Chapter 1

Introduction

Optimization encompasses a range of mathematical principles and techniques to find the maxima/minima of a function or a criterion over a set of constraints. Classical optimization is featured by a vast collection of useful algorithms and software. It has become an essential tool for robust modelling, optimal design, industrial operations, management, finding the trajectory of a rocket, and many more fields. Optimization theory and methods have numerous applications in various fields, such as computational mathematics, science, engineering, business management, space technology, etc. The process involves studying the problem's optimality conditions, constructing the problem's models, determining the algorithmic methods for finding the solution, establishing the convergence theory of the algorithms, and constructing numerical experiments involving typical and real-life problems.

The statement of optimization models in mathematics consists of three key components: decision variables, objective function, and constraints.

- *Decision variables* represent values that can vary within the context of a given optimization problem.

- *Constraints* constitute the permissible values or logical conditions for the variables in an optimization problem.

- An *objective function* defines the primary criterion of the problem, which has to be either minimized or maximized over a set of feasible alternatives.

In an optimization problem, the nature of the mathematical relationships among decision variables, the objective function, and constraints influences the complexity of the problem and determines the suitable algorithms to find the optimum solution to the given optimization problem. Despite the availability of numerous theories and tools to find the optimal solutions, it is not always possible to represent real-world situations with classical mathematics. This is due to uncertainty in collected data, measurement errors or random events. Sometimes, the data may be roughly estimated, or it is more appropriate to assume that it belongs to some set or interval. The imperfection or uncertainty stems from two main causes: imprecision and randomness.

In this thesis, we aim to address two types of optimization problems: Interval-valued optimization problems and set-valued optimization problems. We develop theories for smooth and nonsmooth interval-valued functions and derive the optimality conditions for interval optimization problems. After that, we propose suitable methods to find efficient solutions to set-valued optimization problems.

1.1 Interval Analysis

Digital computers have limited numerical representation capabilities and can operate with rounded floating-point numbers only. However, some numbers have infinite digits in their numerical representation, such as π and $\sqrt{2}$, and most rational numbers are represented in a rounded form. In 1611, Kepler proposed an important conjecture stating that close packing, whether cubic or hexagonal, is the densest possible sphere packing, with both having maximum densities of approximately $\frac{\pi}{3\sqrt{2}}$ or about 74.048%. The proof of this conjecture has the data in the form of intervals and relies extensively on methods from interval arithmetic, linear programming, and global optimization. To address such work, it is essential to understand real interval spaces and interval analysis. The set of real numbers, $\mathbb{R}$, is an unbounded, totally ordered set with the usual structure, allowing any number to be bounded by two adjacent numbers. Mathematical analysis with intervals involves exploring the algebraic and analytical structures of spaces associated with these intervals.

Interval analysis is a growing branch of applied mathematics pioneered by Moore in 1966 [2, 9]. Methods of interval analysis contain both ordinary machine arithmetic results and infinite precision arithmetic results. This provides a general mechanism for bounding the accumulation of roundoff errors in any machine computation. In real-world engineering fields, many dynamic problems can be formulated by dynamic models, such as motor servo systems, navigation control, and so forth. However, this system usually involves multiple uncertain parameters or interval coefficients. The purpose of using intervals is to provide upper and lower bounds of the parameters of such mathematical problems. To deal with such uncertainty of given data, these problems are often modelled by optimization problems with interval-valued objective functions. The optimization problem with interval-valued functions (IVFs) is known as interval optimization problems (IOPs).

1.2 Interval-Valued Function and Interval Optimization Problem

An interval-valued function is a mathematical function of one or more variables whose domain is a subset of Euclidean space $\mathbb{R}^n$ and range is a set of closed and bounded intervals of a real number denoted by $I(\mathbb{R})$. For each argument point $x \in \mathcal{X}$, where $\mathcal{X}$ is a nonempty subset of $\mathbb{R}^n$ and an IVF $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$ is presented by

$$\mathbf{F}(x) = [\underline{f}(x), \overline{f}(x)],\ x \in \mathcal{X},$$

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

In real-world problems, the data collected by the decision-makers are always assumed to be real numbers with a certain value. In this case, the objective function of optimization problems is a real-valued function. However, there are some optimization problems where the objective may be uncertain due to inexact data. For example, suppose that a factory can produce two goods, say G_1 and G_2, in input quantities x_1 and x_2, subject to budget constraint $S \subseteq \mathbb{R}^2$. For selling the goods G_1 and G_2 in the market, we assume that the factory can earn c_1 and c_2 dollars per units, respectively. In this case, the purpose is to maximize the objective function $c_1x_1 + c_2x_2$ subject to the budget constraint set $S \subseteq \mathbb{R}^2$. However, the prices of goods may fluctuate from time to time in the financial market. It seems more reasonable to assume the prices to be uncertain quantities. There are three kinds of methodologies that can model uncertain quantities: random variables, fuzzy numbers and intervals. If the coefficients c_1 and c_2 are assumed to be random variables, then the problem becomes a stochastic optimization problem. Birge and Louveaux in 1997 [10] explained the mainstream of stochastic optimization problems and also introduced some useful methods to solve these optimization problems. If the coefficients c_1 and c_2 are assumed to be fuzzy, then the problem becomes a fuzzy optimization problem. Stochastic and fuzzy optimization problems are not easy to solve. The usual way is to transform them into conventional optimization problems; frequently, these problems are very complicated. On the other hand, if we assume the coefficients c_1 and c_2 are assumed to be compact intervals of real numbers, then although the prices may fluctuate from time to time, we can always make sure that the prices will fall into the corresponding intervals, in this case, problem becomes an interval optimization problem (IOP). Under this assumption, the optimization problem will be easier to be solved than a stochastic or fuzzy optimization problem. As we know that in most of the cases, the coefficients of the objective function in the stochastic optimization problems are considered as random variables with known distributions. However, the specifications of the distributions are very subjective. For example, many researchers invoke the Gaussian (normal) distributions with different parameters in stochastic optimization problems. These specifications do not completely match the real problems. Therefore, interval-valued optimization problems may provide an alternative choice to consider the uncertainty in the optimization problems. That is to say, the coefficients of the objective function in the interval-valued optimization problems are considered as compact intervals. Although the specifications of compact intervals may still be judged as subjective viewpoint, we can argue that the bounds of uncertain data (i.e., determining the compact intervals which give the upper and lower bounds of the possible observed data) are easier to be handled than specifying the Gaussian distributions in stochastic optimization problems. Mathematically, an interval optimization problem is presented by can argue that the bounds of uncertain data (i.e., determining the compact intervals which give the upper and lower bounds of the possible observed data) are easier to be handled than specifying the Gaussian distributions in stochastic optimization problems. Mathematically, an interval optimization problem is presented by

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

where $\mathcal{X} \subset \mathbb{R}^n$ and $\mathbf{F} : \mathcal{X} \to I(\mathbb{R})$. In the rigorous study on IOPs, we often need the following notions.

- Decision space: A decision is characterized by the DM's choice between different possible courses of action, called alternatives. Set of all alternatives constitute the set $\mathcal{X}$ which is called as a feasible set. An alternative can be defined by a vector of real numbers $x = (x_1, x_2, \ldots, x_n)^\top$. A vector for representing an alternative is said to be a decision vector. The components of this vector are called decision variables. Each decision variable is related to a particular aspect of the alternatives. The space of the decision vectors is called as decision space.

The set $\mathcal{X}$ is also known as the decision feasible region. A point x in $\mathcal{X}$ is known as a decision feasible point.

- Objective space: For any point x in the decision feasible region X, an interval $\mathbf{F}(x)$ in $I(\mathbb{R})$ is obtained. The interval space in which the points $\mathbf{F}(x)$ lie is known as objective space. The set of all intervals $\mathbf{F}(x)$ where x in $\mathcal{X}$ is known as a feasible set in the objective space or objective feasible region. In solving IOPs, DM is always more interested in the objective space than the decision space because DM is often more interested in the objective values. We note that the image of the feasible set $\mathcal{X}$ under the interval-valued function $\mathbf{F}$ is the feasible set in the objective space. In IOPs, the interval-valued objective function can be observed as a bunch of infinitely many real-valued objective functions. So, optimal solutions of IOPs behave like optimal solutions of multi-objective optimization problems. Unlike a single objective problem, there may not exist a unique solution of an IOP, since otherwise, there does not arise any conflict among all infinitely many real objectives of IOP, and it loses the essence of IOP. The basic difference between IOPs and single objective optimization problems is that the feasible region in the objective space of a single objective optimization problem is a totally ordered subset of the real line, whereas an IOP constitutes an infinite dimensional objective feasible region, which is not a totally ordered set in general. Thus, all solutions can be completely ordered according to the objective function in single objective optimization problems. In contrast, for an IOP, the solutions can only be ordered partially. As a consequence, in the case of single-objective problems, only one global optimum exists; but in the case of IOPs, conflicting real objectives can cause a situation where no solution is superior to the others. Thus, usually, there are many solutions to an IOP. The feasible solutions which can be improved without causing simultaneous deterioration in at least one criterion can not certainly be the optimal solution of the considered IOP. This concept leads to the foundation of non-dominated solutions.

- Non-dominated solution: A non-dominated solution of IOP is a feasible point of interval objective space where any improvement in one criterion in a bunch of infinitely many real objectives of IOP can take place only through the worsening of at least one another criterion in that infinitely many real objectives. The concept of nondominated solution of IOP is of primordial importance to recognize the conflicting nature of a bunch of infinitely many criteria of IOP. For a non-dominated point of IOP, there is no other feasible point in interval objective space that makes every criterion strictly better off. Each non-dominated point is equally acceptable as a solution to the IOP. Non-dominated point determines efficient solution from the entire feasible region

1.3 Set-Valued Analysis

This class of mathematical problems deals with the minimization of set-valued mappings acting between two normed spaces, where the image space is partially ordered by a closed, convex, and pointed cone. These problems have various applications in finance [11, 12], robotics, socio-economics [13, 14], and robust multiobjective decision-making [15]. The solution concept is based on comparing sets with respect to a binary relation (usually a preorder) defined on the power set of the image space and finding the minimal solution accordingly. The optimization problems acting between two sets are known as set-valued optimization problems (SOPs).

