introduce the notion of critical points for a particular type of set optimization problem and establish global convergence proofs for our three proposed trust-region-based methods. Lastly, in our numerical contributions to set optimization, we introduce the trust region method for the first time and extend it by developing two non-monotone versions that exhibit improved performance characteristics.

First, we consider the interval optimization problem involving interval-valued functions of interval variables, introducing the concepts of normal and tangent cones. These concepts play a crucial role in establishing optimality conditions for both smooth and non-smooth optimization problems. The tangent cone is useful for expressing constraint qualifications, while the normal cone serves to characterize the constraints themselves. However, the field of interval optimization currently lacks such established concepts, and there has been no significant work in this direction. This motivates our study in this thesis. To the best of our knowledge, our work represents the first attempt to explore normal and tangent cones specifically tailored for interval optimization involving interval variables.

Next, we examine the conditions for efficient solutions in non-smooth and non-convex IOPs. As with classical optimization problems, the existing concept of the gH-subgradient is less effective in characterizing the optimal solution of nonconvex IOPs. This is because the gH-subgradient, which refers to a supporting hyperplane, does not always support the graph of a non-convex interval-valued function. At this point, the weak-subgradient concept can address this issue by utilizing a supporting conic surface to the graph of non-convex functions. Consequently, in this thesis, we propose the weak subgradient concept for interval-valued functions and its application in optimality criteria for non-smooth and non-convex IOPs. Additionally, to demonstrate the advantages of our proposed gH-weak subgradient, we introduce a new algorithm based on this concept to solve IOPs, which is easy to implement and more efficient in terms of computational complexity.

Next, we develop the concept of the gH-Dini-Hadamard subdifferential. Since the concept of subdifferential is closely connected to directional derivatives, the first step in generalizing a subdifferential is to generalize the directional derivative. One such generalized directional derivative is the Dini-Hadamard derivative, which is particularly used to identify the monotonic nature of lower-semicontinuous functions. However, the Dini-Hadamard derivative does not adhere to common algebraic rules (e.g., sum rule, product rule, chain rule) that are essential for deriving optimality conditions for lower-semicontinuous functions. The Dini-Hadamard subdifferential can be useful in overcoming this problem. Therefore, one of the primary motivation of this thesis is to improve the concept of the Dini-Hadamard subdifferential for IOPs involving gH-lower-semicontinuous interval-valued functions, as well as non-convex and non-smooth interval-valued functions.

While subdifferentials are used for convex functions, generalized subdifferentials are used for functions exhibiting generalized convexity. One such generalized subdifferential is the Clarke subdifferential, which plays a crucial role in duality theory. For IOPs involving either gH-Lipschitz IVFs or gH-lower semicontinuous IVFs, the Clarke subdifferential concept may help characterize optimality conditions under the assumption of generalized convexity. Therefore, one of the objectives of this thesis is to enhance the Clarke subdifferential for characterizing optimal solutions in interval-valued duality theory.

Beyond theoretical advancements, it is crucial that numerical developments in terms of algorithms keep pace with theory. Interval optimization can be viewed as a subset of the broader and more robust framework of set optimization, as intervals are connected and ordered subsets of $\mathbb{R}$. Therefore, developing numerical algorithms for set optimization is more beneficial, as these algorithms will inherently apply to interval optimization as well. Algorithms designed exclusively for interval optimization cannot be generalized to other mathematical domains. Consequently, in the second part of this thesis, we shift the focus to set optimization and develop numerical algorithms for it. Although set optimization is more impactful, it also presents greater analytical challenges. Therefore, as a preliminary step, we address a specific type of set optimization problem and offer numerical advancements, laying the groundwork for tackling more general and abstract set optimization problems in the future.

