University of Granada

Doctoral Program in Mathematical and Applied Statistics

Doctoral Thesis

Statistical analysis with fuzzy data of the Chinese economy

Gerardo Alfonso Pérez

Granada, 2017

Statistical analysis with fuzzy data of the Chinese economy

A DISSERTATION

SUBMITTED TO THE DOCTORAL SCHOOL OF SCIENCES,
TECNOLOGIES AND ENGINEERING

FULFILLMENT OF THE REQUIREMENTS

FOR THE DOCTORAL DEGREE

(PH.D.)

IN MATHEMATICAL AND APPLIED STATISTICS

Gerardo Alfonso Pérez

ADVISORS

Dr. Concepción Beatriz Roldán López de Hierro

Department of Statistics and Operations Research

Dr. Antonio Francisco Roldán López de Hierro

Department of Quantitative Methods for Economics and Business

University of Granada

Editor: Universidad de Granada. Tesis Doctorales
Autor: Gerardo Alfonso Pérez
ISBN: 978-84-9163-882-7
URI: http://hdl.handle.net/10481/51617

The doctoral candidate **Gerardo Alfonso Pérez** and the thesis supervisors: **Concepción Beatriz Roldán López de Hierro y Antonio Francisco Roldán López de Hierro**

Guarantee, by signing this doctoral thesis, that the work has been done by the doctoral candidate under the direction of the thesis supervisors and, as far as our knowledge reaches, in the perfomance of the work, the rights of other authors to be cited (when their results or publications have been used) have been respected.

Place and data:

Granada 29 June 2017

Doctoral candidate:

Gerardo Alfonso Pérez

Thesis supervisors:

Dra. D^{a}. Concepción B. Roldán and Dr. D. Antonio F. Roldán

Abstract

Different sources of imprecision and uncertainty are encountered in practical problems and, thus, many elements may need to be imprecisely observed, defined or treated. In this setting, one of the most successfully applied techniques to describe possible relationships between fuzzy variables is the regression methodology. In this dissertation, we introduce a fuzzy regression procedure involving a class a fuzzy numbers defined by some level sets called finite fuzzy numbers. We give a characterization of the image of a finite fuzzy number in terms of the extremes of its level sets and we present a parametric family of fuzzy semidistances between them that let us to consider a total fuzzy error of estimation (described as a fuzzy sum of squares of residuals in particular cases). The estimation process consist in finding a regression model that minimizes, in a fuzzy sense, such fuzzy error. Although spreads of finite fuzzy numbers can take some values very close to zero, which complicate the task of finding nonnegative models, the presented algorithm is able to guarantee that the predicted response is a fuzzy variable. Finally a numerical example based on fuzzy economic data of China is given to illustrate the use of the proposed method.

Keywords

Finite fuzzy numbers, Fuzzy random variable, Fuzzy distance measure, Fuzzy regression

Acknowledgements

To my parents, wife and children. Thank you for making everyday a wonderful experience.

I would like to thank also my supervisors, Dra. D^a Concepción B. Roldán López de Hierro and Dr. D. Antonio F. Roldán López de Hierro, for their patience and guidance in this dissertation, and Shenyin & Wanguo Co., Ltd for its support to this research.

Dedicado a mis padres, esposa e hijos. Gracias por convertir cada día en una experiencia fantástica.

Quiero también dar las gracias a mis directores, la Dra. D^a Concepción B. Roldán López de Hierro y el Dr. D. Antonio F. Roldán López de Hierro, por su paciencia y comprensión a lo largo de esta tesis doctoral, y a Shenyin & Wanguo co. Ltd, por el apoyo recibido en este proyecto.

International Journal Publications

G. Alfonso , A.F. Roldán-López-de-Hierro, C. Roldán. A fuzzy regression model based on finite fuzzy numbers and its application to real-world financial data,*Journal of Computational and Applied Mathematics.* 318 (2017) 47–58.

2016 Impact factor: 1.357 (Position 63/255 in 'Mathematics, Applied'), Q1.

International Conference Publications

A.F. Roldán-López De Hierro, C.B. Roldán-López De Hierro, J. Martínez Moreno, G. Alfonso-Pérez. A determination coefficient in a fuzzy random environment. 6th International Conference on Approximation Methods and Numerical Modelling in Environment and Natural Resources, Proceedings of MAMERN VI, pp. 575–584, 2015.

C.B. Roldán-López De Hierro, A.F. Roldán-López De Hierro, G. Alfonso-Pérez. A fuzzy regression model based on finite fuzzy numbers. 6th International Conference on Approximation Methods and Numerical Modelling in Environment and Natural Resources, Proceedings of MAMERN VI, pp. 585–596, 2015.

A.F. Roldán-López De Hierro, C.B. Roldán-López De Hierro, J. Martínez Moreno, G. Alfonso-Pérez. On a new methodology for ranking fuzzy numbers and its application to real economic data. 15th International Conference on Computational and Mathematical Methods in Science and Engineering, Proceedings of CMMSE15, pp. 973–983, 2015.

A.F. Roldán-López De Hierro, G. Alfonso-pérez, C.B. Roldán-López De Hierro. A novel fuzzy regression model based on finite fuzzy numbers and its application to the stock market of China. 16th International Conference Computational and Mathematical Methods in Science and Engineering, Proceedings of CMMSE16, 1108, 2016.

Contents

Resumen

En muchas situaciones de la vida real, la información asociada a algunos experimentos puede medirse con gran precisión. No obstante, en otras ocasiones, sólo es imprecisa, subjetiva o evaluable. Recientes investigaciones científicas han puesto de manifiesto la creciente inestabilidad de los sistemas económicos y empresariales, y la necesidad de crear nuevos instrumentos para representar y gestionar dicha inestabilidad. Elaborar modelos económicos que sean capaces de explicar la realidad observada requiere la capacidad de gestionar diferentes fuentes de imprecisión que surgen de manera natural por la falta de conocimiento preciso. Los tipos de incertidumbre y sus fuentes son muy variados y pueden clasificarse atendiendo a diversos puntos de vista de diferentes formas. Por ejemplo, podemos considerar la aleatoriedad (o incertidumbre estocástica) en el resultado de un experimento aleatorio (que se formaliza matemáticamente a través del Cálculo de Probabilidades) o la imprecisión (o incertidumbre no estocástica) que surge debido a la falta de información sobre un proceso concreto, o la asociada a un proceso de agregación, de transmisión de datos, etc.