The study of set-valued functions in optimization has evolved significantly with the contribution of various key researchers. Clarke's seminal work on generalized gradients laid the foundation for the analysis of nonsmooth and set-valued optimization problems [16, 17]. After that, Aubin and Ekeland [18] expanded these ideas by integrating measurable multifunctions within the convex analysis scenarios which helped in the stability and sensitivity analysis of optimization problems. Moreover, the comprehensive work by Rockafellar in [19, 20] extended these developments by providing a unified framework for understanding the role of set-valued functions across various optimization contexts. These contributions have collectively shaped the modern landscape of set-valued optimization and established it as a critical area of study with applications across diverse fields.

1.4 Set-Valued Function

A general statement of a set-valued optimization problem is given by

$$\text{Minimize } F(x) \;\text{ subject to }\; x \in X,$$

where F is a set-valued map from a nonempty subset X of $\mathbb{R}^n$ to $\mathbb{R}^m$. For the solution concepts of set optimization problems, there are two main approaches, namely the vector approach [21, 22] and the set approach [21, 23]. In the vector approach, the decision maker's preference is based on comparing the vectors in the image set $F(X)$. Thus, in this case, an optimal set is selected by identifying just one of its elements, without taking into account the rest of the set. An alternative definition of solutions of set optimization problems was studied [21, 23] via comparing the sets $F(x)$ with respect to a binary relation (usually a preorder) for all $x \in X$.

1.5 Set-Valued Optimization Problem

In real-life situations, set optimization accounts for multiple criteria in their choices. For example, a company needs to optimize its supply chain, deciding how much inventory to hold at different locations. Since the demand and supply levels can fluctuate, so instead of using fixed numbers, the company can use set-valued data representing a range of possible demand and supply. Set optimization helps in finding strategies that minimize cost while ensuring service levels under all possible scenarios. We discuss an example that demonstrates the natural occurrences of set-valued optimization problems. We begin with the following example emerging in game theory:

We discuss a vector-valued two-person game. Let M and N be nonempty sets, and Y be a linear space and let $F : M \times N \to Y$ be a single-valued map. Suppose that Player 1 chooses $a \in M$, and Player 2 chooses $b \in N$. Then, $f(a, b)$ describes the loss for Player 1. Consider a set-valued map $F : M \rightrightarrows Y$ defined by

$$F(a) := \{F(a, b) \mid b \in N\}.$$

Player 2 is said to be cooperative towards Player 1 and look for $\preceq^l_K$-minimal solutions $\bar{a} \in M$ in the sense of Definition 1.32. Then, the vector-valued game can be formulated as a set-valued optimization problem and aim to solve the following SOP:

$$\preceq^l_K -\text{minimize } F(a), \text{ subject to } a \in M.$$

When Player 2 is cooperative, then the solution is obtained by looking for $\preceq^u_K$-minimal solutions $\bar{a} \in M$ in the sense of Definition 1.32. Then, the vector-valued game can be formulated as a set-valued optimization problem and aim to solve the following SOP:

$$\preceq^u_K -\text{minimize } F(a), \text{ subject to } a \in M.$$

The concept of finding minimizes of set-valued maps is given in Definition 1.33. To achieve a more refined solution to a set-valued optimization problem, the concept of comparison among the values of objective maps has been given. Various applications of set-valued approaches are given in the work by Fernandez et. al [24] and references therein.

1.6 Literature Survey

1.6.1 Literature on Interval Analysis and Calculus of Interval-Valued Functions

In the literature of interval analysis, Sunaga in 1958 [25], and Moore in 1966 [2] independently developed interval arithmetic. Although Sunaga were the first to create interval arithmetic, the recognition of interval analysis was given by the book of Moore on interval analysis 1966 [2]. Further, basic contributions in interval analysis are given by Apostolatos and Kulisch in 1967 [26], Hansen in 1965 [27], Kruckeberg in 1966 [28], Nickel in 1966 [29], and others. By using Moore's interval arithmetic, Mayer in 1970 [30] described the concept of quasilinear space for compact intervals. In Moore's interval arithmetic, there are a few limitations (see [5] for details), such as, the additive inverse of a degenerate intervals, i.e., an interval whose upper and lower limits are equal, exist only. For the same reason, many conventional properties for real numbers are not true for compact intervals, for instance, for two compact intervals $\mathbf{A}$ and $\mathbf{C}$, $(\mathbf{A} \ominus_{gH} \mathbf{C}) \oplus \mathbf{C} \neq \mathbf{A}$ (see Remark 2.3.1 of [5]). Thus, to develop a theoretical framework of the calculus of interval analysis, Hukuhara in 1967 [31] introduced a new concept for the difference of compact intervals, known as Hukuhara difference (H-difference) of intervals and derive several properties of compact intervals by using this difference. Although H-difference provides the additive inverse of compact intervals, this difference of a compact interval B from a compact interval A can be calculated only when the width of A is greater than or equal to that of B (see details in [32]). To overcome this difficulty, the nonstandard difference of intervals is introduced by Markov in 1979 [33], and for the same reason, Stefanini in 2008 [34] introduced a strong concept of difference of two intervals as generalized Hukuhara difference (gH-difference) of intervals. The gH-difference provides an additive inverse of any compact interval and is applicable for all pairs of compact intervals. By using the gH-difference of intervals and Moore's interval arithmetic, Stefanini proved the cancellation law for the addition of intervals and the distributive law for subtraction of intervals by the scalar. Further, in [35], with the help of norm of an interval defined by Moore and gH-difference of intervals, it is explained that set of compact intervals is quasi normed linear space. Recently, to generalize the concepts of smooth and nonsmooth analysis for interval-valued function and to derive interval variational inequalities, Ghosh in 2019 [5] and proved some inequalities of intervals by using dominance relation and gH-difference of intervals. Apart from Moore's interval arithmetic, another concept of interval arithmetic has been developed by Piegat and Landowski [36], namely RDM interval arithmetic, which also ensures the existence of an additive inverse for any compact interval. Generally, all the properties of RDM interval arithmetic are similar to Moore's interval arithmetic except the subtraction of an interval from itself. In this thesis, we use Moore's interval arithmetic with gH-difference instead of RDM interval arithmetic (see Note 2 in [5] for the reason).

To observe the properties of an IVF, calculus plays an essential role. Initially, to develop the calculus of IVFs, Hukuhara in 1967 [31] introduced the concept of differentiability of IVFs with the help of H-difference of intervals. However, the definition of Hukuhara differentiability (H-differentiability) is found to be restrictive (see [32]). To remove the deficiencies of H-differentiability, Bede and Gal in 2005 [34] defined strongly generalized derivative (G-derivative) for IVFs and derived a Newton-Leibnitz-type formula. In order to formulate the mean-value theorem for IVFs, Markov in 1979 [33] introduced a new concept of difference of intervals and defined differentiability of IVFs by using this difference. In 2009 [34], Stefanini and Bede defined the generalized Hukuhara differentiability (gH-differentiability) of IVFs by using the concept of generalized Hukuhara difference. In defining the calculus of IVFs, the concepts of gH-derivative, gH-partial derivative, gH-gradient, and gH-differentiability for IVFs have been developed in [34, 37].

To derive a Karush-Kuhn-Tucker (KKT) condition for IOPs, Guo et al. in 2019 [38] defined gH-symmetric derivative for IVFs. Ghosh in 2016 [39], analyzed the notion of gH-differentiability of multi-variable IVFs to propose the Newton method for IOPs. The concept of second-order differentiability of IVFs is introduced by Van [40] to study the existence of a unique solution of interval differential equations. Lupulescu [35] defined delta generalized Hukuhara differentiability on time scales by using gH-difference. Chalco et al. [32] introduced the concept of gH-derivative for IVFs that generalizes Hukuhara derivative and G-derivative, and proved that this derivative is equivalent to gH-derivative. In [34], Stefanini and Bede defined level wise gH-differentiability and generalized fuzzy differentiability by LU-parametric representation for fuzzy-valued functions. Kalani et al. [41] analyzed the concept of interval-valued fuzzy derivative for perfect and semi-perfect interval-valued fuzzy mappings to derive a method for solving interval-valued fuzzy differential equations using the extension principle. Recently, Ghosh et al. [5] have provided the idea of gH-directional derivative, gH-Gateaux derivative, and gH-Frechet derivative of IVFs to derive the optimality conditions for IOPs.

1.6.2 Literature on Interval Optimization

In recent years, the interval analysis method was developed to model the uncertainty in uncertain optimization problems, in which the bounds of the uncertain coefficients are only required, not necessarily knowing the probability distributions or membership functions. Tanaka et al. in 1984 [42] discussed the linear programming problem with interval coefficients in the objective function. Chanas and Kuchta in 1996 [43, 44] suggested an approach based on an order relation of interval number to convert the linear optimization problem with uncertainty into a deterministic optimization problem. Liu and Da in 1999 [45] proposed an interval number optimization method based on a fuzzy constraint satisfactory degree to deal with linear problems. Sengupta et al. in 2001 [46] studied the linear interval number programming problems in which the coefficients of the objective function and inequality constraints are all interval numbers.

They proposed the concept of the "acceptability index" and gave one solution for the uncertain linear programming. Zhang et al. in 1999 [47] assumed interval numbers as random variables with uniform distributions and constructed a possibility degree to solve the multi-criteria decision problem. The above methods point out a new way for uncertain optimization. Wu [48], used the concept of Hukuhara differentiability to study KKT conditions of optimization problems with IVFs. Wu [49], has also illustrated the solution concept of optimization problems with IVFs by imposing a partial ordering on the set of all closed intervals and applying the existing calculus of IVFs. In 2013, the KKT conditions, based on gH-differentiability, of optimization problems with IVFs have been illustrated by Chalco-Cano and others [50]. After that, Bhurji and Panda [51] have defined interval-valued function in the parametric form and studied its properties, and developed a methodology to study the existence of the efficient solution of an optimization problem with IVFs. Ghosh [37] has introduced a new definition of gH-differentiability and proposed a Newton method [37] and an updated Newton method [52] to capture the efficient solution of an optimization problem with IVFs. Recently, Ghosh et al. [53] have proposed the theory of alternatives and hence the KKT optimality condition for IOPs. Importantly, it is shown in [53] that KKT optimality conditions appear with the inclusion relations instead of equations.

