FUZZY MEASURE THEORY

Zhenyuan Wang
and
George J. Klir

Library of Congress Cataloging-in-Publication Data

Wang, Zhenyuan.
Fuzzy measure theory / Zhenyuan Wang and George J. Klir.
p. cm.
Includes bibliographical references and indexes.
ISBN 0-306-44260-4
1. Measure theory. 2. Fuzzy sets. I. Klir, George J., 1932-
II. Title.
QA312.W36 1992
515'.42--dc20 92-34026
CIP

ISBN 0-306-44260-4

Printed in the United States of America

FUZZY MEASURE THEORY

Preface

The principal purpose of this book is to present a comprehensive treatment of a relatively new mathematical subject referred to as fuzzy measure theory. The emergence of fuzzy measure theory (in the late 1970s) exemplifies a significant current trend in mathematics, a trend toward generalizations of existing mathematical concepts and theories. Each generalization enriches not only our insights but also our capabilities to properly model the intricacies of the real world.

Fuzzy measure theory is a generalization of classical measure theory. This generalization is obtained by replacing the additivity axiom of classical measures with weaker axioms of monotonicity and continuity. The development of fuzzy measure theory has been motivated by the increasing apprehensiveness that the additivity property of classical measures is in some application contexts too restrictive and, consequently, unrealistic.

Mathematical results presented in this book are almost exclusively those of Zhenyuan Wang. They are the results of more than a decade of concentrated research. Although most of the results were published in various journal articles and conference proceedings, some results are published here for the first time.

The book was written primarily as a text for a one-semester graduate or upper-division course. Such a course is suitable not only for programs in mathematics, where it might be offered at the junior or senior level, but also for programs in a host of other disciplines. Most notable among these disciplines, in which the utility of fuzzy measure theory is increasingly recognized, are systems, computer, information, and cognitive sciences, as well as artificial intelligence, quantitative management, mathematical social sciences, and some areas of engineering.

Although a solid background in calculus is required for understanding the material presented, the book is otherwise self-contained. Knowledge of classical measure theory, whose basic concepts and results are overviewed in App. A, is helpful but not essential. Relevant aspects of set theory, which play an important role in developing fuzzy measure theory, are introduced in Chap. 2. Basic concepts and results of fuzzy set theory are overviewed in App. B.

After a brief conceptual and historical discussion of fuzzy measure theory in Chap. 1, and relevant prerequisites from set theory in Chap. 2, the essence of fuzzy measure theory is covered in Chaps. 3–8. The applicability of the theory is then illustrated by simple examples in Chap. 9. Individual chapters are accompanied by notes, whose purpose is to provide the reader with relevant bibliographical and historical information, and exercises, by which the reader can test his or her comprehension of the material covered in each chapter.

The nine chapters of the book are supplemented with six appendices. As already mentioned, two of the appendices (A and B) provide the reader with relevant background in classical measure theory and fuzzy set theory; two appendices (C and D) are glossaries of key concepts and symbols; and two of them contain six reprinted articles that are significantly connected with the text: Three of the articles open new directions in fuzzy measure theory (App. E), and three of them describe significant applications of fuzzy measure theory (App. F). The following are copyright owners of the articles, whose permission to reproduce them is gratefully acknowledged: The Institute of Electrical and Electronic Engineers (IEEE), Elsevier, and Springer-Verlag.

The manuscript of the book was used three times (in Fall 1989, Spring 1990, and Fall 1991) as a text in a graduate course "Fuzzy Measure Theory," which was taught by Zhenyuan Wang at the Department of Systems Science, Thomas J. Watson School, of the State University of New York at Binghamton. We are grateful to all members of the faculty of systems science and the Dean of the Watson School, Lyle Feisel, for their support of this innovative course. The three offerings of the course, which were highly successful, demonstrated that the text is amenable to students who do not major in mathematics, in spite of its rigorous, mathematical treatment of the material. We are grateful to several graduate students, who took the course, for their suggestions and help with proofreading of the manuscript: Kevin Hufford, Cliff Joslyn, Yunxia Qi, Mark Scarton, Ute St. Clair, William Tastle, and Bo Yuan.