Desde este punto de vista, hemos de admitir que las magnitudes tanto físicas como socioeconómicas que observamos en la realidad están asociadas a mediciones imperfectas, lo que puede interpretarse como un conocimiento vago, impreciso, incierto, ambiguo, inexacto, o probabilístico, por naturaleza. La lógica difusa, como su propio nombre indica, es una lógica alternativa a la lógica clásica

que introduce un grado de incertidumbre en los objetos que evalúa a través de funciones evaluadas en el intervalo $[0,1]$, que representan el grado de certeza de que se dispone sobre dicho objeto. La teoría de conjuntos difusos (introducida en 1965 por Zadeh [65]) es una herramienta que resulta de gran utilidad para el modelado de estas situaciones. Dar un enfoque difuso a un problema planteado permite incorporar información imprecisa o incompleta, y permite entender la realidad incorporando informaciones de gran valor tales como juicios razonados (y/o subjetivos) acerca de lo que va a suceder. Se trata de un campo de investigación muy activo en la actualidad que cuenta con numerosas publicaciones científicas asociadas.

El término *difuso* surge de la traducción al castellano de la expresión inglesa *fuzzy*, que también puede ser interpretada como *borroso*. A lo largo de la presente Memoria utilizaremos el adjetivo *difuso* por considerarlo más frecuente en el campo científico. Una de las nociones más importantes que se han adaptado al ambiente difuso es la generalización del concepto de número real, dando lugar a los *números difusos* (véase [19–21]). Éstos pueden interpretarse como entidades probabilísticas que determinan el grado de certidumbre que se puede tener de que cierto valor de una variable esté comprendido dentro de un intervalo real concreto. Mizumoto y Tanaka [37] introdujeron una aritmética para operar con números difusos que extiende a las operaciones habituales con números reales. En la práctica, la aritmética de intervalos también ha demostrado ser un método eficaz cuando se opera con números difusos.

Debido a su posible utilidad en algunas aplicaciones, en 2001, Voxman [60] resaltó una clase de números difusos que vienen caracterizados por poseer un soporte finito, y los denominó números difusos *discretos*. Desde entonces, numerosos/as investigadores/as han han utilizado esta clase de números difusos en el desarrollo de sus estudios de investigación (véase, por ejemplo, [8, 56–59]).

En 2014, Roldán *et al.* [47] introdujeron una noción diferente del concepto de número difuso *discreto* suponiendo que su imagen (como aplicación evaluada en $[0,1]$) es un conjunto finito o numerable. El principal interés de esta clase de números difusos radica en el hecho de que, en muchos casos, el cálculo habitual que se realiza con números difusos sólo hace referencia a ciertos niveles de confianza (es decir, a una aproximación finita de un número difuso arbitrario) y, en la práctica, la mayoría de ejemplos de estructuras difusas (espacios métricos probabilísticos, espacios métricos difusos en varios sentidos y espacios métricos difusos intuicionistas) se construyen utilizando este tipo de números. Una de las propiedades más importantes de la familia los números difusos cuya imagen es un conjunto numerable (o finito) es que ésta es cerrada bajo las operaciones habituales, lo que posibilita la seguridad de que, operando con esta clase de números difusos, el resultado siempre será un número difuso del mismo tipo de los que se están considerando. Esta propiedad no se cumple, ni siquiera, con números difusos tan elementales como los triangulares, pues su multiplicación lleva a números difusos de tipo LR.

La presencia de incertidumbre en los procesos naturales o en su medición nos hace plantearnos si la mejor forma de medir distancias entre números difusos puede realizarse a través de números reales. A lo largo de los años, se han propuesto distintas nociones acerca del concepto de *espacio métrico difuso* (véase Roldán *et al.* [44, 48], y las referencias incluidas en estos trabajos) que, en gran medida, están basadas en asociar números reales a parejas de números difusos. Sin embargo, ninguno de estos conceptos trata, de manera particular, el problema de cómo definir una noción de distancia (o, más bien, de medida de similitud) entre dos números difusos arbitrarios (véase [1, 2, 4, 5, 13, 27, 55]). De manera natural, la distancia entre números difusos debe ser medida, de nuevo, mediante otro número difuso. La diferencia entre números difusos (utilizando el *principio de extensión de Zadh*) no es un método eficaz a la hora de determinar si dos números difusos se parecen o no, ya que su soporte puede ser muy amplio aún cuando los números difusos sean muy parecidos. De esta forma, no existe un procedimiento canónico universalmente aceptado que asocie, a cada par de números difusos, otro número difuso que mida, en cierta forma, el grado de cercanía o lejanía entre ellos (es por ello que, en muchas ocasiones, esta similitud entre números difusos se mide utilizando números reales). La dificultad asociada a este problema es doble: por un lado, siguiendo la noción clásica, una métrica en un conjunto debe tomar valores en el intervalo de los números reales no negativos, lo cual no tiene, en principio, un análogo difuso; por otro lado, y relacionado directamente con lo anterior, dado que el conjunto de los números difusos carece de un orden parcial universalmente aceptado que extienda al orden entre números reales, la desigualdad triangular se transforma en un axioma difícilmente generalizable. La posibilidad de ordenar números difusos arbitrarios que representan cantidades inciertas (e incluso opiniones subjetivas) es un tema de gran interés en el campo científico (véase [3, 39]) pues permitiría establecer ranking entre objetos que, de forma natural, no guardan una relación ni siquiera de orden parcial entre ellos.