1.6.3 Literature on Set-Valued Analysis and Set Optimization

Now, we discuss the literature on set-valued analysis. There are mainly two approaches for defining the optimal solutions of a set optimization problem; these are the vector approach and the set approach. In the vector approach, one looks for the minimal points in the image set of the set-valued objective mapping [15]. Thus, in this case, an optimal set is selected by identifying just one of its elements, without taking into account the rest of the set. However, in some practical scenarios, this feature poses an important drawback from the modelling point of view. The set approach is an important attempt to overcome this problem. The idea lies in comparing sets with respect to a binary relation defined on a power set of the image space and finding the minimal solution. To the best of our knowledge, the first of these set relations were introduced by Young [54] and Nishnianidze [55]. After that, Kuroiwa modified the concept in [56] and [57]. Recently, Jahn and Ha [15], and Karaman et al. [58] have defined new set order relations. After that, set optimization has been developed in various topics such as existence of solutions [15, 57, 59–61], well-posedness [62–69], optimality conditions and algorithms [21, 70–72]. Several other references are given in [11, 73, 74].

Scalarization techniques are a fundamental tool in vector and set optimization from theoretical as well as computational point of view. In this technique, the set or vector optimization problem is reduced to a parametric family of scalarization problems. Scalarization problems are formed by minimizing the composition of so-called scalarizing functional with set-valued objective mapping of the set optimization problem along with some additional constraints. In vector optimization, there are several scalarization techniques discussed in the literature, see [75]. Many of these have been extended to the set-valued context. Several generalizations of nonlinear separating functionals have been discussed in [76,77]. In this thesis, we propose a Newton-type and a quasi-Newton type method for a particular class of set optimization problems.

The research using the solution concept of the set approach was started with the works in [23, 57, 78], which considers preorder relations for comparing sets. A detailed discussion on this field is given in [79]. The existing methods in the literature to solve set optimization problems fall into one of the following groups:

- Algorithms *based on scalarization* have been discussed in [80, 81]. The methods proposed in these papers address a particular class of set-valued mappings characterized by robust counterparts of vector optimization problems. In [80, 81], a linear scalarization technique was utilized to analyse the optimistic solution of set optimization problem and extended the ϵ-constraint method for the ordering cones with nonnegative orthant to deal with the set optimization problem.

- Algorithms of *sorting type* have been discussed in [82–85]. These methods deal with set optimization problems having a finite feasible set and rely completely on comparing the images of the set-valued objective mappings. In [84, 85], Köbis et al., extended the work of Jahn given in [15, 86] for vector optimization problems. They employ the forward and backward reduction procedures to the algorithms of [15, 86], which effectively reduce the number of comparisons compared to the native implementations where every pair of sets needs to be compared. After that, an extension of the algorithm developed by Günther and Popovici [87] for vector problems has been discussed in [82, 83]. The approach involves first finding an enumeration of the images of the set-valued mapping, ensuring that their values, when scalarized using a strongly monotone functional, are in an increasing order. This is followed by a forward iteration procedure.

- Eichfelder et al. [88] proposed a *branch and bound* technique for solving multiobjective optimization problems with decision uncertainty based on set-order relation to handling the uncertainties.

- Methods based on derivative-free strategy from [89]. Jahn [90,91] reported *derivative free* descent methods for set optimization using the concepts from [89]. These algorithms are designed to handle unconstrained problems without assuming any particular structure of the set-valued mapping. In [90], the method focuses on scenarios where both the epigraphical and typographical multifunction of the set-valued objective mapping exhibit convex values. After that, a relaxation of the convexity assumption was explored in [91] for upper set less relation. In this, several directions are chosen at once instead of only one. These methods follow a tree generation concept, with roots as the initial point and leaves as the possible solutions. This method is called the rooted tree method.

- Bouza et al. [1] introduced the study of conventional gradient-based classical approaches (starting with the steepest descent method) to solve set optimization problems with finite cardinality.

Recently, Bouza et al. [1] pioneered generalizing classical gradient-based algorithm (started with the steepest descent method) for set optimization problems. This steepest descent method is known to have a linear convergence rate.

1.7 Preliminaries

In this section, we have discussed the basic definitions and fundamental results on of interval-valued and set-valued functions. The discussion is presented in the following subsections.

1.7.1 Interval Arithmetic and Some Basic Properties on Intervals

In this subsection, we discuss Moore's interval arithmetic [2] followed by the concepts of gH-difference of two intervals [34] and ordering of intervals [9]. Throughout the thesis, we denote the elements of $I(\mathbb{R})$ by bold capital letters: $\mathbf{A}, \mathbf{B}, \mathbf{C}, \ldots$. An element $\mathbf{A}$ of $I(\mathbb{R})$ is represented by the corresponding small letter: $\mathbf{A} = [\underline{a}, \overline{a}]$, where $\underline{a}$ and $\overline{a}$ are in $\mathbb{R}$ and $\underline{a} \leq \overline{a}$.

Let $\mathbf{A}, \mathbf{B} \in I(\mathbb{R})$ and $\mu \in \mathbb{R}$. The addition of $\mathbf{A}$ and $\mathbf{B}$ is denoted by $\mathbf{A} \oplus \mathbf{B}$, is defined by

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

The subtraction of $\mathbf{B}$ from $\mathbf{A}$, denoted $\mathbf{A} \ominus \mathbf{B}$, is defined by

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

The multiplication of $\mathbf{A}$ and $\mathbf{B}$ is denoted by $\mathbf{A} \odot \mathbf{B}$, is defined by

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

The multiplication by a real number μ to $\mathbf{A}$, denoted $\mu \odot \mathbf{A}$ or $\mathbf{A} \odot \mu$, is defined by

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

Let $0 \notin \mathbf{B}$. The division of $\mathbf{A}$ by $\mathbf{B}$ is denoted by $\mathbf{A} \oslash \mathbf{B}$ defined by

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

Next, we define the following definition for the difference between a pair of intervals since it is the most general definition of difference (see [37]).

Definition 1.1 (gH-difference of intervals) [34]). *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be two elements of $I(\mathbb{R})$. The gH-difference between $\boldsymbol{A}$ and $\boldsymbol{B}$ is defined as the interval $\boldsymbol{C}$ such that*

$$\boldsymbol{C} = \boldsymbol{A} \ominus_{gH} \boldsymbol{B} \Longleftrightarrow \begin{cases} \boldsymbol{A} = \boldsymbol{B} \oplus \boldsymbol{C} \\ or \\ \boldsymbol{B} = \boldsymbol{A} \ominus \boldsymbol{C}. \end{cases}$$

For $\mathbf{A} = [\underline{a}, \overline{a}]$ and $\mathbf{B} = [\underline{b}, \overline{b}]$, $\mathbf{A} \ominus_{gH} \mathbf{B}$ is given by

$$\mathbf{A} \ominus_{gH} \mathbf{B} = [\min\{\underline{a} - \underline{b}, \overline{a} - \overline{b}\}, \max\{\underline{a} - \underline{b}, \overline{a} - \overline{b}\}].$$

Definition 1.2 (Dominance of intervals) [49]). *Let $\boldsymbol{A} = [\underline{a}, \overline{a}]$ and $\boldsymbol{B} = [\underline{b}, \overline{b}]$ be two elements of $I(\mathbb{R})$. Then,*

(i) $\boldsymbol{B}$ is said to be dominated by $\boldsymbol{A}$ if $\underline{a} \leq \underline{b}$ and $\overline{a} \leq \overline{b}$, and then we write $\boldsymbol{A} \preceq \boldsymbol{B}$;

(ii) $\boldsymbol{B}$ *is said to be strictly dominated by* $\boldsymbol{A}$ *if* $\boldsymbol{A} \preceq \boldsymbol{B}$ *and* $\boldsymbol{A} \neq \boldsymbol{B}$*, and then we write* $\boldsymbol{A} \prec \boldsymbol{B}$*. Equivalently,* $\boldsymbol{A} \prec \boldsymbol{B}$ *if and only if any of the following 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 $\boldsymbol{A} \preceq \boldsymbol{B}$ *nor* $\boldsymbol{B} \preceq \boldsymbol{A}$*, we say that none of* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *dominates the other, or* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are not comparable. Equivalently,* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are not comparable if either '*$\underline{a} < \underline{b}$ *and* $\overline{a} > \overline{b}$*' or '*$\underline{a} > \underline{b}$ *and* $\overline{a} < \overline{b}$*';*

(iv) $\boldsymbol{B}$ *is said to be not dominated by* $\boldsymbol{A}$ *if either* $\boldsymbol{B} \preceq \boldsymbol{A}$ *or* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are not comparable, and then we write* $\boldsymbol{A} \nprec \boldsymbol{B}$*. Similarly, a real number* a *is said to be not dominated by* $\boldsymbol{A}$ *if either* $[a, a] \preceq \boldsymbol{A}$ *or* $\boldsymbol{A}$ *and* $[a, a]$ *are not comparable, and then we write* $\boldsymbol{A} \nprec a$*.*

The algebraic operations on the product space $I(\mathbb{R})^n = I(\mathbb{R}) \times I(\mathbb{R}) \times \cdots \times I(\mathbb{R})$ (n times) are defined as follows. For two elements $\widehat{\mathbf{C}} = (\mathbf{C}_1, \mathbf{C}_2, \ldots, \mathbf{C}_n)^\top$ and $\widehat{\mathbf{D}} = (\mathbf{D}_1, \mathbf{D}_2, \ldots, \mathbf{D}_n)^\top$ of $I(\mathbb{R})^n$, the operation $\widehat{\mathbf{C}} \star \widehat{\mathbf{D}}$ is defined by

$$\widehat{\mathbf{C}} \star \widehat{\mathbf{D}} = (\mathbf{C}_1 \star \mathbf{D}_1, \mathbf{C}_2 \star \mathbf{D}_2, \ldots, \mathbf{C}_n \star \mathbf{D}_n)^\top,$$

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

Remark 1.1 *For any two elements* $\widehat{\boldsymbol{C}} = (\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_n)^\top$ *and* $\widehat{\boldsymbol{D}} = (\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots, \boldsymbol{D}_n)^\top$ *in* $I(\mathbb{R})^n$*,*

$$\widehat{\boldsymbol{C}} \preceq \widehat{\boldsymbol{D}} \iff \boldsymbol{C}_j \preceq \boldsymbol{D}_j \textit{ for all } j = 1, 2, \ldots, n.$$

Definition 1.3 (Maximum and minimum of intervals). *Let* $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_p$ *be elements of* $I(\mathbb{R})$ *with* $\boldsymbol{Z}_1 \preceq \boldsymbol{Z}_2 \preceq \cdots \preceq \boldsymbol{Z}_p$*. Then,*

$$\max\{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_p\} = \boldsymbol{Z}_p \textit{ and } \min\{\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_p\} = \boldsymbol{Z}_1.$$

Definition 1.4 (Norm on $I(\mathbb{R})$ and $I(\mathbb{R})^n$ [2]). *Let* $\boldsymbol{Y} = [\underline{y}, \overline{y}]$ *and* $\widehat{\boldsymbol{Y}} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_n)^\top$ *be generic elements in* $I(\mathbb{R})$ *and* $I(\mathbb{R})^n$*, respectively. The following two functions* $\|\cdot\|_{I(\mathbb{R})} : I(\mathbb{R}) \to \mathbb{R}_+$ *and* $\|\cdot\|_{I(\mathbb{R})^n} : I(\mathbb{R})^n \to \mathbb{R}_+$ *are referred to norm on* $I(\mathbb{R})$ *and* $I(\mathbb{R})^n$*, respectively:*

$$\|\boldsymbol{Y}\|_{I(\mathbb{R})} = \max\{|\underline{y}|, |\bar{y}|\} \text{ and } \|\widehat{\boldsymbol{Y}}\|_{I(\mathbb{R})^n} = \sum_{j=1}^{n} \|\boldsymbol{Y}_j\|_{I(\mathbb{R})}.$$

Next, we discuss the following basic properties of interval analysis with the help of the dominance relation of intervals, the norm of an interval, and the gH-difference of two intervals, which are used throughout the thesis.