ZHENYUAN WANG
GEORGE J. KLIR

Binghamton, New York

Contents

CHAPTER 1

Introduction

Fuzzy measure theory, the subject of this text, is an offspring of *classical measure theory*. The latter has its roots in metric geometry, which is characterized by assigning numbers to lengths, areas, or volumes. In antiquity, this assignment process, or *measurement*, was first conceived simply as a comparison with a standard unit. Soon, however, *the problem of incommensurables* (exemplified by the problem of measuring the length of the diagonal of a square whose sides each measure one unit) revealed that measurement is more complicated than this simple, intuitively suggestive process. It became clear that measurement must inevitably involve infinite sets and infinite processes.

Prior to the emergence and sufficient development of the calculus, the problem of incommensurables had caused a lot of anxiety since there were no satisfactory tools to deal with it. *Integral calculus*, based upon the Riemann integral, which became well developed in the second half of the 19th Century, was the first tool to deal with the problem. Certain measurements that are contingent upon the existence of associated limits could finally be determined by using appropriate techniques of integration.

In the late 19th Century, there was a growing need for more precise mathematical analysis, induced primarily by the rapidly advancing science and technology. As a result, new questions regarding measurement emerged. Considering, for example, the set of all real numbers between 0 and 1, which may be viewed as points on a real line, mathematicians asked: When we remove the end points, 0 and 1, from this set, what is the measure of the remaining set (or the length of the remaining open interval on the real line)? What is the measure of the set obtained from the given set by removing some rational numbers, say number 1, 1/2, 1/3, 1/4,...? What is the measure of the set obtained by removing all rational numbers?

Questions like these and many more difficult questions were carefully examined by Émile Borel (1871–1956), a French mathematician. He developed a theory [Borel, 1898] to deal with these questions, which was an important step toward a more general theory that we now refer to as the *classical measure theory*.

Borel's theory deals with families of subsets of the set of real numbers that are closed under the set union of countably many sets and the set complement. He defines a measure that associates a positive real number with each bounded subset in the family, which, in the case of an interval, is exactly equal to the length of the interval. The measure is additive in the sense that its value for a bounded union of a sequence of pairwise disjoints sets is equal to the sum of the values associated with the individual sets.

Borel did not connect his theory with the theory of integration. This was done a few years later by Henri Lebesgue (1875–1941), another French mathematician. In a paper published in 1900, he defined an integral, more general than the Riemann integral, which is based on a generalized measure that subsumes the Borel measure as a special case. These generalized concepts of a measure and an integral (further developed in Lebesgue's doctoral dissertation published in 1902), which are now referred to as the *Lebesgue measure* and the *Lebesgue integral*, are the cornerstones of classical measure theory.

Perhaps the best nontechnical exposition of the motivation behind the Lebesgue measure and the Lebesgue integral, and a discussion of their physical meaning, was prepared by Lebesgue himself; it is available in a book edited by K. O. May, which also contains a biographical sketch of Lebesgue and a list of his key publications [Lebesgue, 1966].

Classical measure theory is closely connected with *probability theory*. A probability measure, as any other classical measure, is a set function that assigns 0 to the empty set and a nonnegative number to any other set, and that is additive. However, a probability measure requires, in addition, that 1 be assigned to the universal set in question. Hence, probability theory may be viewed as a part of classical measure theory.

The concept of a *probability measure* (or, simply, a *probability*) was formulated axiomatically in 1933 by Andrei N. Kolmogorov (1903–1987), a Russian mathematician [Kolmogorov, 1950]. This concept of probability is sometimes called a *quantitative* or *numerical probability* to distinguish it from other types of probability, such as *classificatory* or *comparative probabilities* [Fine, 1973; Walley and Fine, 1979; Walley, 1991]. Nevertheless, the term "probability theory," with no additional qualifications, refers normally to the theory based upon Kolmogorov's axioms.

