Literature Reviews

Reviewed by S. Chakraverty

E-mail: chakravertys@nitrkl.ac.in

Please send the PDF files, or copies of the books or papers you want to see reviewed here, by e-mail, or by regular mail to me.

1 Some Results on Boundary Value Problems for Fuzzy Differential Equations with Functional Dependence
by: Juan J. Nieto and Rosana Rodríguez-López
Published 2013
Fuzzy Sets and Systems, Vol. 230, pp.92–118, 2013
27pp, ISSN: 0165-0114

As truly said in the above paper that the interest in the investigation of practical problems modelled as delay or impulsive differential equations are increasing significantly in recent years. Moreover, the actual essence of the physical problems needs to study the problem when we consider the system as uncertain. Due to the non-decreasing character of the level sets' diameter of the solutions to fuzzy differential equations from the point of view of Hukuhara differentiability, the study of boundary value problems presents certain difficulties which the authors tried to overcome by introducing impulses. As such, in the above paper, the authors solved explicitly an impulsive linear first order fuzzy differential equation subject to boundary value conditions. Then, they derived conclusions with proofs on the existence of a unique Hukuhara-differentiable solution to a boundary value problem for a class of fuzzy functional differential equations.

2 How to Determine Basis Stability in Interval Linear Programming
by: Milan Hladík
Published 2014
Optimization Letters, Vol. 8, No. 1, pp.375–389, 2014
15pp, ISSN: 1862-4472 (print version),
ISSN: 1862-4480 (electronic version)

Linear programming problems have been of interest due to its practical applications. Real-life data are often subject to uncertainties and measurement errors, so interval linear programming (ILP) was introduced to tackle these troubles. Diverse methods have been developed for solving ILP problems. As mentioned in this paper that some of the methods are based on interval arithmetic and extensions of the simplex algorithm to the case of interval data, while others use a direct inspection.

The paper rightly mentions that the basic tasks in ILP such as calculating the optimal value bounds or set of all possible solutions may be computationally very expensive. However, if some basis stability criterion holds true then the problems become easier to solve. In the above paper, the author proposed a new algorithm for basis stability checking. Even though the complexity is (and due to NP-hardness probably must be) exponential, it runs quickly in many cases and outperforms the known methods as shown by example. The algorithm works also for non-degenerate basis stability. Finally, the author writes about some open problems in non-standard form (inequalities, mixed equations and inequalities, etc.).

Literature Review

Reviewed by Hend Dawood

E-mail: hend.dawood@sci.cu.edu.eg

Computing Statistics under Interval and Fuzzy Uncertainty: Applications to Computer Science and Engineering
by: H.T. Nguyen, V. Kreinovich, B. Wu and G. Xiang
Published 2012
Studies in Computational Intelligence, Vol. 393
by Springer Verlag
Heidelberger Platz 3, 14197 Berlin, Germany, 432pp
ISBN: 978-3-642-24904-4

Scientific knowledge is not perfect exactitude: it is learning with uncertainty. Computing under uncertainty is thus of great importance in both fundamental research and practical problems and many approaches have been developed to deal with uncertainty and get reliable knowledge about the real world. Within this scope, this interesting book is meant to provide an exposition of the contemporary approaches to computing under uncertainty. A great deal is said in this rich and rewarding book which draws mainly on research papers conducted by the authors during the previous decade, and perhaps this is the reason why it is slightly fragmented. However, it is a work of great interest and importance. As the authors put it in the preface, their aim in this book is "to compute statistics under interval and fuzzy uncertainty". In more precise terms, their objective is to estimate the difference

$$\begin{aligned}\Delta C &= \tilde{C} - C \\ &= C(\tilde{x}_1, ..., \tilde{x}_n) - C(x_1, ..., x_n),\end{aligned}$$

under the assumption that they only have interval or fuzzy estimates of the differences $\Delta x_i = \tilde{x}_i - x_i$.

Before turning to the strengths and weaknesses, let me first describe the content and organisation of the book. The book is structured in five parts and 47 chapters.

Part I, 'Computing statistics under interval and fuzzy uncertainty: formulation of the problem and an overview of general techniques which can be used for solving this problem', contains the formulation of the main problem of the book along with an overview of the general techniques available for solving this problem. From the available techniques, the authors describe interval and fuzzy computations with explaining the interesting *reduction* of computing statistics under fuzzy uncertainty to computing statistics under interval uncertainty. The authors end up this part by discussing the most practically useful statistical characteristics.