Lemma 1.1 (See [92]). *Let* $\boldsymbol{A}, \boldsymbol{B}$, *and* $\boldsymbol{C}$ *be elements of* $I(\mathbb{R})$. *If* $\boldsymbol{A} \preceq \boldsymbol{B}$ *and* $\boldsymbol{B} \preceq \boldsymbol{C}$, *then* $\boldsymbol{A} \preceq \boldsymbol{C}$.

Lemma 1.2 (See [93]). *Let* $\boldsymbol{A}, \boldsymbol{B} \in I(\mathbb{R})$.

(i) If $\boldsymbol{A} \not\prec \boldsymbol{B}$*, then* $\boldsymbol{A} \ominus_{gH} \boldsymbol{B} \not\prec \boldsymbol{0}$

(ii) If $\boldsymbol{A} \prec \boldsymbol{B}$*, then* $\boldsymbol{A} \ominus_{gH} \boldsymbol{B} \prec \boldsymbol{0}$.

Lemma 1.3 (See [9]). *For* $\boldsymbol{A}, \boldsymbol{B} \in I(\mathbb{R})$ *and* $y, z \in \mathbb{R}$*, we have*

(i) $y \odot (\boldsymbol{A} \oplus \boldsymbol{B}) = (y \odot \boldsymbol{A}) \oplus (y \odot \boldsymbol{B})$ *and*

(ii) $(y + z) \odot \boldsymbol{A} \subseteq (y \odot \boldsymbol{A}) \oplus (z \odot \boldsymbol{A})$.

Lemma 1.4 *Let* $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, *and* $\boldsymbol{D}$ *be elements of* $I(\mathbb{R})$*, and* $\epsilon, \delta \in \mathbb{R}$*. Then, the following relations hold.*

(i) If $\boldsymbol{A} \oplus \boldsymbol{B} \preceq \boldsymbol{C} \oplus \boldsymbol{D}$*, then* $\boldsymbol{A} \ominus_{gH} \boldsymbol{C} \preceq \boldsymbol{D} \ominus_{gH} \boldsymbol{B}$.

(ii) $\boldsymbol{A} \preceq \boldsymbol{B} \oplus [\epsilon, \epsilon] \iff \boldsymbol{A} \ominus_{gH} \boldsymbol{B} \preceq [\epsilon, \epsilon]$,

(iii) $(\boldsymbol{A} \ominus_{gH} \boldsymbol{B}) \oplus ([\epsilon, \epsilon] \ominus_{gH} [\delta, \delta]) = (\boldsymbol{A} \oplus [\epsilon, \epsilon]) \ominus_{gH} (\boldsymbol{B} \oplus [\delta, \delta])$*, and*

(iv) *if* $\boldsymbol{A} \preceq \boldsymbol{B} \oplus \boldsymbol{C}$ *and* $\boldsymbol{C} \preceq \boldsymbol{0}$, *then* $\boldsymbol{A} \preceq \boldsymbol{B}$.

Proof: See Appendix A.1. $\square$

Remark 1.2 *(i) Let* $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, *and* $\boldsymbol{D}$ *be elements of* $I(\mathbb{R})$. *If* $(\boldsymbol{A} \oplus \boldsymbol{B}) \ominus_{gH} (\boldsymbol{C} \oplus \boldsymbol{D}) = (\boldsymbol{A} \ominus_{gH} \boldsymbol{C}) \ominus_{gH} (\boldsymbol{D} \ominus_{gH} \boldsymbol{B})$, *then part (i) of Lemma 1.4 is an obvious property. However,* $(\boldsymbol{A} \oplus \boldsymbol{B}) \ominus_{gH} (\boldsymbol{C} \oplus \boldsymbol{D})$ *is not always equal to* $\big(\boldsymbol{A} \ominus_{gH} \boldsymbol{C}\big) \ominus_{gH} \big(\boldsymbol{D} \ominus_{gH} \boldsymbol{B}\big)$. *For instance, consider* $\boldsymbol{A} = [-3, 2]$, $\boldsymbol{B} = [0, 0]$, $\boldsymbol{C} = [4, 10]$, *and* $\boldsymbol{D} = [-7.5, -6]$. *Then,*

$$(\boldsymbol{A} \oplus \boldsymbol{B}) \ominus_{gH} (\boldsymbol{C} \oplus \boldsymbol{D}) = [-2, 0.5] \text{ and } (\boldsymbol{A} \ominus_{gH} \boldsymbol{C}) \ominus_{gH} (\boldsymbol{D} \ominus_{gH} \boldsymbol{B}) = [-1, -0.5].$$

Therefore, part (i) of Lemma 1.4 is not an obvious property.

(ii) For any $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$, *and* $\boldsymbol{D}$ *of* $I(\mathbb{R})$, *if we consider* $\boldsymbol{B} = \boldsymbol{0}$ *in part (i) of Lemma 1.4, then* $\boldsymbol{A} \preceq \boldsymbol{C} \oplus \boldsymbol{D} \implies \boldsymbol{A} \ominus_{gH} \boldsymbol{C} \preceq \boldsymbol{D}$.

(iii) For any $\boldsymbol{A}, \boldsymbol{B}$, *and* $\boldsymbol{C}$ *of* $I(\mathbb{R})$ *if* $\boldsymbol{A} \preceq \boldsymbol{B}$, *then* $\boldsymbol{A} \oplus \boldsymbol{C} \preceq \boldsymbol{B} \oplus \boldsymbol{C}$. *Thus, from part (i) of Lemma 1.4, we obtain*

$$\boldsymbol{A} \ominus_{gH} \boldsymbol{C} \preceq \boldsymbol{B} \ominus_{gH} \boldsymbol{C} \quad \text{or} \quad \boldsymbol{C} \ominus_{gH} \boldsymbol{B} \preceq \boldsymbol{C} \ominus_{gH} \boldsymbol{A}.$$

1.7.2 Sequence of Intervals

Definition 1.5 (Infimum of a set of intervals [7]). *Let* $\boldsymbol{P} \subseteq \overline{I(\mathbb{R})}$. *An interval* $\mathbf{X} \in I(\mathbb{R})$ *is said to be a lower bound of* $\boldsymbol{P}$ *if*

$$\mathbf{X} \preceq \boldsymbol{Y} \text{ for all } \boldsymbol{Y} \in \boldsymbol{P}.$$

A lower bound $\mathbf{X}$ *of* $\boldsymbol{P}$ *is called an* infimum *of* $\boldsymbol{P}$, *denoted by* $\inf \boldsymbol{P}$, *if*

$$\boldsymbol{Z} \preceq \mathbf{X} \text{ for all lower bounds } \boldsymbol{Z} \text{ of } \boldsymbol{P} \text{ in } I(\mathbb{R}).$$

Definition 1.6 (Supremum of a set of intervals [7]). *Let $\boldsymbol{P} \subseteq \overline{I(\mathbb{R})}$. An interval $\boldsymbol{X} \in I(\mathbb{R})$ is said to be an upper bound of $\boldsymbol{P}$ if*

$$\boldsymbol{Y} \preceq \mathbf{X} \ \text{ for all } \boldsymbol{Y} \in \boldsymbol{P}.$$

An upper bound $\mathbf{X}$ of $\boldsymbol{P}$ is called a supremum *of $\boldsymbol{P}$, denoted by $\sup \boldsymbol{P}$, if*

$$\mathbf{X} \preceq \boldsymbol{Z} \text{ for all upper bounds } \boldsymbol{Z} \text{ of } \boldsymbol{P} \text{ in } I(\mathbb{R}).$$

Remark 1.3 (See [7]). *Let $\boldsymbol{P} = \left\{ [a_\mu, b_\mu] \in \overline{I(\mathbb{R})} : \mu \in \Lambda \text{ and } \Lambda \text{ being an index set} \right\}$. Then, by Definitions 1.5 and 1.6, it follows that $\inf \boldsymbol{P} = \left[\inf_{\mu \in \Lambda} a_\mu,\ \inf_{\mu \in \Lambda} b_\mu \right]$ and $\sup \boldsymbol{P} = \left[\sup_{\mu \in \Lambda} a_\mu,\ \sup_{\mu \in \Lambda} b_\mu \right]$.*

Remark 1.4 *Let $\boldsymbol{P} \subseteq I(\mathbb{R})$ be a finite set of comparable intervals. Then, the infimum and supremum of $\boldsymbol{S}$ coincide with the minimum and maximum of the set $\boldsymbol{S}$, respectively.*

1.7.3 Fundamental Results on Interval-Valued Functions

Let $\mathcal{Y}$ be a nonempty subset of $\mathbb{R}^n$. A function $\mathbf{F} : \mathcal{Y} \to I(\mathbb{R})$ is said to be an IVF. For each $y \in \mathcal{Y}$, $\mathbf{F}$ is presented by

$$\mathbf{F}(y) = \left[\underline{f}(y), \overline{f}(y) \right],$$

where $\underline{f}$ and $\overline{f}$ are real-valued functions on $\mathcal{Y}$ such that $\underline{f}(y) \leq \bar{f}(y)$ for all $y \in \mathcal{Y}$.