Como aproximación a este problema, Aguilar *et al.* [2] introdujeron una familia de aplicaciones que, de muchas formas diferentes, permiten determinar un número difuso que puede interpretarse como la distancia (en el sentido de cercanía o lejanía) entre dos números difusos. En lugar de *métricas*, los autores prefirieron utilizan el término *medidas de similitud* ya que, aunque pueden cumplir varios de los axiomas que definen el concepto de métrica sobre un conjunto arbitrario, su principal objetivo no es éste, sino el de expresar, en términos difusos, lo parecidos o lo diferentes que son dos números difusos arbitrarios. Las principales ventajas de esta familia son las siguientes: se definen en el conjunto formado por todos las números difusos (es decir, no sólo pueden ser consideradas entre números difusos de una misma clase, como son los triangulares o los trapezoidales); algunos de los subconjuntos más útiles del conjunto de todos los números difusos (incluyendo números difusos triangulares y trapezoidales) son cerrados bajo estas medidas de similitud; en algunos casos, como en el conjunto de los números difusos trapezoidales, son auténticas métricas difusas (pues verifican propiedades análogas a los axiomas métricos); en estos casos, estos subconjuntos están dotados de una topología Hausdorff que verifica el primer axioma de numerabilidad; por otra parte, una vez que hemos fijado los elementos geométricos que intervienen en la definición de medida de similaridad, el cálculo que conlleva la medida de la distancia es un proceso muy sencillo, intuitivo y con un coste computacional muy bajo; finalmente, esta familia nos permite resolver problemas que se plantean en un ambiente difuso aportando, en muchos casos, nuevas soluciones y, en otros, obteniendo resultados tan buenos o, al menos, similares, a los obtenidos por otras técnicas más complejas.

Otro de los temas que han suscitado gran interés en la comunidad científica es el estudio de las posibles conexiones entre las nociones difusas y probabilísticas. De esta convergencia ha surgido, por ejemplo, el concepto de *variable aleatoria difusa* (véase [41]) y las técnicas estadísticas que utilizan este tipo de información. Entre ellos, el problema de análisis de regresión en un entorno difuso ha sido discutido en la literatura desde diferentes puntos de vista y teniendo en cuenta una gran variedad de datos de entrada y de salida (véase [17, 23, 31, 52–54]). En Roldán *et al.* [45, 46] se presentan nuevas metodologías para resolver el problema de regresión difusa empleando una familia de semidistancias difusas entre números difusos y considerando el método de mínimos cuadrados.

El objetivo principal de esta Memoria es introducir una metodología de regresión difusa utilizando semidistancias difusas entre números difusos con imagen finita. Este enfoque puede ser útil cuando se conocen unos cuantos conjuntos de nivel de los números difusos que intervienen en la muestra (por ejemplo, cuando la recopilación de datos es costosa y/o requiere mucho tiempo), o cuando sólo queremos hacer una estimación de algunos conjuntos de nivel de la variable de respuesta difusa (véase [47]). Para ello, primero se introduce y se estudia la clase formada por los números difusos que vamos a manejar, justificando el interés que esta clase de números difusos podría suscitar en el futuro. En segundo lugar, también se desarrollan las herramientas que necesitaremos para el ajuste de un modelo de regresión difusa. A continuación se describe el algoritmo difuso propuesto con el fin de modelizar una variable difusa que toma valores en el conjunto de los números difusos finitos. Finalmente, mostraremos un ejemplo numérico basado en datos financieros de la economía china para ilustrar el uso de la metodología propuesta. De acuerdo con estos objetivos, la memoria se ha organizado en varios capítulos atendiendo a los siguientes objetivos.

En el Capítulo 2 está dedicado a la presentación de las nociones básicas tanto de la teoría de conjuntos difusos como de la teoría de regresión.

En el Capítulo 3 se estudia la familia formada por todos los números difusos finitos, es decir, los números difusos con imagen finita como funciones valuadas en el intervalo $[0,1]$ de los números reales. Aunque a primera vista se puede considerar que esta clase es muy restrictiva, esta idea es falsa. Por ejemplo, dado un número real $\alpha_0 \in (0,1)$, la familia formada por todos los números difusos cuya imagen es el conjunto $\{0,\alpha_0,1\}$ es biyectiva a la clase formada por los números difusos trapezoidales (que es ampliamente utilizada en la mayoría de los trabajos científicos que tratan temas difusos), y si suponemos que la imagen del número difuso puede contener cuatro puntos distintos, entonces obtenemos una familia aún más amplia, caracterizada por seis números reales ordenados de menor a mayor, lo que sobrepasa al conjunto formado por los números difusos trapezoidales. Pero ésta no es la única ventaja de este tipo de números difusos. Cuando utilizamos números difusos triangulares para modelizar la imprecisión inherente a un experimento aleatorio, asumimos que todos los conjuntos de nivel de los resultados son conocidos y que éstos varían de forma lineal. Sin embargo, los números difusos finitos pueden ser útiles para modelizar experimentos en los que sólo conocemos algunos conjuntos de nivel (tal vez porque son difíciles de obtener o porque es costoso acceder a ellos), o incluso situaciones en las que nuestro conocimiento de dos o tres conjuntos de nivel es suficiente para determinar que los números difusos involucrados no son triangulares o trapezoidales. Otra ventaja de los números difusos finitos es el hecho de que pueden ser fácilmente caracterizados por un número real (que sirve como valor central y más probable) y una lista finita de números reales no negativos (llamados *amplitudes a la izquierda* y *a la derecha*). Esta caracterización hace que la aritmética con este tipo de números difusos sea fácil e intuitiva. Además, a diferencia de los números difusos triangulares o trapezoidales (cuyos productos bajo la aritmética usual no producen un número difuso del mismo tipo), siguiendo las técnicas introducidas en [47], observamos que la familia de números difusos finitos es cerrada bajo las operaciones habituales introducidas por Mizumoto y Tanaka [37].