After more than 50 years of the existence and steady development of the classical measure theory, the additivity property of classical measures became a subject of controversy. Some mathematicians felt that additivity is too restrictive in some application contexts. It is too restrictive to capture adequately the full scope of measurement. While additivity characterizes well many types of measurements under idealized, error-free conditions, it is not fully adequate to characterize most measurements under real, physical conditions, when measurement errors are unavoidable. Moreover, some measurements, involving, for example, subjective judgements or nonrepeatable experiments, are intrinsically nonadditive.

Numerous arguments have been or can be raised against the necessity and adequacy of the additivity axiom of probability theory. One such argument was presented by Viertl [1987]. It is based on the fact that all measurements are inherently fuzzy due to unavoidable measurement errors. Consider, for example, two disjoint events, A and B, defined in terms of adjoining intervals of real numbers, as shown in Fig. 1.1a. Observations in close neighborhoods (within a measurement error) of the endpoint of each event are unreliable and should be properly discounted, for example, according to the discount rate functions shown in Fig. 1.1a. That is, observations in the neighborhood of the end-points should carry less evidence than those outside them. When measurements are taken for the union of the two events, as shown in Fig. 1.1b, one of the discount rate functions is not applicable. Hence, the same observations produce more evidence for the single event $A \cup B$ than for the two disjoint events A and B. This implies that the degree of belief in $A \cup B$ (probability of $A \cup B$) should be greater than the sum of the degrees of belief in A and B (probabilities of A and B). The additivity axiom is thus violated.

The earliest challenge to classical measure theory came from a theory proposed by a French mathematician G. Choquet [1954], for which he coined the name *theory of capacities.* A Choquet capacity is a set function that associates a real number (not necessarily nonnegative) with each subset of the universal set employed and is continuous and monotonic with respect to set inclusion.

Although the Choquet theory of capacities is a broad framework, encompassing various types of nonadditive measures, it is too general for

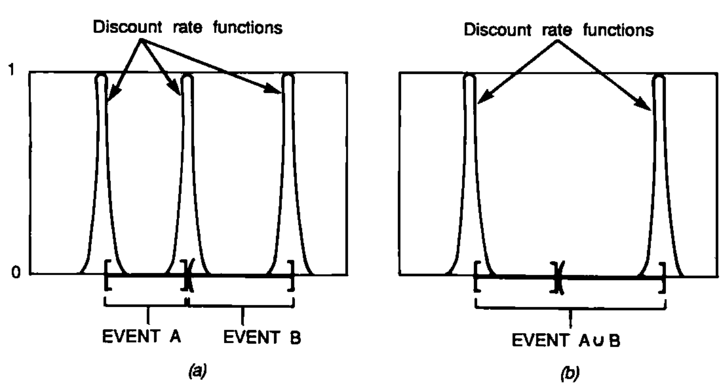

Figure 1.1. An example illustrating the violation of the additivity axiom of probability theory.

most practical applications. Other, more useful types of nonadditive measures emerged later in more specific contexts.

In the context of probability theory, a generalized theory based upon two types of nonadditive measures was originated by Dempster [1967] and, later, fully developed by Shafer [1976]. These types of measures are obtained by replacing the additivity requirement of probability measures with either a superadditivity requirement or a subadditivity requirement. The superadditive measures, which are also upper semicontinuous, are usually called *belief measures.* The subadditive measures, which are also lower semicontinuous, are usually referred to as *plausibility measures.*

Given a measure of either of the two types, it induces a unique measure of the other type. Taken together, belief and plausibility measures form a theory that is usually called the *Dempster–Shafer theory* or *evidence theory.* It stems from the initial work by Dempster [1967] that belief and plausibility measures have a natural interpretation as lower and upper probabilities, respectively. Since belief measures are always smaller than or equal to the corresponding plausibility measures, the intervals between belief and plausibility values may be viewed as ranges of admissible probabilities. The Dempster–Shafer theory may thus be viewed as a theory that is capable of dealing with interval-valued probabilities.