Definition 1.7 (Interval-valued convex function [48]). *If $\mathcal{Y}$ is convex, then the IVF $\boldsymbol{F}$ is said to be convex on $\mathcal{Y}$ if for any y_1 and $y_2 \in \mathcal{Y}, \beta_1,\ \beta_2 \in [0, 1]$ with $\beta_1 + \beta_2 = 1$,*

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

Lemma 1.5 (See [48]). *If an IVF $\boldsymbol{F}$ is convex on a convex set $\mathcal{Y} \subseteq \mathbb{R}^n$, then $\underline{f}$ and $\overline{f}$ are convex on $\mathcal{Y}$ and vice-versa.*

Definition 1.8 (gH-continuity [37]). *The IVF $\boldsymbol{F}$ is called gH-continuous at $\bar{y} \in \mathcal{Y}$ if*

$$\lim_{\|h\| \to 0} \left(\boldsymbol{F}(\bar{y} + h) \ominus_{gH} \boldsymbol{F}(\bar{y}) \right) = \boldsymbol{0}.$$

If $\boldsymbol{F}$ is gH-continuous at each y in $\mathcal{Y}$, then $\boldsymbol{F}$ is said to be gH-continuous on $\mathcal{Y}$.

Definition 1.9 (Proper IVF). *Let $\mathcal{Y}$ be a nonempty subset of $\mathbb{R}^n$ and $\boldsymbol{F} : \mathcal{Y} \to \overline{I(\mathbb{R})}$ be an extended IVF. If there exists $\bar{y} \in \mathcal{Y}$, such that*

$$\boldsymbol{F}(\bar{y}) \prec +\boldsymbol{\infty} \text{ and } -\boldsymbol{\infty} \prec \boldsymbol{F}(y) \text{ for all } y \in \mathcal{Y},$$

then $\boldsymbol{F}$ is called a proper IVF.

Definition 1.10 (Domain of an IVF). *Let $\mathcal{Y}$ be a nonempty subset of $\mathbb{R}^n$. For an extended IVF $\boldsymbol{F} : \mathcal{Y} \to \overline{I(\mathbb{R})}$, the domain of $\boldsymbol{F}$, denoted as dom $\boldsymbol{F}$, is defined by*

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

Definition 1.11 (Effective domain of an IVF). *Let $\boldsymbol{F} : \mathcal{Y} \to \overline{I(\mathbb{R})}$ be an extended IVF. The effective domain of $\boldsymbol{F}$ is defined as*

$$\text{dom}(\boldsymbol{F}) = \left\{ y \in \mathcal{Y} : \|\boldsymbol{F}(y)\|_{I(\mathbb{R})} < +\infty \right\}.$$

Definition 1.12 (Linear IVF [5]). *Let $\mathcal{Y}$ be a linear subspace of $\mathbb{R}^n$. The function $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ is said to be linear if*

(i) $\boldsymbol{F}(\lambda y) = \lambda \odot \boldsymbol{F}(y)$ *for all* $y \in \mathcal{Y}$ *and for all* $\lambda \in \mathbb{R}$ *and*

(ii) for all $y,\ w \in \mathcal{Y}$, either $\boldsymbol{F}(y) \oplus \boldsymbol{F}(w) = \boldsymbol{F}(y+w)$ or none of $\boldsymbol{F}(y) \oplus \boldsymbol{F}(w)$ and $\boldsymbol{F}(y+w)$ dominate the other.

Definition 1.13 (gH-derivative [32]). *Let $\mathcal{Y}$ be a nonempty subset of $\mathbb{R}$. The gH-derivative of an IVF $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ at a point $\bar{y} \in \mathcal{Y}$ is defined by*

$$\boldsymbol{F}'(\bar{y}) = \lim_{h \to 0} \frac{1}{h} \odot (\boldsymbol{F}(\bar{y}+h) \ominus_{gH} \boldsymbol{F}(\bar{y})), \text{ provided the limit exists.}$$

Remark 1.5 (See [32]). *The gH-derivative of an IVF $\boldsymbol{F}$ at a point $\bar{y} \in \mathcal{Y}$ exists if the derivatives of $\underline{f}$ and $\overline{f}$ at $\bar{y}$ exists and*

$$\boldsymbol{F}'(\bar{y}) = \left[\min\{\underline{f}'(\bar{y}), \overline{f}'(\bar{y})\}, \max\{\underline{f}'(\bar{y}), \overline{f}'(\bar{y})\}\right].$$

However, the converse is not true in general.

Definition 1.14 (Partial gH-derivative [37]). *Let $\mathcal{Y}$ be a nonempty subset of $\mathbb{R}^n$. Then, for a given IVF $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$, define a function $\boldsymbol{H}_j$ by*

$$\boldsymbol{H}_j(y_j) = \boldsymbol{F}(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_{j-1}, y_j, \bar{y}_{j+1}, \ldots, \bar{y}_n),$$

where $\bar{y} = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)^\top \in \mathcal{Y}$. If $\boldsymbol{H}_j'(\bar{y}_j)$ exists, it is called the partial gH-derivative of $\boldsymbol{F}$ at $\bar{y}$, denoted as $D_j\boldsymbol{F}(\bar{y})$, i.e.,

$$D_j\boldsymbol{F}(\bar{y}) = \boldsymbol{H}_j'(\bar{y}_j) \text{ for all } j = 1, 2, \ldots, n.$$

Definition 1.15 (gH-gradient [37]). *The gH-gradient of $\boldsymbol{F}$ at $\bar{y}$ is denoted by $\nabla\boldsymbol{F}(\bar{y})$ and defined by the interval vector*

$$\nabla\boldsymbol{F}(\bar{y}) = (D_1\boldsymbol{F}(\bar{y}), D_2\boldsymbol{F}(\bar{y}), \ldots, D_n\boldsymbol{F}(\bar{y}))^\top.$$

Definition 1.16 (gH-directional derivative [5]). *Let $\boldsymbol{F}$ be an IVF on a nonempty subset $\mathcal{Y}$ of $\mathbb{R}^n$. Let $\bar{y} \in \mathcal{Y}$ and $h \in \mathbb{R}^n$ be such that $\bar{y} + \beta h \in \mathcal{Y}$ and*

$$\lim_{\beta \to 0+} \frac{1}{\beta} \odot \big(\boldsymbol{F}(\bar{y} + \beta h) \ominus_{gH} \boldsymbol{F}(\bar{y})\big) \text{ exists finitely.}$$

Then, the limit at $\bar{y}$ in the direction h is said to be gH-directional derivative of $\boldsymbol{F}$, and it is denoted by $\boldsymbol{F}_{\mathscr{D}}(\bar{y})(h)$.

Definition 1.17 (Efficient point [5]). *A point $\bar{y} \in \mathcal{Y}$ is called an efficient point of the IOP $\min\limits_{y \in \mathcal{Y}} \boldsymbol{F}(y)$ if $\boldsymbol{F}(y) \nprec \boldsymbol{F}(\bar{y})$ for all $y \in \mathcal{Y}$.*

Definition 1.18 (Weak efficient point). *A point $\bar{y} \in \mathcal{Y}$ is said to be a weak efficient point of the IOP $\min\limits_{y \in \mathcal{Y}} \boldsymbol{F}(y)$ if $\boldsymbol{F}(\bar{y}) \preceq \boldsymbol{F}(y)$ for all $y \in \mathcal{Y}$.*

Definition 1.19 (Infimum and supremum of an IVF [7]). *Let $\mathcal{S}$ be a nonempty subset of $\mathcal{Y}$ and $\boldsymbol{F} : \mathcal{S} \to \overline{I(\mathbb{R})}$ be an extended IVF. Then, the infimum of $\boldsymbol{F}$ denoted as $\inf\limits_{y \in \mathcal{S}} \boldsymbol{F}(y)$ is equal to the infimum of range set of $\boldsymbol{F}$, i.e.,*

$$\inf_{y \in \mathcal{S}} \boldsymbol{F}(y) = \inf\{\boldsymbol{F}(y) : y \in \mathcal{S}\}.$$

Similarly, the supremum of an IVF is defined by

$$\sup_{y \in \mathcal{S}} \boldsymbol{F}(y) = \sup\{\boldsymbol{F}(y) : y \in \mathcal{S}\}.$$

Definition 1.20 (Lower limit of an extended IVF [7]). *The lower limit of an extended IVF $\boldsymbol{F}$ at $\bar{y} \in \mathcal{Y}$, denoted $\liminf\limits_{y \to \bar{y}} \boldsymbol{F}(y)$, is defined by*

$$\begin{aligned} \liminf_{y \to \bar{y}} \boldsymbol{F}(y) &= \lim_{\delta \downarrow 0}(\inf\{\boldsymbol{F}(y) : y \in B_\delta(\bar{y})\}) \\ &= \sup_{\delta > 0}(\inf\{\boldsymbol{F}(y) : y \in B_\delta(\bar{y})\}). \end{aligned}$$

Lemma 1.6 (See [7]). *Let $\boldsymbol{F}_1$, $\boldsymbol{F}_2$ be two proper extended IVFs and $\mathcal{S}$ be a nonempty subset of $\mathcal{Y}$. Then,*

(i) $\inf_{y\in\mathcal{S}} \boldsymbol{F}_1(y) \oplus \inf_{y\in\mathcal{S}} \boldsymbol{F}_2(y) \preceq \inf_{y\in\mathcal{S}}\{\boldsymbol{F}_1(y) \oplus \boldsymbol{F}_2(y)\}$ *and*

(ii) $\sup_{y\in\mathcal{S}}\{\boldsymbol{F}_1(y) \oplus \boldsymbol{F}_2(y)\} \preceq \sup_{y\in\mathcal{S}} \boldsymbol{F}_1(y) \oplus \sup_{y\in\mathcal{S}} \boldsymbol{F}_2(y)$.