The existing literature on numerical methods for set optimization is very limited. Only a few methods are available, such as SetOpt, Newton method, derivative-free method, sorting-based and branch-and-bound methods, and steepest-descent method. However, these methods often suffer from various drawbacks, including restrictions on the type of objective function, assumptions about metric regularity, limitations on the feasible set, slow convergence rates, and a lack of global convergence (see Chapter 6 for more details). In this context, trust region methods present a promising class of algorithms capable of addressing these issues, as demonstrated in single-objective and multi-objective optimization literature. Consequently, one of the main objectives of this thesis is to develop and analyze a trust region-based algorithm for set optimization problems. Additionally, since the trust region method identifies critical points rather than optimal points, it is essential to define the concept of a critical point for set optimization, which is currently absent in the literature. Therefore, our work is the first to establish the concept of critical point for set optimization.

Finally, we point out that the trust region method has a monotonic restriction that may hinder its optimal performance. Drawing inspiration from single and multi-objective optimization literature, we incorporate non-monotone modifications to develop two additional trust region algorithms for set optimization. These are termed the max-type and average-type non-monotone trust region methods. Through extensive numerical experiments and performance profiling, we demonstrate that these new methods outperform the original trust region method.

1.4 Organization of the thesis

The thesis is composed of eight chapters, including an introductory chapter and a concluding chapter that discusses future research possibilities. The introductory chapter provides a thorough review of relevant literature. The overall structure of the thesis is as follows.

Chapter 2 delves into the concept of the normal cone for a set of intervals and examines its fundamental properties. It explores the concepts of tangent cones, their characteristics, and their characterization using the distance function of interval variables. The dual relationship between normal and tangent cones is also investigated. Additionally, a Lagrangian multiplier approach for constrained IOPs is introduced. The chapter concludes with an example illustrating the practical application of a normal cone for a set of intervals in support vector machine type IOPs.

Chapter 3 introduces the concept of gH-weak subdifferential for IVFs and discusses their properties, including convexity, closedness, and non-emptiness. The significance of gH-weak subdifferential in deriving the necessary condition for weak efficiency for gH-weak subdifferentiable IVFs is highlighted. The necessary condition for obtaining an efficient solution of the difference between two IVFs is examined. A sup-relation between gH-directional derivative and gH-weak subgradients is defined. Utilizing this relation, a $\mathcal{W}$-gH-weak subgradient method for finding a weak efficient solution to an unconstrained IOP is presented, along with its algorithmic implementation and convergence analysis.

Chapter 4 focuses on gH-Dini Hadamard subdifferential and superdifferential for IVFs. It explores the relationship between these concepts and examines key properties such as convexity and closedness. The chapter provides several calculus rules, including the partial chain rule and subadditive rule, as well as a smooth variational type description of gH-Dini Hadamard subdifferential. An FJ-type and a KKT-type necessary optimality criterion for constrained IOPs with gH-Dini Hadamard superdifferentiable objectives and gH-Dini Hadamard subdifferentiable constraints are derived. Additionally, the gH-Dini Hadamard subdifferential is applied to analyze the asymptotic controllability of non-smooth interval dynamical control systems.

Chapter 5 introduces the concept of gH-Clarke subdifferential for IVFs and explores generalizations of strong convexity for IVFs. It establishes sufficient optimality conditions for a specific class of non-smooth single-objective optimization problems involving support functions of compact convex sets. The chapter formulates the dual problem and provides weak and strong duality theorems to ensure strict efficiency.

Chapter 6 introduces the concept of critical points and descent directions for set-valued maps, and proposes a trust region method for set optimization. The proposed approach is described in detail, including a step-by-step presentation of Algorithm 3 with explanations of each step. The chapter examines the convergence of the proposed trust-region method in terms of global optimality. Numerical experiments are conducted, and results from running Algorithm 3 on several test cases are presented.

Chapter 7 introduces two non-monotone schemes, Max-NTRM and Avg-NTR, and presents their respective pseudo-codes in Algorithm 4 and Algorithm 5. Through extensive numerical experiments and performance profiling, the improvements offered by the non-monotone trust region methods are verified.

Finally, Chapter 8 concludes this study by summarizing the final thoughts from the perspectives of interval and set optimization, and suggests potential directions for future research.

****