Properties of belief and plausibility measures are studied in Sec. 3.4. While current literature dealing with these measures (including the classical book by Shafer [1976]) is predominantly based on the assumption that the universal set on which the measures are defined is finite, our treatment of the subject is not restricted in this way.

Another theory based upon nonadditive measures, referred to as *possibility theory,* emerged from the concept of a *fuzzy set,* which was proposed by Zadeh [1965]. A fuzzy set is a set whose boundary is not sharp. That is, the change from nonmembership to membership is gradual rather than abrupt. This gradual change is expressed by a *membership grade function* of the fuzzy set, which assigns to each individual of the universal set a value in the unit interval [0, 1]. This value represents the grade of membership of the individual in the fuzzy set. A fuzzy set is called *regular* if the maximum of its membership grade function is 1. To distinguish between fuzzy and nonfuzzy sets, the latter are usually referred to as *crisp sets.*

Given a normal fuzzy set, Zadeh [1978] defines a *possibility distribution function* associated with the set as numerically equal to its membership grade function. Then, he defines a *possibility measure* by taking the supremum of the possibility distribution function in each crisp set of concern.

It turns out that possibility measures emerge not only from the context of fuzzy sets, but also from the context of evidence theory, as special plausibility measures [Dubois and Prade, 1988; Klir and Folger, 1988]. It is this latter context that is employed in our study of possibility measures and the associated *necessity measures* (special belief measures) in Sec. 3.5.

The two principal themes of this text, *fuzzy measures* and *fuzzy integrals*, also emerged in the context of fuzzy sets, as suggested by their names. These concepts were envisioned by Sugeno [1974, 1977] in his efforts to compare membership grade functions of fuzzy sets with probabilities. Since no direct comparison is possible, Sugeno conceived of the generalization of classical measures into fuzzy measures as an analogy of the generalization of classical (crisp) sets into fuzzy sets.

Fuzzy measures, according to Sugeno, are obtained by replacing the additivity requirement of classical measures with weaker requirements of *monotonicity* (with respect to set inclusion) and *continuity*. The requirement of continuity was later found to be still too restrictive and was replaced with a weaker requirement of *semicontinuity*. In fact, belief and plausibility measures, as well as necessity and possibility measures, are only semicontinuous. In this text, we cover both continuous and semicontinuous fuzzy measures.

Similarly as Choquet's capacities, fuzzy measures are too loose to allow us to develop a theory that would capture their full generality and, yet, were of pragmatic utility. On the other hand, some special types of fuzzy measures, such as superadditive and subadditive measures, appear to be unnecessarily restrictive in some application contexts. These considerations led to a more systematic investigation of useful structural characteristics of set functions, primarily by Wang [1984a, 1985a], as presented in Chap. 5. These characteristics are essential for capturing mathematical properties of measurable functions on fuzzy measure spaces (Chap. 6), and that, in turn, is requisite for developing a *theory of fuzzy integrals* (Chap. 7), as well as a more general *theory of pan-integrals* (Chap. 8).

There have been many additional developments pertaining to various aspects of fuzzy measure theory (Chaps. 3–8) that we do not deem necessary to cover in this Introduction. Since most of these developments are rather technical and involve special terminology, we leave their historical and bibliographical coverage to Notes accompanying the individual chapters.

Notes

1.1. An overview of relevant concepts and results of classical measure theory is given in Appendix A. For further study, we recommend the classic text by Halmos [1967]. An excellent text on classical measure theory by Billingsley [1986] is recommended to readers that are interested particularly in probability measures.