Lemma 1.7 *Let $\mathcal{S}$ be a nonempty subset of $\mathcal{Y}$ and $\boldsymbol{F} : \mathcal{S} \to \overline{I(\mathbb{R})}$ be an extended IVF. Then, for $\mathcal{S}_1, \mathcal{S}_2 \subseteq \mathcal{S}$ with $\mathcal{S}_1 \subseteq \mathcal{S}_2$ and $\delta \geq 0$,*

(i) $\inf_{y\in\mathcal{S}_2} \boldsymbol{F}(y) \preceq \inf_{y\in\mathcal{S}_1} \boldsymbol{F}(y)$,

(ii) $\sup_{y\in\mathcal{S}_1} \boldsymbol{F}(y) \preceq \sup_{y\in\mathcal{S}_2} \boldsymbol{F}(y)$,

(iii) $\inf_{y\in\mathcal{S}} (\delta \odot \boldsymbol{F})(y) = \delta \odot \inf_{y\in\mathcal{S}} \boldsymbol{F}(y)$, *and*

(iv) $\sup_{y\in\mathcal{S}} (\delta \odot \boldsymbol{F})(y) = \delta \odot \sup_{y\in\mathcal{S}} \boldsymbol{F}(y)$.

Proof: See Appendix A.2. □

Definition 1.21 (Monotonicity of IVFs [5]). *An IVF $\boldsymbol{F}$ from a nonempty subset $\mathcal{Y}$ of $\mathbb{R}^n$ to $I(\mathbb{R})$ is said to be monotonically increasing if for all $y_1, y_2 \in \mathcal{Y}$,*

$$y_1 \leq y_2 \implies \boldsymbol{F}(y_1) \preceq \boldsymbol{F}(y_2).$$

The IVF $\boldsymbol{F}$ from a nonempty subset $\mathcal{Y}$ of $\mathbb{R}^n$ to $I(\mathbb{R})$ is said to be monotonically decreasing if for all $y_1, y_2 \in \mathcal{Y}$,

$$y_1 \leq y_2 \implies \boldsymbol{F}(y_2) \preceq \boldsymbol{F}(y_1).$$

Definition 1.22 (Sequence in $I(\mathbb{R})^n$ [6]). *An IVF* $\widehat{\boldsymbol{F}} : \mathbb{N} \to I(\mathbb{R})^n$ *is called a sequence in* $I(\mathbb{R})^n$.

Definition 1.23 (Convergence of a sequence in $I(\mathbb{R})^n$ [6]). *A sequence* $\{\widehat{\boldsymbol{G}}_k\}$ *in* $I(\mathbb{R})^n$ *is said to be* convergent *to* $\widehat{\boldsymbol{G}} \in I(\mathbb{R})^n$ *if for each* $\epsilon > 0$, *there exists an* $m \in \mathbb{N}$ *such that*

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

The interval $\widehat{\boldsymbol{G}}$ *is called limit of the sequence* $\{\widehat{\boldsymbol{G}}_k\}$ *and is presented by* $\lim\limits_{k\to\infty} \widehat{\boldsymbol{G}}_k = \widehat{\boldsymbol{G}}$.

Remark 1.6 (See [6]). *It is to be acclaimed that if a sequence* $\{\widehat{\boldsymbol{G}}_k\}$ *in* $I(\mathbb{R})^n$ *converges to some* $\widehat{\boldsymbol{G}} \in I(\mathbb{R})^n$, *where* $\widehat{\boldsymbol{G}}_k = (\boldsymbol{G}_{1_k}, \boldsymbol{G}_{2_k}, \ldots, \boldsymbol{G}_{n_k})^\top$ *and* $\widehat{\boldsymbol{G}} = (\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_n)^\top$, *then from the Definition 1.4, the sequence* $\{\boldsymbol{G}_{j_k}\}$ *in* $I(\mathbb{R})$ *converges to* $\boldsymbol{G}_j \in I(\mathbb{R})$ *for each* $j = 1, 2, \ldots, n$.

Lemma 1.8 *Let* $\{\boldsymbol{X}_k\}$ *and* $\{\boldsymbol{Y}_k\}$ *be two sequence in* $I(\mathbb{R})^n$ *and* $\lim\limits_{k\to\infty} \boldsymbol{X}_k = \boldsymbol{X}$ *and* $\lim\limits_{k\to\infty} \boldsymbol{Y}_k = \boldsymbol{Y}$. *If* $\boldsymbol{X}_k \preceq \boldsymbol{Y}_k$ *for all* k, *then* $\boldsymbol{X} \preceq \boldsymbol{Y}$.

Proof: Let $\mathbf{X}_k \preceq \mathbf{Y}_k$ for all k, then we have

$$\begin{aligned}
& \underline{x}_k \leq \underline{y}_k \text{ and } \overline{x}_k \leq \overline{y}_k \\
\implies & \lim_{k\to\infty} \underline{x}_k \leq \lim_{k\to\infty} \underline{y}_k \text{ and } \lim_{k\to\infty} \overline{x}_k \leq \lim_{k\to\infty} \overline{y}_k \\
\implies & \underline{x} \leq \underline{y} \text{ and } \overline{x} \leq \overline{y} \\
\implies & \mathbf{X} \preceq \mathbf{Y},
\end{aligned}$$

which is the required result. $\square$

Lemma 1.9 *If a sequence* $\{\widehat{\boldsymbol{G}}_k\}$ *in* $I(\mathbb{R})^n$ *converges to* $\widehat{\boldsymbol{G}} \in I(\mathbb{R})^n$, *then every subsequence of* $\{\widehat{\boldsymbol{G}}_k\}$ *converges to the same limit* $\widehat{\boldsymbol{G}}$.

Proof: Let $\{\widehat{\mathbf{G}}_{k_j}\}$ be a subsequence of the convergent sequence $\{\widehat{\mathbf{G}}_k\}$. Since $\{\widehat{\mathbf{G}}_k\}$ has a limit $\widehat{\mathbf{G}}$, then for each $\epsilon > 0$, there exists an $m \in \mathbb{N}$ such that

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

As $\{k_j\}$ is increasing, there exists $p \in \mathbb{N}$ such that $k_j \geq m$ for each $j \geq p$. Therefore,

$$\|\widehat{\mathbf{G}}_{k_j} \ominus_{gH} \widehat{\mathbf{G}}\|_{I(\mathbb{R})^n} < \epsilon \text{ for all } j \geq p.$$

This implies that $\{\widehat{\mathbf{G}}_{k_j}\}$ converges to $\widehat{\mathbf{G}}$. □

Definition 1.24 (gH-subgradient [6]). *For a convex IVF $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ on a convex set $\mathcal{Y} \subseteq \mathbb{R}^n$, an element $\widehat{\boldsymbol{G}} = (\boldsymbol{G}_1, \boldsymbol{G}_2, \ldots, \boldsymbol{G}_n)^\top \in I(\mathbb{R})^n$ is called a gH-subgradient of $\boldsymbol{F}$ at $\bar{y} \in \mathcal{Y}$ if*

$$(y - \bar{y})^\top \odot \widehat{\boldsymbol{G}} \ \preceq \ \boldsymbol{F}(y) \ominus_{gH} \boldsymbol{F}(\bar{y}) \text{ for all } y \in \mathcal{Y}.$$

The collection of all subgradients of $\boldsymbol{F}$ at $\bar{y} \in \mathcal{Y}$ is called gH-subdifferential and is denoted by $\partial\boldsymbol{F}(\bar{y})$.

Definition 1.25 (gH-Gâteaux derivative [5]). *Let $\boldsymbol{F}$ be an IVF on a nonempty open subset $\mathcal{Y}$ of $\mathbb{R}^n$. If for each $h \in \mathbb{R}^n$ and at $\bar{y} \in \mathcal{Y}$, the limit*

$$\boldsymbol{F}_{\mathscr{G}}(\bar{y})(h) = \lim_{\lambda \to 0+} \frac{1}{\lambda} \odot \left(\boldsymbol{F}(\bar{y} + \lambda h) \ominus_{gH} \boldsymbol{F}(\bar{y})\right)$$

exists and $\boldsymbol{F}_{\mathscr{G}}(\bar{y})$ is a gH-continuous linear IVF from $\mathbb{R}^n$ to $I(\mathbb{R})$, then $\boldsymbol{F}_{\mathscr{G}}(\bar{y})$ is called gH-Gâteaux derivative of $\boldsymbol{F}$ at $\bar{y}$. If $\boldsymbol{F}$ has a gH-Gâteaux derivative at $\bar{y}$, then $\boldsymbol{F}$ is called gH-Gâteaux differentiable at $\bar{y}$.

Definition 1.26 (gH-Fréchet derivative [5]). *Let $\boldsymbol{F}$ be an IVF on be a nonempty open subset $\mathcal{Y}$ of $\mathbb{R}^n$. For $\bar{y} \in \mathcal{Y}$ and $h \in \mathbb{R}^n$, if there exists a gH-continuous and linear* *mapping* $\boldsymbol{F}_{\mathscr{F}} : \mathcal{Y} \to I(\mathbb{R})$ *with the following property*

$$\lim_{\|h\|\to 0} \frac{\|\boldsymbol{F}(\bar{y}+h) \ominus_{gH} \boldsymbol{F}(\bar{y}) \ominus_{gH} \boldsymbol{F}_{\mathscr{F}}(h)\|_{I(\mathbb{R})}}{\|h\|} = 0,$$

then $\boldsymbol{F}$ *is said to have a gH-Fréchet derivative at* $\bar{y}$*, denoted by* $\boldsymbol{F}_{\mathscr{F}}$*.*

Theorem 1.1 (See [5]). *Let* $\mathcal{Y}$ *be a nonempty open subset of* $\mathbb{R}^n$ *and the Fréchet derivative of IVF* $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ *exists at some* $\bar{y} \in \mathcal{Y}$*. Then, the gH-Gâteaux derivative of* $\boldsymbol{F}$ *at* $\bar{y}$ *exists along any* $h \in \mathbb{R}^n$ *and values of both the derivatives are equal.*

Theorem 1.2 (See [5]). *Let* $\mathcal{Y}$ *be a nonempty subset of* $\mathbb{R}^n$*. If* $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ *has a gH-Gâteaux derivative at* $\bar{y}$ *in every direction* $h \in \mathbb{R}^n$ *and* $\bar{y} \in \mathcal{Y}$ *be an efficient point of* $\boldsymbol{F}$*, then*

$$0 \in \boldsymbol{F}_{\mathscr{G}}(\bar{y})(h) \text{ for all } h \in \mathbb{R}^n.$$