In the first chapter, 'Formulation of the problem', the authors discuss the importance of statistical analysis and computing statistical characteristics. The authors formulate the problem as follows.

Let C be the sample-based estimate for the *desired* statistical characteristic and c be its (unknown) *actual* value. By definition, we have

$$C \stackrel{def}{=} C(x_1, ..., x_n).$$

There are many techniques for estimating the difference $C - c$. The existing statistical estimates for the difference $C - c$ are based on the assumption that for each object i, the exact value x_i of the corresponding quantity is known. In practice, the values x_i come from *measurements* or from *expert estimates*. Therefore, what we want to know is how accurate are the estimates $\tilde{C} \stackrel{def}{=} C(\tilde{x}_1, ..., \tilde{x}_n)$, i.e., what is the size of the quantity $\tilde{C} - c$. Statistical analysis enables us to estimate $C - c$, with $C \neq \tilde{C}$. Thus, to estimate the value of the quantity $\tilde{C} - c$ we need to estimate the difference $\tilde{C} - C$. Estimates for the difference $\tilde{C} - C$ are also known in statistics, for the case when we know the *exact* probability distribution of all the measurement errors Δx_i. Sometimes we have no information about the probabilities of Δx_i; what we only have is the *upper bound* on the measurement error.

The manufacturers of a measuring device usually provide an upper bound Δ_i for the possible measurement errors, i.e., with the bound Δ_i for which it is guaranteed that $|\Delta x_i| \leq \Delta_i$. Therefore, we can guarantee that the actual (unknown) value of the desired quantity belongs to the interval

$$\mathbf{x}_i \stackrel{def}{=} [\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i].$$

In case of *interval uncertainty*, we do not know the actual value x_i of the i^{th} quantity. Instead, we know the *interval* $[\tilde{x}_i - \Delta_i, \tilde{x}_i + \Delta_i]$ that contains x_i. For some measuring instruments, *noise* and *discretisation* errors are negligible, and the *detection* error is the main source of measurement uncertainty. In some other cases, noise and detection limits are negligible, and discretisation is the main source of measurement errors. In many numerical computations, there is another important source of interval uncertainty: *roundoff errors*. In these cases, all we know about the result z is that it belongs to the interval $[\tilde{z} - \varepsilon, \tilde{z} + \varepsilon]$, where ε is the guaranteed upper bound on the roundoff accuracy. Here, the authors introduce a generalised claim that in comparison with measurement errors, roundoff errors can be safely ignored.

In case of *fuzzy uncertainty*, we, in some practical situations, only have expert estimates for the inputs x_i. Sometimes, experts provide guaranteed bounds on x_i, and even the probabilities of different values within these bounds. However, such cases are rare. Usually, the experts' opinions about the uncertainty of their estimates are described by imprecise '*fuzzy*' words from natural language.

In chapter 2, 'Computing statistics under probabilistic and interval uncertainty: a brief description', the authors repeat again what they mentioned in chapter 1, about computing under probabilistic uncertainty (with *known* probability distributions for measurement errors) and interval uncertainty, with slightly more details.

Chapter 3, 'Computing statistics under fuzzy uncertainty: formulation of the problem', is devoted to explaining the problem of computing with experts' fuzzy information. The authors discuss why we need to process such fuzzy information with clarifying the basic idea underlying such computation. Towards explaining the details, they introduce concepts like *degrees of belief*, *membership function* $\mu_S(x_i)$ (to describe the degree to which x_i is consistent with the statement *S*), *t-norms* and *t-conorms* ('*and*' and '*or*' operations to compute the degrees of belief for composite statements). The chapter ends with studying what is known as 'Zadeh's extension principle', by deriving the main formula that describes extending a statistic with real inputs to fuzzy inputs. Finally, a proof is introduced that the only idempotent t-norm is the *min* operation and the only idempotent t-conorm is the *max* operation.

The fourth chapter, 'Computing under fuzzy uncertainty can be reduced to computing under interval uncertainty', introduces the interesting *reduction* of fuzzy computations to interval computations. It starts by defining the concept of *α-cut*; the set of possible elements for some threshold $\alpha \in (0, 1]$. Then, it explains the relation between fuzzy numbers and intervals in terms of these α-cuts. The chapter continues to reformulate the '*and*' and '*or*' operations in terms of sets and α-cuts. Finally, it explains the main relation between fuzzy and interval computations. Assuming *continuity* of the desired statistic $y = C(x_1, \ldots, x_n)$ and continuity of the membership functions $\mu_i(x_i)$ with compact interval α-cuts, the α-cut for the output can be derived from the α-cuts for the inputs as the range of the statistic function y with inputs vary through the corresponding α-cuts, and hence the computation reduces to the problem of interval computations.