En el Capítulo 4 pueden diferenciarse claramente dos partes. La primera parte del Capítulo 4 se centra en la introducción de las herramientas que están involucradas en la mayoría de procedimientos de regresión. En un primer paso, partiendo de un modelo teórico que depende de ciertos parámetros, necesitamos medir la similitud entre los valores observados y estimados (los *residuos*). Los desarrollos clásicos emplean el cuadrado de la diferencia entre ambos números para ajustar un modelo (es el conocido *método de mínimos cuadrados*). Esta función de pérdida no es útil en un ajuste difuso porque la diferencia entre dos números difusos no es una buena medida de la similitud entre ellos. De hecho, hay varias maneras de extender la diferencia entre números reales a los números difusos, pero ninguna de ellas es compatible con la suma (en el sentido de que, en general, el número difuso $(\mathcal{A}-\mathcal{B})+\mathcal{B}$ es distinto de $\mathcal{A}$ para cualesquiera números difusos $\mathcal{A}$ y $\mathcal{B}$). Ante este problema, la solución propuesta en algunos enfoques consiste en utilizar números reales para medir la distancia entre dos números difusos. Sin embargo, desde nuestro punto de vista, esta interpretación no es coherente con la interpretación de los números difusos como entidades inciertas. En un entorno difuso, las medidas de similitud (es decir, procedimientos que asocian, a cada par de números difusos, otro número difuso que se interpreta como la similitud o la diferencia entre ellos) que se utilizan suelen producir mejores resultados que las métricas (en el sentido clásico). En un segundo paso, después de determinar el proceso que se va a emplear para estimar los errores, observamos que cada modelo ajustado conlleva una suma total difusa de los cuadrados de los residuos y es necesario decidir el modelo óptimo, es decir, el modelo que minimiza estos números difusos. Sin embargo, en un entorno difuso, no hay un método universalmente aceptado para ordenar números difusos (por el contrario, hay muchas maneras de clasificar los números difusos). En [45], los autores introdujeron, al mismo tiempo, una familia de relaciones binarias y medidas de distancia, que dependían de varios parámetros (con una interpretación geométrica) y que permitían a los/as investigadores/as desarrollar un proceso de regresión personalizado que puede variar según sus intereses, con aplicaciones en el ámbito económico y en otros campos de la ciencia en general. Tales relaciones binarias y medidas de distancia se convierten en órdenes parciales y auténticas métricas difusas en muchos casos, lo que da una coherencia intuitiva al procedimiento de regresión difusa. Desafortunadamente, aunque las familias introducidas son apropiadas para manejar una gran clase de números difusos (triangulares, trapezoidales, etc.), no producen resultados óptimos cuando consideramos números difusos finitos. En el Capítulo 4 mostramos cómo extender las familias introducidas en [45] de tal manera que, preservando sus buenas propiedades, la nueva familia también es capaz de manejar números difusos finitos en el sentido de que los resultados obtenidos son coherentes con nuestra intuición. Estas familias han sido especialmente diseñadas para trabajar de forma sencilla con números difusos finitos cuando se expresan a través de la caracterización dada en el capítulo anterior. Además, la nueva familia de medidas de similitud es más amplia que la anterior y proporciona a los/as investigadores/as interesados/as en este área nuevas y diferentes opciones para medir similitudes entre números difusos.

En la segunda parte del Capítulo 4 se hace uso de las herramientas previamente desarrolladas y se plantea una nueva aproximación al problema de la regresión difusa que puede ser muy útil para muchos/as investigadores/as de diferentes campos: una metodología de regresión que es capaz de manejar con soltura una amplia variedad de números difusos, incluidos los números difusos finitos. De la misma manera que el método clásico, el enfoque presentado considera un entorno en el que una variable aleatoria difusa finita está explicada a partir de un vector aleatorio de variables. Sin embargo, la solución propuesta a lo largo de la presente Memoria es novedosa por varias razones. En primer lugar, los estudios previos que se han realizado hasta el momento nunca han considerado números difusos finitos y, por otro lado, las herramientas desarrolladas en trabajos anteriores no pueden aplicarse directamente, de una forma sencialla y eficaz, a este tipo de números (o no funcionan correctamente en este contexto). Ésta es la razón principal para considerar una versión extendida de las nociones anteriores. En segundo lugar, esta Memoria también da una nueva solución al problema del ajuste de modelos de amplitudes no negativas para las amplitudes con objeto de garantizar que la respuesta predicha sea un número difuso.

Finalmente, el Capítulo 5 ilustra la metodología propuesta utilizando datos reales. El objetivo de este capítulo es dar a los/as inversores/as una herramienta cuantitativa para ayudarles a tomar decisiones de inversión y obtener un modelo de regresión difusa que dé valor al proceso de inversión. Este ejemplo se utiliza para ilustrar la metodología propuesta y demostrar la eficiencia de la misma. Cabe señalar que el objetivo ha sido en ningún momento crear un algoritmo para indicarle a un/a inversor/a cuándo debe comprar o vender el bono convertible, sino una herramienta cuantitativa que le ayude a hacerlo de acuerdo con sus propias expectativas sobre las variables independientes que se consideraron, el precio de las acciones en Hong Kong y Shangai y el valor SHIBOR. Cabe destacar en relación al ejemplo numérico que no podemos utilizar un análisis clásico porque los datos no se publican simultáneamente en el mismo día (no es posible obtener información diaria el mismo día para todas las variables financieras consideradas). Si resumiésemos los datos del mes utilizando una medida de tendencia central, estaríamos perdiendo mucha información de gran interés sobre la evolución del valor a lo largo del mes. De esta manera, el análisis clásico no nos proporciona una respuesta razonable a este problema. Para evitar la pérdida de información, trabajadores/as de la empresa Shenwan Hongyuan Securities (HK) Limited, como expertos/as en análisis financiero, describieron las variables en términos de números difusos finitos. A partir de los datos de cada mes, obtuvieron medidas diferentes (tales como la media, la desviación estándar, la mediana, los percentiles, etc.), y sobre en base a dicha información, propusieron un valor central representativo, dos extremos (sin incluir valores atípicos) para definir el soporte del número difuso finito y dos valores intermedios, de modo que aproximadamente el 50% de los datos caen dentro del rango de puntuaciones definidas por ellos/as. Si el/la investigador/a no fuese un/a experto/a en el campo, podría considerar un gráfico de caja y bigotes (por ejemplo, el primer y el tercer cuartil pueden ser utilizados como valores intermedios y percentiles apropiados, tales como los correspondientes al 2%, 9%, 91% y 98%, para los extremos del gráfico, que determinarían los valores correspondientes al soporte del número difuso).