Remark 1.7 (See [5]). *Let* $\boldsymbol{F}$ *be an IVF on a nonempty open subset* $\mathcal{Y}$ *of* $\mathbb{R}^n$*. Let* $\boldsymbol{F}$ *has gH-Gâteaux derivative at* $\bar{y} \in \mathcal{Y}$*. Then,* $\boldsymbol{F}$ *has gH-directional derivative at* $\bar{y}$ *in every direction* $h \in \mathbb{R}^n$ *also.*

Theorem 1.3 (See [5]). *Let* $\mathcal{Y}$ *be a nonempty open convex subset of* $\mathbb{R}^n$ *and* $\boldsymbol{F}$ *be an IVF which is gH-differentiable at* $\bar{y} \in \mathcal{Y}$*. If the IVF* $\boldsymbol{F}$ *is convex on* $\mathcal{Y}$*, then*

$$(y - \bar{y})^\top \odot \nabla \boldsymbol{F}(\bar{y}) \preceq \boldsymbol{F}(y) \ominus_{gH} \boldsymbol{F}(\bar{y}) \text{ for all } y, \bar{y} \in \mathcal{Y}.$$

Theorem 1.4 (See [6]). *Let* $\mathcal{Y} \subseteq \mathbb{R}^n$ *and the IVF* $\boldsymbol{F} : \mathcal{Y} \to I(\mathbb{R})$ *be gH-differentiable at* $\bar{y} \in \mathcal{Y}$*. Then,* $\boldsymbol{F}$ *has gH-directional derivative at* $\bar{y}$ *for every direction* $h \in \mathbb{R}^n$ *and*

$$\partial \boldsymbol{F}(\bar{y}) = \{\nabla \boldsymbol{F}(\bar{y})\}.$$

1.7.4 Preliminaries of Set-Valued Functions

Definition 1.27 *A nonempty set $K \subseteq \mathbb{R}^m$ is called a cone if for every $t \in K$ and $\lambda \in \mathbb{R}_+$, we have $\lambda t \in K$. The cone K is called*

(i) convex if for all $t_1, t_2 \in K$, we have $t_1 + t_2 \in K$,

(ii) pointed if $K \cap (-K) = \{0\}$,

(iv) solid if $\operatorname{int}(K) \neq \emptyset$.

Definition 1.28 (Partial ordering on $\mathbb{R}^m$ [94]). *Let the cone K is convex and pointed. Then, K generates a partial order $\preceq$ on $\mathbb{R}^m$, and if K is solid, it generates a strict order $\prec$ on $\mathbb{R}^m$ defined as follows: for any $y, z \in \mathbb{R}^m$,*

$$y \preceq z \iff z - y \in K, \text{ and } y \prec z \iff z - y \in \operatorname{int}(K).$$

If $y \preceq z$, we often present it by $z \succeq y$. Similarly, $y \prec z$ is often presented by $z \succ y$.

Definition 1.29 (Lipschitz continuity on $\mathbb{R}^n$). *Let S be a nonempty subset of $\mathbb{R}^n$. A function $F : S \to \mathbb{R}^m$ is said to be Lipschitz continuous on S if there exists a constant $L > 0$ such that*

$$\|F(y) - F(x)\| \leq L\|y - x\| \text{ for all } x, y \in S.$$

Definition 1.30 (Minimal and weakly minimal elements of a set [95]). *Let $A \in \mathcal{P}(\mathbb{R}^m)$. The set of minimal and weakly minimal elements of A with respect to K is defined by*

$$\operatorname{Min}(A, K) = \{z \in A : (z - K) \cap A = \{z\}\} \text{ and}$$
$$\operatorname{WMin}(A, K) = \{z \in A : (z - \operatorname{int}(K)) \cap A = \emptyset\}, \text{ respectively.}$$

Proposition 1.1 [95] *Any compact set $A \in \mathcal{P}(\mathbb{R}^m)$ satisfies the domination property with respect to K, i.e., $A + K = \operatorname{Min}(A, K) + K$.*

Next, we recall the concept of the Gerstewitz scalarizing function, which performs a significant role in the main results of the thesis.

Definition 1.31 (Gerstewitz function [96]). *For an element $e \in \operatorname{int}(K)$ and $z \in \mathbb{R}^m$, the Gerstewitz function $\Psi_e : \mathbb{R}^m \to \mathbb{R}$ associated with e and K is defined by*

$$\Psi_e(z) = \min\{t \in \mathbb{R} \ : \ te \in z + K\}.$$

Proposition 1.2 (See [79]). *For a given element $e \in \operatorname{int}(K)$, the function $\Psi_e : \mathbb{R}^m \to \mathbb{R}$ has the following properties:*

(i) Ψ_e *is sublinear on* $\mathbb{R}^m$.

(ii) Ψ_e *is positive homogenous of degree 1 on* $\mathbb{R}^m$.

(iii) Ψ_e *is Lipschitz continuous on* $\mathbb{R}^m$.

(iv) Ψ_e *is monotone, i.e., for all* $p, q \in \mathbb{R}^m$,

$$p \preceq q \implies \Psi_e(p) \leq \Psi_e(q) \quad \text{ and } \quad p \prec q \implies \Psi_e(p) < \Psi_e(q).$$

(v) Ψ_e *satisfies the representability property, i.e.,*

$$-K = \{z \in \mathbb{R}^m : \Psi_e(z) \leq 0\} \text{ and } -\operatorname{int}(K) = \{z \in \mathbb{R}^m : \Psi_e(z) < 0\}.$$

(vi) Ψ_e *has the translativity property, i.e.,* $\Psi_e(y + te) = \Psi_e(y) + t \ \forall \ y \in \mathbb{R}^m$.

Next, we recall the set order relations between the nonempty subsets of $\mathbb{R}^m$.

Definition 1.32 (Lower set less and strict lower set less relations [78]). *Let A and B be two sets in $\mathcal{P}(\mathbb{R})^m$. The lower set less ($\preceq^l$) and strict lower set less ($\prec^l$) relations* *for a given cone K are defined by*

$$A \preceq^l B \iff B \subseteq A + K \quad \textit{and} \quad A \prec^l B \iff B \subseteq A + \operatorname{int}(K), \ \textit{respectively.}$$

In this work, we aim to derive a Newton and a quasi-Newton method to identify weakly minimal solutions of the following unconstrained set optimization problem. Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a nonempty set-valued mapping. The unconstrained set optimization problem that we study is as follows:

$$\preceq^l \text{-}\min_{x \in \mathbb{R}^n} F(x), \tag{SOP}$$

where the solution concept is the notion of weakly minimal solutions with respect to a given ordering cone K in $\mathbb{R}^m$ as given below.

Definition 1.33 (Weakly minimal solution of (SOP) [1]). *A point $\bar{x} \in \mathbb{R}^n$ is a local weakly minimal solution of (SOP) if there exists a neighbourhood $U \subset \mathbb{R}^n$ of $\bar{x}$ such that there does not exist any $x \in U$ with $F(x) \prec^l F(\bar{x})$. The point $\bar{x}$ is called a weakly minimal solution of (SOP) if $U = \mathbb{R}^n$.*

Assumption 1 *For the sake of deriving Newton and quasi-Newton method for* (SOP)*, we employ the following assumption for dealing with* (SOP)*. The function $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ in* (SOP) *is given by finitely many vector-valued functions:*

$$F(x) = \left\{ f^1(x), f^2(x), \ldots, f^p(x) \right\} \ \textit{for all } x \in \mathbb{R}^n, \tag{1.1}$$

where $f^1, f^2, \ldots, f^p$ are twice continuously differentiable.

1.8 Optimality Conditions on Set-Valued Maps

In this section, we recall some results on optimality conditions for weakly minimal solutions of (SOP) under Assumption 1. These results are the foundation for constructing the proposed Newton and quasi-Newton method to capture weakly minimal solutions of (SOP). We start with some index-related set-valued mappings analogous to those given in [1].

Definition 1.34 (Active indices for set-valued maps [1]).

(i) The active index of minimal elements associated with the set-valued mapping F of (SOP) is $I : \mathbb{R}^n \rightrightarrows [p]$, defined as

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

(ii) The active index of weakly minimal elements associated with the set-valued mapping F is $I_W : \mathbb{R}^n \rightrightarrows [p]$, defined by

$$I_W(x) = \{i \in [p] : f^i(x) \in \text{ WMin}(F(x), K)\}.$$

(iii) For a given $r \in \mathbb{R}^m$, the set-valued mapping $I_r : \mathbb{R}^n \rightrightarrows [p]$ is given by

$$I_r(x) = \{i \in I(x) : f^i(x) = r\}.$$

It is to notice that for any $u \in \mathbb{R}^m$, $I_u(x) = \emptyset$ for $u \notin \text{Min}(F(x), K)$; $I_u(x) \cap I_v(x) = \emptyset$ for any $u \neq v \in \mathbb{R}^m$ and $I(x) = \bigcup\limits_{u \in \text{Min}(F(x),K)} I_u(x)$.

Definition 1.35 (Cardinality of a set of minimal elements [1]). *The map $w : \mathbb{R}^n \to \mathbb{N} \cup \{0\}$, which is defined by*

$$w(x) = |\text{Min}(F(x), K)|$$

is called the cardinality of the set of minimal elements of $F(x)$ with respect to K. Further, at any $\bar{x} \in \mathbb{R}^n$, we denote $w(\bar{x}) = \bar{w}$ for simplicity.

Definition 1.36 (Partition set at a point [1]). *Let us consider an element $x \in \mathbb{R}^n$ and an enumeration $\{r_1^x, r_2^x, \ldots, r_{w(x)}^x\}$ of the set $\operatorname{Min}(F(x), K)$. The partition set at x is defined by*

$$P_x = \prod_{j=1}^{w(x)} I_{r_j^x}(x), \ \ where \ \prod_{j=1}^{w(x)} I_{r_j^x}(x) = I_{r_j^x}(x) \times I_{r_j^x}(x) \times \cdots \times I_{r_j^x}(x).$$

Throughout the thesis, for a given iterative point $x_k \in \mathbb{R}^n$, a generic element of the partition set P_{x_k} is denoted by a^k. For every $j \in [w(x_k)]$, we denote the j-th component of a^k by a_j^k, where $k = 1, 2, 3, \ldots$. Specifically, if $|P_{x_k}| = p_k$ and $\operatorname{Min}(F(x_k), K) = \{r_1^{x_k}, r_2^{x_k}, \ldots, r_{w(x_k)}^{x_k}\}$, then

$$P_{x_k} = \left\{a^1, a^2, \ldots, a^{p_k}\right\},$$

where for each $k = 1, 2, 3, \ldots, p_k$,

$$a^k = \left(a_1^k, a_2^k, \ldots, a_{w(x_k)}^k\right), \ a_j^k \in I_{r_j^{x_k}}, \ j \in [w(x_k)].$$

Now, we describe a family of vector optimization problems that will help us to find weakly minimal solutions of (SOP). At a given iterate, one of the vector optimization problems from this family is solved, and the solution is checked for optimality. If the optimality condition is met, then we stop. Otherwise, we solve another vector optimization problem from the family to progress the iteration. Towards finding a stopping condition, we study the following result from [1].