There follows chapter 5, 'Computing under interval uncertainty: traditional approach based on uniform distributions', on computing statistics for the statistically independent variables $x_1, \ldots, x_n$ where the error of measurement for each variable, Δx_i, is *uniformly* distributed on the interval $[-\Delta_i, \Delta_i]$. With this approach, the authors give an example that shows how the computations with a uniform distribution could be *misleading*.

In chapter 6, 'Computing under interval uncertainty: when measurement errors are small', the authors explain the case with interval uncertainty when the measurement errors Δx_i are relatively *small*. The *linearisation* technique is introduced to simplify the computations. In this technique, a function y is expanded using only the linear terms of its Taylor series, yielding

$$\begin{aligned} y &= f\left(x_i, \ldots, x_n\right) \\ &= f\left(\tilde{x}_1 - \Delta x_i, \ldots, \tilde{x}_n - \Delta x_n\right) \\ &\approx f\left(\tilde{x}_1, \ldots, \tilde{x}_n\right) - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i. \end{aligned}$$

Hence,

$$\begin{aligned} \Delta y &= \tilde{y} - y \\ &\approx f\left(\tilde{x}_1, \ldots, \tilde{x}_n\right) - f\left(\tilde{x}_1, \ldots, \tilde{x}_n\right) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i, \end{aligned}$$

so

$$\Delta y \approx \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \Delta x_i .$$

From the last formula, it is easy to conclude that the maximum value of Δy, denoted Δ, is computed by

$$\Delta = \sum_{i=1}^{n} \left| \frac{\partial f}{\partial x_i} \right| \Delta_i ,$$

because $\Delta x_i \in [-\Delta_i, \Delta_i]$. The smallest possible value is therefore $-\Delta$. Hence, $\Delta y \in [-\Delta, \Delta]$ which yields

$$y \in \left[\tilde{y} - \Delta, \tilde{y} + \Delta \right].$$

Towards computing Δ, the authors discuss the usual method, for computing the partial derivatives $\frac{\partial f}{\partial x_i}$, that depends on finding the limit as $h_i \to 0$. The computation is simplified by choosing a small value $h_i = \Delta_i$. This technique takes $n + 1$ times calling f. So, the authors provide at the end of this chapter a much more faster computation using *Monte-Carlo* techniques.

In chapter 7, 'Computing under interval uncertainty: general algorithms', the authors introduce the theory of intervals and discuss why we need interval computations. Limitations of traditional numerical methods are discussed. A brief history of interval arithmetic is introduced. The basic operations of interval arithmetic; *addition*, *subtraction*, *multiplication*, and *reciprocal*; are discussed. Towards computing with intervals, they introduce the method of 'naïve' interval computation with studying its limitations. Also, other techniques are provided including the *mean value*, *bisection*, and *Taylor* techniques, along with discussing the effectiveness of these techniques compared with the naive method.

Chapter 8, 'Computing under interval uncertainty: computational complexity', provides a good discussion of the computational complexity for the problem of estimating the range of a function with interval arithmetic. For a *linear* function, it is shown that the range can be computed in linear time. However, for a *quadratic* function, the authors state that it is an NP-hard problem. The remaining of the chapter delves into the definitions and characteristics of NP-hard problems, compared with P, NP, and NP-complete problems.

Since, under interval uncertainty, it is not always possible to design a polynomial-time algorithm for *all* statistics, the authors discuss the necessity of selecting the *practically useful* statistical characteristics. Towards this, chapter 9, 'Towards selecting appropriate statistical characteristics: the basics of decision theory and the notion of utility', introduces the basics of making decisions about the most appropriate statistical characteristics. Chapter 10, 'How to select appropriate statistical characteristics', discusses the standard way of making a decision by selecting the action with the *largest* possible expected utility (gain). The authors then explain how to compute the expected value of *smooth* and *non-smooth* utility functions. At the end of this chapter, the authors conclude that the most appropriate statistical characteristics are the *moments*, *mean*, *variance*, *median*, and other characteristics of the distribution.

After providing an overview of the general techniques in part I, there follows four parts on computational algorithms, the quality of the input data, typical applications, and other types of input uncertainty.