De manera resumida, las principales ventajas de la metodología propuesta en esta Memoria son las siguientes:

- Proporciona a los/as investigadores/as un enfoque fácil para el problema de analizar relaciones de regresión cuando los datos observados pueden verse afectados por diferentes fuentes de incertidumbre.
- No se consideran restricciones de no negatividad desde el principio.
- Los/as investigadores/as pueden modelizar relaciones estadísticas entre variables difusas con un método que se puede aplicar en el ámbito económico y financiero y en otros contextos diferentes.
- La técnica no se limita a considerar modelos lineales.

- La metodología propuesta no se limita a intervalos numéricos o variables difusas triangulares (como en otros trabajos).
- El método descrito puede considerarse como genuinamente difuso ya que no se limita a variables explicativas de tipo real ni a funciones de pérdida de tipo real.

CHAPTER 1

Introduction

1.1 Presentation

In many real-life situations the information associated with some random experiments is perfectly measurable but, in many cases, is only valuable. Recent scientific research shows the increasing instability of economic and business systems and the need to create new instruments to represent and manage this instability. The development of current economic models demand the ability to manage different sources of imprecision that arise naturally due to the lack of precise knowledge. Nowadays the management of uncertainty (also called imprecision, vagueness or ambiguity) considers different approaches that are not competitive but complementary: the probabilistic approach and fuzzy approach.

The theory of fuzzy sets (introduced in 1965 by Zadeh [65]) is a very useful tool for modeling these situations. To give a fuzzy approach to a classical problem allows the incorporation of inaccurate or incomplete information and uncertainty on parameters, properties, geometry, initial conditions, etc. We can also understand real-world problems incorporating valuable information, such as reasoned judgments about what will happen and about what has happened.

The term fuzzy is used as a designation for a particular class of sets. However we observe that this term caused different reactions, because it sound like something "unclear". However, the apparent internal contradiction of its own name has not been an inconvenience to a increase of theoretical developments and applications of fuzzy logic to the field of science, technology and, even, empirical analysis in the social sciences. Moreover, many authors agree that "there is nothing fuzzy about fuzzy logic". It is the theory of fuzziness, not being fuzzy itself. On the contrary, there is a powerful mathematical support behind it.

One of the most important notions that have been adapted to a fuzzy context is the generalization of the concept of real number, giving rise to the term fuzzy number (see [19–21]). Fuzzy numbers do not refer to one single value but rather to a set of possible values, where each possible value has its own weight, in general, between 0 and 1. That is, using or proposing a membership function which associates these weights appropriately. Mizumoto and Tanaka [37] introduced a way to operate with FNs that also extends the usual operations with real numbers. However, some difficulties are found when we try to use this definition. In practice, the interval arithmetic has also proven to be a powerful method when operating with FNs. In 2001, Voxman [60] introduced the conception of *discrete FN* (which is useful in some applications) assuming that its support is finite and gave out a canonical representation. Note that his notion of discrete FN is not a FN in the sense that we will use. Since then, a lot of work in this direction have been done and many results have been obtained (for instance, see [8, 59]).

1.2 Motivation and previous works

In 2014 Roldán et al. [47] introduce a slightly notion of *discrete FN* assuming that its image is a finite or countable set. This kind of FNs is interesting since, in many cases, the usual computation of FNs is only referred to certain data (a finite approximation of a FN) and, in practice, most of examples of fuzzy structures (probabilistic metric spaces, fuzzy metric spaces in several senses and intuitionistic fuzzy metric spaces) are constructed using these classes of FNs. The family of all FNs whose image is a countable (or finite) set is closed under the usual operations (notice that, in general, arithmetic operations between discrete FNs in the sense of Voxman [60] do not preserve the closeness of the operations when we use the Zadeh's Extension Principle).

The presence of uncertainty also leads us to consider whether a real number is the best way to measure distances between fuzzy numbers. Over the years, different notions have been proposed about the concept of fuzzy metric space (see Roldan *et al.* [44, 48], and the references therein). However, none of these concepts deals in particular with the problem of how to define a notion of distance or measure of similarity between two arbitrary fuzzy numbers [1,2,4,5,13,27,55]. In principle, it is not clear how to consider a canonical way to associate to two fuzzy numbers another fuzzy number that can be interpreted as the proximity or distance between them. Therefore, in many problems, this similarity between fuzzy numbers is measured using real numbers.

The problem is twofold: on the one hand, following the classical notion, a metric on a set must take values in the range of nonnegative real numbers, but the substraction between fuzzy numbers is not clear; on the other hand, and directly related to the above, the absence of a partial order leads the triangular inequality becomes an axiom hardly generalizable. In relation to the latter the introduction of new ranking concepts and the study of their properties is also a topic of interest [3, 39].

To solve this problem, Aguilar *et al.* [2] introduced a family of measures of similarity that, in many different ways, allow to obtain a fuzzy number that can be interpreted as the closeness between two fuzzy numbers . The main advantages of this family are the following: they are defined in the set of all the fuzzy numbers (not only among trapezoid diffuse numbers); some of the most useful subsets of the set of all fuzzy numbers (including triangular and trapezoidal diffuse numbers) are closed under these measures of similarity; in some cases, as in the set of the trapezoidal diffuse numbers, are authentic fuzzy metrics; in these cases, this set is endowed with a Hausdorff topology that verifies the first axiom of numerability; on the other hand, once we have fixed the geometric elements involved in our definition, the calculation of the distance measurement is a very simple, intuitive process and with a very low computational cost; finally, this family allows us to solve problems that arise in a fuzzy environment by providing, in some cases, new solutions and, in others, obtaining results as good or at least similar to those obtained by other more complex techniques.

On the other hand there is a great interest in studying the connections between fuzzy and probabilistic concepts (see [44]). From this convergence arise, for example, the concept of fuzzy random variable [41] and statistical techniques that use this kind of information. Among them, the problem of regression analysis in a fuzzy environment has been discussed in the literature from different points of view and taking into account a variety of input and output (see [36, 45, 46]). Roldán *et al.* [45, 46] introduced new methods to solve the problem of fuzzy regression using a family of fuzzy semi-distances between fuzzy numbers and considering the method of least squares.