Theorem 1.5 (See [1]). *Let $P_{\bar{x}}$ be the partition set at $\bar{x}$ and $\bar{w} = w(\bar{x})$. For every $a = (a_1, a_2, \ldots, a_{\bar{w}}) \in P_{\bar{x}}$, define a vector-valued function $\widetilde{f}^a : \mathbb{R}^n \to \prod_{j=1}^{\bar{w}} \mathbb{R}^m = \mathbb{R}^{m\bar{w}}$ by*

$$\widetilde{f}^a(x) = \left(f^{a_1}(x), f^{a_2}(x), \ldots, f^{a_{\bar{w}}}(x)\right)^\top.$$

Let $\widetilde{K} \in \mathcal{P}(\mathbb{R}^{m\bar{w}})$ be the cone given by $\widetilde{K} = \prod_{j=1}^{\bar{w}} K$, and $\preceq_{\widetilde{K}}$ denote the partial order in $\mathbb{R}^{m\bar{w}}$ induced by $\widetilde{K}$. Then, $\bar{x}$ is a local weakly minimal solution of (SOP) if and only if for every $a \in P_{\bar{x}}$, $\bar{x}$ is a local weakly minimal solution of the vector optimization problem

$$\preceq_{\widetilde{K}} \text{-}\min_{x \in \mathbb{R}^n} \widetilde{f}^a(x). \tag{VOP}$$

Next, to find a necessary optimality condition for weakly minimal solutions of (SOP), we provide the concept of a stationary point for (SOP).

Definition 1.37 (Stationary points of (SOP) [1]). *A point $\bar{x}$ is called a stationary point of (SOP) if for every $a = (a_1, a_2, \ldots, a_{\bar{w}}) \in P_{\bar{x}}$ and $u \in \mathbb{R}^n$, there exists $j \in [\bar{w}]$ such that*

$$\nabla f^{a_j}(\bar{x})^\top u \notin -\operatorname{int}(K), \ \textit{i.e.}, \ \Psi_e(\nabla f^{a_j}(\bar{x})^\top u) \geq 0. \tag{1.2}$$

Definition 1.38 (Stationary point for (VOP) [97]). *A point $\bar{x}$ is called a stationary (or critical) point of (VOP) if for every $a = (a_1, a_2, \ldots a_{\bar{w}}) \in P_{\bar{x}}$ and $u \in \mathbb{R}^n$, there exists $j \in [\bar{w}]$ such that*

$$\mathcal{R}(\mathcal{J} f^{a_j}(\bar{x}) \cap (-\operatorname{int}(K)) = \emptyset,$$

Therefore, $\bar{x}$ is stationary if and only if for every $a = (a_1, a_2, \ldots a_{\bar{w}}) \in P_{\bar{x}}$ and for all $u \in \mathbb{R}^n$, we have

$$\nabla f^{a_j}(\bar{x})^\top u \notin -\operatorname{int}(K) \textit{ for all } j \in [\bar{w}].$$

Lemma 1.10 *A point $\bar{x} \in \mathbb{R}^n$ is a stationary point of (SOP) if and only if for every $a \in P_{\bar{x}}$, $\bar{x}$ is a stationary point of (VOP) for every $a \in P_{\bar{x}}$.*

Proof: Let $\bar{x}$ be a stationary of (SOP). We prove that $\bar{x}$ is a stationary point of (VOP). On the contrary, suppose $\bar{x}$ is not a stationary point of (VOP) for some $a \in P_{\bar{x}}$. Then, there exists $\bar{u} \in \mathbb{R}^n$ such that

$$\nabla f^{a_j}(\bar{x})^\top \bar{u} \in -\operatorname{int}(K) \ \forall \ j \in [\bar{w}]$$

$$\text{i.e., } 0 \in \nabla f^{a_j}(\bar{x})^\top \bar{u} + \text{int}(K) \ \forall\ j \in [\bar{w}], \tag{1.3}$$

which implies that

$$\{0\} \subseteq \{0\} + K \overset{(1.3)}{\subseteq} \{\nabla f^{a_j}(\bar{x})^\top \bar{u}\}_{j\in[\bar{w}]} + \text{int}(K) + K \subseteq \{\nabla f^{i}(\bar{x})^\top \bar{u}\}_{i\in[p]} + \text{int}(K).$$

Therefore, $\{\nabla f^{i}(\bar{x})^\top \bar{u}\}_{i\in[p]} \prec^l \{0\}$, which is contradictory to $\bar{x}$ being a stationary point of (SOP). Hence, $\bar{x}$ must be a stationary point of (VOP) for every $a \in P_{\bar{x}}$.
Conversely, suppose $\bar{x}$ is a stationary point of (VOP) for every $a \in P_{\bar{x}}$. We prove that $\bar{x}$ is a stationary point of (SOP). To prove this, let us assume that $\bar{x}$ is not a stationary point of (SOP). Thus, there exists $\bar{u} \in \mathbb{R}^n$ such that

$$\begin{aligned} & \nabla f^{i}(\bar{x})^\top \bar{u} \prec 0 \ \forall\ i \in [p] \\ \implies & 0 \in \nabla f^{i}(\bar{x})^\top \bar{u} + \text{int}(K) \ \forall\ i \in [p] \\ \implies & \nabla f^{a_j}(\bar{x})^\top \bar{u} \in -\text{int}(K) \ \forall\ j \in [\bar{w}] \subseteq [p], \end{aligned}$$

which is contradictory to $\bar{x}$ being a stationary point of (VOP) for every $a \in P_{\bar{x}}$. Hence, the result follows. $\square$

1.9 Objective of the Thesis

The objectives of the thesis are the following:

1. To analyse the gH-subdifferential calculus for interval-valued functions and its applications in nonconvex composite interval optimization problems. To derive Fritz-John-Type and KKT-type optimality conditions for IOPs.

2. To define the concept of interval-valued value function for IVFs and its application in IOPs.

3. To propose the concept of gH-Dini Hadamard ϵ-subdifferential and H_ϵ-subgradient and analysing their application in IOPs.

4. To explore and characterize the efficient solution to IOPs by using the above-defined generalized subdifferentials.

5. Next, we extend our work to set-valued optimization problems. In this, we propose a Newton and a quasi-Newton method to find weakly minimal solutions to set-valued optimization problems.

1.10 Organization of the Thesis

The thesis organization is as follows. This thesis consists of seven chapters including an introductory chapter and a chapter comprised of conclusion and future scopes. In this chapter, which is the introductory chapter, a concise but adequate literature of these topics has been discussed. It also defines the objective of this thesis. The outline of the thesis is as follows.

Chapter 1 begins with an introduction to some properties of interval analysis, smooth and nonsmooth analysis of interval-valued functions (IVFs), and optimality conditions for interval optimization problems (IOPs). Interval arithmetic and some important properties of intervals are also explained. The definitions of continuity and convexity for IVFs and their basic results are explained. Next, an introduction to basic definitions and fundamental results of set-valued optimization is discussed.

Chapter 2 contributes to gH-subdifferential calculus for IVFs. The concept of weak efficient solution of IOPs is defined. With the help of this concept, various optimality conditions for weak efficient solutions of nonsmooth IOPs, such as a Fermat-type, a Fritz-John-type, and a KKT-type, are established without assuming strong convexity.

A relation is proposed to estimate the weak efficient solution of a nonconvex composite problem to an IOP with the help of gH-subdifferential calculus of IVFs. The entire study is supported by suitable illustrative examples.

Chapter 3 analyses the effect of perturbations on an interval-valued objective function. The objective function due to perturbations has been named as interval-valued value function. Towards this, the concept of the Lagrangian of IVFs is given, followed by a weak duality theorem for IVFs. After that, the notion of interval-valued value function and the characterization of its gH-subdifferential set for IOPs. After that, the characterization of the stability of a solution to an IOP with gH-subdifferential set of an interval-valued value function is given. Also, an example of applying the results of the interval-valued value function is discussed.

In Chapter 4, the concepts of gH-Dini Hadamard ϵ-subdifferentiability for IVFs and $\mathbf{H}_\epsilon$-subgradient is given. These proposed concepts are observed to be more general than all the existing subdifferentials on IOPs (see [3–8, 98]) and also contains the set of gH-Dini Hadamard ϵ-subdifferential. A few relations between gH-Fréchet differentiability and gH-Dini Hadamard ϵ -subdifferentiability is given. Next, an important concept of $\mathbf{H}_\epsilon$-subgradient is given, which is based on the criterion of sponge of a set. Further, a variational interpretation of gH-Dini Hadamard ϵ-subdifferential based on the sponge of a set is discussed. Furthermore, the concept of ϵ-efficient solution followed by necessary and sufficient efficient conditions for finding an ϵ-efficient solution to an IOP with the gH-Dini Hadamard ϵ-subgradient of its objective function are given. An example of applying proposed results in a sparsity regularizer for IOPs is given.

In Chapter 5, a Newton method for the set optimization problems studied in [1] with a strong convexity assumption is discussed. The proposed method in this work exhibits a quadratic convergence near the optimal solution and works well for highly nonlinear objective functions. After that, its superlinear and quadratic convergence are analysed. Further, we provide the numerical performance of the proposed method in some test examples. Lastly, we compare the results of the proposed algorithm with the results of the steepest descent method presented in [1].

Chapter 6 discusses the quasi-Newton method for the considered set optimization problem. The well-definedness of the proposed algorithm with the existence of Armijo's step length condition and the boundedness of the norm of descent direction is given. After that, the convergence of the proposed quasi-Newton method is analysed. Further, we show the numerical implementation of our method with the help of suitable examples. Finally, we compare the results of the proposed algorithm with the results of the steepest descent method presented in [1].

Finally, in Chapter 7, we summarize the main conclusions and forecast potential directions for future research.

****