1.2. Among many other books on classical measure theory, let us mention a few that are significant in various respects. The book by Caratheodory [1963], whose original German version was published in 1956, is one of the earliest and most highly influential books on classical measure theory. Books by Temple [1971] and Weir [1973] provide pedagogically excellent introductions to classical measure theory; they require only some basic knowledge of calculus and algebra as prerequisites. The book by Constantinescu and Weber [1985], suitable for a mathematically mature reader, attempts to unify abstract and topological approaches. Other valuable books are by Berberian [1965], Kingman and Taylor [1966], and Wheeden and Zygmund [1977]. The book by Faden [1977] is an extensive treatise on the use of measure theory, particularly in the area of economics, which also contains a good introduction to measure theory itself.

1.3. An excellent discussion of the various shortcomings of additive (i.e., precise) probabilities and the reasons why nonadditive (i.e., imprecise) probabilities are needed to overcome these shortcomings is presented by Walley [1991]. It is shown by Klir [1989] that classical (additive) probability measure can capture only one of several types of uncertainty that can clearly be recognized when the additivity property is abandoned. A paper by Billot [1992] contains an interesting historical overview of the use of nonadditive probabilities in economics.

1.4. The history of classical measure theory and Lebesgue's integral is carefully traced in a fascinating book by Hawkins [1975]. He describes how modern mathematical concepts regarding these theories (involving concepts such as a function, continuity, convergence, measure, integration, and the like) developed (primarily in the 19th Century and the early 20th Century) through the work of many mathematicians, including Cauchy, Fourier, Borel, Riemann, Cantor, Dirichlet, Hankel, Jordan, Weierstrass, Volterra, Peano, Lebesgue, Radon, and many others.

1.5. For the history of probability theory, we recommend a book by Hacking [1975] and a paper by Shafer [1978]. From the standpoint of fuzzy measure theory, it is most interesting that Bernoulli (1654–1705) and later Lambert (1728–1777) were already concerned with a calculus of probabilities that are not additive and, consequently, are imprecise. Their work, unfortunately, was forgotten for more than two centuries.

CHAPTER 2

Required Background in Set Theory

2.1. Set Inclusion and Characteristic Function

Let X be a nonempty set. Unless otherwise stated, all sets that we consider are subsets of X. X is called the *universe of discourse*. The elements of X are called *points*. X may contain finite, countably infinite, or uncountably infinite number of points. A set that consists of a finite number of points $x_1, x_2, \ldots, x_n$ (or, a countably infinite number of points $x_1, x_2, \ldots$) may be denoted by $\{x_1, x_2, \ldots, x_n\}$ ($\{x_1, x_2, \ldots\}$, respectively). A set containing no point is called the *empty set*, and is denoted by $\varnothing$.

If x is a point of X and E is a subset of X, the notation

$$x \in E$$

means that x belongs to E, i.e., x is an element of E; and the statement that x does not belong to E is denoted by

$$x \notin E.$$

Thus, for every point x of X, we have

$$x \in X$$

and

$$x \notin \varnothing.$$

A set of sets is called a *class*. If E is a set and $\mathscr{C}$ is a class, then

$$E \in \mathscr{C}$$

means that the set E belongs to the class $\mathscr{C}$.

If, for each x, $\pi(x)$ is a proposition concerning x, then the symbol

$$\{x \mid \pi(x)\}$$

denotes the set of all those points x for which $\pi(x)$ is true, that is,

$$x_0 \in \{x \mid \pi(x)\} \quad \Leftrightarrow \quad \pi(x_0) \text{ is true.}$$

By replacing point x with set E, such a symbol may be used to indicate a class. For example,

$$\{E \mid x \in E\}$$

denotes the class of those sets that contain the point x.

Example 2.1. Let $X = \{1, 2, \ldots\}$. The set $\{x \mid x$ is odd and less then $10\}$ is $\{1, 3, 5, 7, 9\}$.