1.3 Objetives

The main objetives of this report are:

- To introduce and study the class of fuzzy numbers which we are going to handle justifying the need of this class that could be very interesting for further studies.
- To develop the tools we need to fit a fuzzy regression model;
- To introduce and describe a fuzzy regression algorithm using fuzzy semidistances in order to model a fuzzy variable that takes finite fuzzy numbers as outputs. This approach can be useful when a few level sets of the fuzzy numbers are known (for instance, when collecting data is expensive and time consuming) or when we only want to do an estimation of some level sets of the fuzzy response variable.
- To illustrate the use of the proposed method considering a numerical example based on financial data.

1.4 Structure of the document

This manuscript is divided into different parts with distinct objectives.

Chapter 2 presents some basic notions of the theory of regression and the theory of fuzzy sets.

In Chapter 3 we study the family of all finite fuzzy numbers, that is, fuzzy numbers with finite image as real-valued functions. Although at a first sight one can consider that this class is very restrictive, this idea is false. For instance, given a real number $\alpha_0 \in (0, 1)$, the family of all fuzzy numbers whose image is the set $\{0, \alpha_0, 1\}$ is bijective to the class of triangular fuzzy numbers (which is widely used in most of scientific papers treating fuzzy topics), and if we suppose that the image of the fuzzy number can contain four distinct points, then we obtain a family which properly contains a subset bijective to the class of all trapezoidal fuzzy numbers. But this is not the unique advantage of this kind of fuzzy numbers. When we use triangular fuzzy numbers to modelize the imprecision inherent to a random experiment, we are assuming that all the level sets of the results are known and they vary in a linear way. However, finite fuzzy numbers can be useful to modelize experiments in which we only know some level sets (maybe because they are difficult to be obtained or because it is expensive to access them), or even situations in which our knowledge of two or three level sets is sufficient to determine that the involved fuzzy numbers are not triangular or trapezoidal. Another advantage of finite fuzzy numbers is the fact that they can be easily characterized by a real number (which serves as a central and most likely value) and a finite list of nonnegative real numbers (called *left* and *right spreads*). This characterization makes that the computation of this kind of fuzzy numbers is easy and intuitive. Furthermore, unlike triangular or trapezoidal fuzzy numbers (whose products are not a fuzzy number of the same type), following the techniques introduced in [47], we observe that the family of finite fuzzy numbers is closed under the usual operations introduced by Mizumoto and Tanaka [37].

The first part of Chapter 4 is dedicated to introduce the tools that are involved in many regression procedures. In a first step, starting from a theoretical model depending on certain parameters, we need to measure the distance between observed and estimated values (the *residuals*). Classical developments can employ the square of the difference between both numbers to fit a model (the least squares method). This loss function is not useful in a fuzzy setting because the difference between two fuzzy numbers is not a good measure of the distance between them. In fact, there are several ways to extend real difference to fuzzy numbers, but none of them is compatible with the sum (in the sense that, in general, $(\mathcal{A} - \mathcal{B}) + \mathcal{B}$ is distinct to $\mathcal{A}$ for fuzzy numbers $\mathcal{A}$ and $\mathcal{B}$). Furthermore, some approaches uses real numbers to measure the distance between two fuzzy numbers but, from our point of view, this interpretation is not coherent with the meaning of fuzzy numbers as uncertain entities. In the fuzzy setting similarity measures (that is, procedures that associate, to each pair of fuzzy numbers, another fuzzy number that is interpreted as the similarity of difference between them) are used to produce better results than metrics (in the classical sense). In a second step, after determining the process that is going to be employed to estimate the errors, we have that every fitted model leads to a fuzzy total sum of squares of residuals and we need to decide the optimal model, that is, the model which minimize these fuzzy numbers. However in a fuzzy setting, there is not an universally accepted method for ordering fuzzy numbers (on the contrary, there are a lot of ways for ranking fuzzy numbers). In [45], the authors introduced, at the same time, a family of binary relations and distance measures, involving many variables that permit the researchers to develop a customized regression process depending on their interests, with many applications to economics and other fields in general. Such binary relations and distance measures became partial orders and genuine fuzzy metrics in many cases, which gives an intuitive coherence to the fuzzy regression procedure. Unfortunately, although the introduced families are appropriate to handle a large class of fuzzy numbers (triangular, trapezoidal, etc.), they do not produce optimal results when we consider finite fuzzy numbers. In Chapter 4 we show how to extend the families given in [45] in such a way that, preserving their good properties, the novel family is also able to handle finite fuzzy numbers in the sense that obtained results are coherent with human intuition. Such families have been especially designed to work in a simple way with finite fuzzy numbers when they are expressed through the characterization given in Section 3. Furthermore, the new family of similarity measures is wider than the previous one and it gives to researchers interested in this topic new different possibilities for measuring similarities between fuzzy numbers.

In the second part of Chapter 4 the tools previously developed are considered to introduce an approach to the problem of fuzzy regression which can be very useful for many researchers of different fields: a regression methodology involving finite fuzzy numbers. In the same way that the classical method, the presented approach considers a setting in which a finite fuzzy random variable is intended to be explained from a random vector of variables. However the solution proposed in this report is novel for several reasons. Firstly previous studies have never considered finite fuzzy numbers and the tools developed in previous papers cannot be directly applied to this kind of numbers (or they do not work properly in this setting). This is the main reason to consider an extended version of previous notions. Secondly, this dissertation also gives a new solution to the problem of fitting nonnegative models for the spreads to guarantee that the predicted response is a fuzzy variable.

Finally, Chapter 5 illustrates the proposed methodology using real data. We emphasize that we cannot use a classical analysis because the data are not published simultaneously on the same day (it not possible to obtain daily information or on the same day for all the considered financial variables). However if we decide to summarize the data of the month using a central tendency measure, we are losing lots of information. In this way, classical analysis does not provide us a reasonable response to this problem. To avoid the loss of information, experts in financial analysis of Shenwan Hongyuan Securities (HK) Limited (a Hong Kong-based investment holding company principally engaged in financial businesses) described the variables in terms of finite fuzzy numbers. They obtained from each month different measures (such as mean, standard deviation, median, percentiles, etc.), and based on them, they proposed a representative central value, two extreme values (not including outliers) to define the support of the finite fuzzy number and two intermediate values, such that approximately 50% of the data fall within the range of scores defined by them. If the researcher is not an expert in the field, he/she could consider a box and whisker plot (the first and third quartiles can be used as intermediate values and appropriate percentiles like 2%, 9%, 91%, 98% can be also used for whisker ends, that is, those values corresponding to the support of the fuzzy number).