Assuming that the interval and fuzzy inputs are of *good quality* and consistent with the actual (unknown) values, part II, 'Algorithms for computing statistics under interval and fuzzy uncertainty', studies general algorithms for computing various *useful* statistical characteristics including *mean*, *median*, *variance*, *outlier thresholds*, *higher moments*, *covariance*, *correlation*, *expected value*, and *entropy*.

Part III, 'Towards computing statistics under interval and fuzzy uncertainty: gauging the quality of the input data', explains how to gauge the quality of the input data by measuring the *reliability* and *accuracy* of the input data. Reliability checking is an NP-hard problem in general. Based on the notion of utility, chapter 30 is devoted to measuring the divergence from the desired accuracy.

After discussing various techniques for computing statistics under interval and fuzzy uncertainty, part IV, 'Applications', delves into different practical applications of these techniques. The discussed fields of applications include *bioinformatics*, *computer science*, *information management*, *signal processing*, *computer engineering*, *mechanical engineering*, and *geophysics*.

In the fifth and final part, 'Beyond interval and fuzzy uncertainty', the authors develop techniques for computing statistics under other types of input uncertainty. This part studies general cases such as when the inputs satisfy a specific constraint, discontinuous operations, more general fuzzy-valued degrees, and so forth.

This book displays many strengths. A major one is that although the book is technically sophisticated, the discussion is very clear and readable with numerous examples to clarify the matters for the non-specialist. For the purpose of legibility, proofs are introduced at the end of each chapter which makes a suitable space for the reader to understand the subject without going into the complications of the detailed mathematical proofs. Moreover, the text is self-contained and the authors successfully packed a great amount of good material in the body of a single book, including the underlying mathematical concepts, computational algorithms, impressive discussions of a wide range of applications, contemporary research trends, and a look ahead.

Now for the problems. The text draws mainly on research papers conducted by the authors, but as a book, the main concern is that it is more or less mixed and fragmented and I sense each author wrote his chapters completely independently. The arrangement of the book is confusing. Cross-references are missing and there are no obvious sections or subsections for the book chapters. Many chapters repeat what is already stated in previous chapters and the chapters are not connected together to make a coherent structure. Chapters' titles are very long and the main part of the titles of some chapters is repeated over and over again along many chapters. For example, chapters 5 through 8 contain 'Computing under interval uncertainty' as a part of their titles. Most of us are guilty of typos and mistakes, but this book contains relatively many typos and other minor oversights that may lead to confusion. Unfortunately, the editor neither took care of typos nor pulled out the various repetitions and redundancies.

My second concern has to do with how the authors compare roundoff errors to measurement errors. On page 8, the authors say: "in comparison with measurement errors, roundoff errors can be safely ignored" and "The additional reason why roundoff errors can be safely ignored is that our objective is to estimate the values of the statistical characteristics like mean, variance, etc., based on the sample data...". While roundoff errors can cause disasters in the result, the authors simply ignore them claiming that they are small compared with measurement errors without sufficient mathematical justification to convince the reader.

There are some other technical and notational concerns that should be brought to attention. On page 7, the authors use the expression "the actual computational result z" which makes no sense since the actual (exact) result z is not usually the same as the computational result $\tilde{z}$ which they use to express the interval of confidence $[\tilde{z}-\varepsilon, \tilde{z}+\varepsilon]$. In chapter 3, the authors first use the symbol S_i to denote the *statement* of the expert's knowledge about the i^{th} input and they even use it to express the *membership function* $\mu_S(x_i)$ (the degree of belief or confidence). However, on page 13 of the same chapter they deploy another symbol A_i to denote the same notion, which is very confusing. Moreover, through chapter 3, the authors use different variations for the notation that expresses the degree of belief. These include $\mu_S(x_i)$, $\mu_i(x_i)$, $\mu(y)$, and d_i. This *multi-denoting*, of course, confuses the reader and leads to ambiguity.

There are a few parts of the book that lack clarifications. Among these, chapter 4, one of the most important chapters, needs a more detailed explanation with clarifying examples for the reduction from fuzzy computations to interval computations. Furthermore, chapters 7 and 8 are misplaced and they should be placed before any chapter that talks about interval uncertainty. This confusing order seems like using the theory before discussing what the theory is and what are its limitations.

To sum up, this book, apart from some minor concerns, is written by masters of their craft who have admirably succeeded to pack an amazing amount of valuable material within the same covers. The authors provide the reader with clear mathematical foundations, detailed algorithms, impressive applications, as well as a very helpful bibliography. This book can be confidently recommended to any student, researcher, or practitioner who wishes to go beyond the fundamentals and delve into the rich techniques and applications of computing under uncertainty.