Example 2.2. Let X be the set of all real numbers, which is often referred to as the real line or one-dimensional Euclidean space. The class $\{(a, b) \mid -\infty < a < b < \infty\}$ is the class consisting of all open intervals on the real line.

If E and F are sets, the notation

$$E \subset F \quad \text{or} \quad F \supset E$$

means that E is a subset of F, i.e., every point of E belongs to F. In this case, we say that F *includes* E, or that E is included by F. For every set E, we have

$$\emptyset \subset E \subset X.$$

Two sets E and F are called *equal* iff

$$E \subset F \quad \text{and} \quad F \subset E;$$

that is, they contain exactly the same points. This is denoted by

$$E = F.$$

The symbols $\subset$ or $\supset$ also may be used for classes. If $\mathscr{E}$ and $\mathscr{F}$ are classes, then

$$\mathscr{E} \subset \mathscr{F}$$

means that every set of $\mathscr{E}$ belongs to $\mathscr{F}$, that is, $\mathscr{E}$ is a subclass of $\mathscr{F}$.

If $E_1, E_2, \ldots, E_n$ are nonempty sets, then

$$E = \{(x_1, x_2, \ldots, x_n) \mid x_i \in E_i, \quad i = 1, 2, \ldots, n\}$$

is called an *n-dimensional product set* and is denoted by

$$E = E_1 \times E_2 \times \cdots \times E_n.$$

Similarly, if $\{E_t \mid t \in T\}$ is a family of nonempty sets, where T is an infinite index set, then

$$E = \{x_t, t \in T \mid x_t \in E_t \text{ for each } t \in T\}$$

is called an *infinite-dimensional product set.*

Example 2.3. Let X_1 and X_2 be one-dimensional Euclidean spaces. Then $X = X_1 \times X_2 = \{(x_1, x_2) | x_1 \in (-\infty, \infty),\ x_2 \in (-\infty, \infty)\}$ is the two-dimensional Euclidean space. The set $\{(x_1, x_2) | x_1 > x_2\}$ is a half (open) plane under the line $x_2 = x_1$, while the set $\{(x_1, x_2) | x_1^2 + x_2^2 < r^2\}$ is the open circle centering at the origin with a radius r, where $r > 0$.

Example 2.4. Let $X_t = \{0, 1\}$, $t \in \{1, 2, \ldots\}$. The space

$$X = X_1 \times X_2 \times \cdots \times X_n \times \cdots = \{(x_1, x_2, \ldots, x_n, \ldots) | x_t \in \{0, 1\} \text{ for each } t \in \{1, 2, \ldots\}\}$$

is an infinite-dimensional product space. Each point $(x_1, x_2, \ldots, x_n, \ldots)$ in this space corresponds to the binary number $0 . x_1 x_2 \ldots x_n \ldots$ in $[0, 1]$. Such a correspondence is not one to one, but it is onto.

If E is a set, the function χ_E, defined for all $x \in X$ by

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E, \end{cases}$$

is called the *characteristic function* of the set E. The correspondence between sets and their characteristic functions is one to one, that is,

$$E = F \quad \Leftrightarrow \quad \chi_E(x) = \chi_F(x), \quad \forall x \in X.$$

It is easy to see that

$$E \subset F \quad \Leftrightarrow \quad \chi_E(x) \leq \chi_F(x), \quad \forall x \in X,$$

and that

$$\chi_X \equiv 1, \qquad \chi_\varnothing \equiv 0.$$

2.2. Operations on Sets

Let $\mathscr{C}$ be any class of subsets of X. The set of all those points of X that belong to at least one set of the class $\mathscr{C}$ is called the *union* of the sets of $\mathscr{C}$. This is denoted by

$$\bigcup \mathscr{C}.$$

If to every t of a certain index set T there corresponds a set E_t, then the union of the sets of class

$$\{E_t | t \in T\}$$

may be also denoted by

$$\bigcup_{t \in T} E_t \quad \text{or} \quad \bigcup_t E_t.$$