CHAPTER 2

Preliminaries

In this chapter we give some established concepts and results which are needed in the development of this report. Firstly we provide some background on binary relations, on the classical regression analysis and on fuzzy set theory. A section of this chapter will be devoted to each of them.

2.1 Sets and binary relations

Throughout this memory, we denote by $\mathbb{R}$, $\mathbb{R}^+ = (0, \infty)$, $\mathbb{R}_0^+ = [0, \infty)$ and $\mathbb{R}_0^- = (-\infty, 0]$ the set of all real numbers, the set of all positive real numbers, the set of all nonnegative real numbers and the set of all non positive real numbers, respectively. From now on, let X and Y be two nonempty sets and let $f : X \to Y$ be a mapping. The *domain* of f is denoted by $\mathrm{dom}(f)$, and its *image* (or *range*) is $f(Y) = \{ f(x) : x \in X \}$. If A is a *subset* of X we will write $A \subseteq X$, and if A and X are distinct, then $A \subset X$.

A *binary relation on* X is a nonempty subset $\mathcal{R}$ of the product space $X \times X$. For simplicity, we denote $x \preccurlyeq y$ if $(x, y) \in \mathcal{R}$, and we will say that $\preccurlyeq$ is the binary relation on X. We shall use $\preccurlyeq$ and $\sqsubseteq$ to denote binary relations on X.

A binary relation $\preccurlyeq$ on X is:

- *reflexive* if $x \preccurlyeq x$ for all $x \in X$;
- *transitive* if $x \preccurlyeq z$ for all $x, y, z \in X$ such that $x \preccurlyeq y$ and $y \preccurlyeq z$;
- *antisymmetric* if $x \preccurlyeq y$ and $y \preccurlyeq x$ imply $x = y$.

A reflexive and transitive relation on X is a *preorder* (or a *quasiorder*) on X. In such a case, $(X, \preccurlyeq)$ is a *preordered space*. If a preorder $\preccurlyeq$ is also antisymmetric, then $\preccurlyeq$ is called a *partial order*, and $(X, \preccurlyeq)$ is a *partially ordered space* (or a *partially ordered set*). For instance, the binary relation $\leq$ is a partial order on any nonempty subset of real numbers.

A binary relation $\preccurlyeq$ is a *total order* if it is transitive, antisymmetric, and it satisfies:

- *totality*: $x \preccurlyeq y$ or $y \preccurlyeq x$ for all $x, y \in X$.

Throughout this memory, we will use some abbreviations: we use "*FN*" rather than "*fuzzy number*" and we abbreviate "trapezoidal fuzzy number" by "*TFN*".

2.2 Background on fuzzy numbers

The uncertainty naturally arises in many real experiments different by the lack of precise knowledge. The types and sources of uncertainty are varied and cannot be classified into a single category. For example, the randomness or probabilistic uncertainty in a experimental study, that appears selecting a sample from a given population, is formalized through the Calculus of Probabilities and studied using the classical concept of random variable. Another type is the non-stochastic inaccuracy or uncertainty which arises due to the lack of information or associated with a process of aggregation, the imprecision in the transmission of data or as a consequence of the experimental errors, etc. In this case, new concepts (such as *fuzzy random variable*) are needed in order to manage the uncertainty.

Consequently, in the real world there is a non-perfect knowledge, i.e., vague, uncertain, ambiguous, inaccurate, or probabilistic knowledge in nature. Human thought and reasoning often leads such information, probably originated from the inherent uncertainty of human concepts and from reasoning based on similar but not identical experiences. The theory of fuzzy sets and fuzzy logic, as its name suggests, is a logical alternative to classical logic which aims to introduce a degree of vagueness in the things that qualify.

The concept of *fuzzy set* and operations between fuzzy sets were introduced by Lotfi Zadeh in 1965. Zadeh was born in Baku in 1921, Azerbaijan, as Lotfi Aliaskerzadeh. In 1943 he decided to emigrate to the United States where he changed his name to Lotfi Asker Zadeh. This mathematician, electrical engineer, computer scientist, artificial intelligence researcher and professor emeritus at the University of Berkeley, is famous for introducing the theory of fuzzy sets or fuzzy logic and he is also considered the father of the theory of possibility.

This theory gives the mathematical framework to treat uncertainty and imprecision. Since its introduction, much time and effort is being devoted to research in this area. Nowadays, this field of research is very important, both for mathematical or theoretical implications as for its practical applications and proof of this importance is the large number of international journals (Fuzzy Sets and Systems, IEEE Transactions on Fuzzy Systems, etc.), conferences (FUZZ-IEEE, IPMU, EUSFLAT, ESTYLF, etc.), and many books dedicated to this topic.

When handling vagueness, there are two points of view. In a first approach the focus is on a real random variable that cannot be observed accurately. For instance, if we want to measure something, a real accurate measure underly which quantification will be performed with a measuring instrument that can have a certain degree of inaccuracy. Another studies in which there is no underlying real variable, the focus is to obtain conclusions based on a proper imprecise variable.

To handle the uncertainty, sometimes it is synthesized into a real value and classical statistical techniques are applied, with the drawback that this process entails that a lot of information is lost, due to experimental inaccuracy. To avoid the loss of information we use the concept of fuzzy set. Formally, any $[0,1]$-valued function determines a fuzzy set. In real problems, the class of fuzzy numbers consisting of upper semi-continuous $[0,1]$-valued functions with compact support is rich enough to cover most of the applications (see [18]). However, this class is still very general, and most researchers use simple shapes, such as triangular fuzzy numbers, trapezoidal and LR-fuzzy numbers which satisfy the need of modeling fuzzy problems and are easier to fix and handle. A function that assigns a fuzzy subset to each possible output of a random experiment is a fuzzy random variable. This concept was introduced by Puri and Ralescu [41] in 1986. In recent decades various studies have been developed on random sets and fuzzy random variables among which the formalization of concepts, strong laws of large numbers, definition of various measures of central tendency and dispersion, hypothesis testing, etc.

Next we will give some fundamental notions of this theory.

2.2.1 Fuzzy sets and fuzzy numbers

One of the key tools of classical sets theory is the concept of *subset*. Given a nonempty set X (interpreted as the *universe*), we will say that A is a *subset of* X, and we will denote it by $A \subseteq X$, if each element of A is also an element of X. Subsets of X are completely characterized by their corresponding *characteristic* *functions*, which are functions whose domains are X and whose ranges are the discrete set $\{0,1\}$. In this way, if A is a subset of X, then its *characteristic function* is $\chi_A : X \to \{0,1\}$, given by

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A, \\ 0, & \text{if } x \in X \diagdown A. \end{cases}$$

Conversely, any map $\chi : X \to \{0,1\}$ is the characteristic function of the subset

$$A = \{\, x \in X : \chi(x) = 1 \,\}.$$

The value of $\chi(x)$ is, from this point of view, interpreted as the membership function that decides if the point x belongs to A: if $\chi(x) = 1$ then x belongs to the subset, and if $\chi(x) = 0$, then x does not belong to the subset. Replacing the finite set $\{0,1\}$ by the continuous interval $[0,1]$ we obtain the notion of *fuzzy set*. Due to its significance and importance throughout this memory, we will denote by $\mathbb{I}$ the compact (closed, bounded) subinterval of real numbers $[0,1]$.

Definition 2.2.1 (Fuzzy set) *A fuzzy set on X is a function $\mathcal{A} : X \to \mathbb{I}$.*

Each value $\mathcal{A}(x) \in \mathbb{I}$ is interpreted as the probability that x belongs to the fuzzy set. Following this conception, if $\mathcal{A}(x) = 1$, we are completely sure that x belongs to the fuzzy set; conversely, if $\mathcal{A}(x) = 0$, then x does not belong to the fuzzy set. For intermediate values $\mathcal{A}(x) \in (0,1)$, we have only a degree of uncertainty about the fact that x belongs (or not) to the fuzzy set.

Due to its great applications, literature on fuzzy sets is very extensive. We refer the reader to [19–21, 42].

Among the whole family of fuzzy sets, we are especially interested in the class of them that, in a natural sense, extend the notion of *real number*. Given a real number $r \in \mathbb{R}$, the characteristic function of the subset $\{r\}$ on $\mathbb{R}$ is

$$\overline{r} : \mathbb{R} \to \mathbb{I}, \qquad \overline{r}(x) = \begin{cases} 1, & \text{if } x = r, \\ 0, & \text{if } x \neq r. \end{cases} \tag{2.1}$$

The fuzzy set $\overline{r}$ on $\mathbb{R}$ uniquely determines the real number r, so any extension of real numbers to the fuzzy framework must necessarily identify any real number r with its corresponding fuzzy set $\overline{r}$. Anyway, there are several notions of *fuzzy number* (see, for instance, [42,43,59,60]) trying to carry out such generalization (for which their supports are compact or not, and imposing or nor the condition of normality). For our purposes, we will employ the following definition (which is widely assumed by most of researchers in this field).

Definition 2.2.2 (Fuzzy number) *A fuzzy number (for short FN) on* $\mathbb{R}$ *is a fuzzy set* $\mathcal{A}$ *on* $\mathbb{R}$ *such that:*

a) Normality condition*: There exits a real number* $x_0 \in \mathbb{R}$ *such that* $\mathcal{A}(x_0) = 1$.

b) *For all* $\alpha \in]0, 1]$, *the* α-level set *(or* α-cut*)*

$$\mathcal{A}_\alpha = \{ x \in \mathbb{R} : \mathcal{A}(x) \geq \alpha \}$$

is a non-empty, closed subinterval of $\mathbb{R}$.

For more details about FNs, see Dubois and Prade [19]- [21], Mizumoto and Tanaka [37] and Wu and Ma [61]. The previous concept of α-*level set* (or α-*cut*) plays a key role in fuzzy set theory.

Definition 2.2.3 (Kernel or *core*) *The kernel (or core) of a FN* $\mathcal{A}$ *is its* 1*-level set, that is,*

$$\ker \mathcal{A} = \mathcal{A}_1.$$

Definition 2.2.4 (Support) *The support of a FN* $\mathcal{A}$ *is the closure (in the Euclidean topology of* $\mathbb{R}$*) of the subset of points in which the FN takes strictly positive real values, that is,*

$$\operatorname{supp}(\mathcal{A}) = \overline{\{ x \in \mathbb{R} : \mathcal{A}(x) > 0 \}}.$$

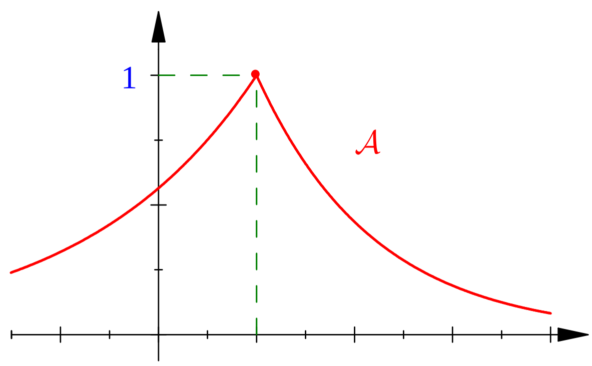

Figure 2.1: A fuzzy number with unbounded support.

Under the previous definitions, the support of a FN is not necessarily a bounded subset of $\mathbb{R}$ (see, for instance, Figure 2.1). However, for our purposes, throughout this memory, we will only consider FNs with bounded support. Therefore, it will be a compact subinterval of $\mathbb{R}$. Coherently, we will denote by $\mathcal{F}$ be the family of all FNs (with compact support). We advise the reader that the FNs we employ henceforth will have compact support.

Under these considerations, for each $\alpha \in \mathbb{I}$, the α-level set $\mathcal{A}_\alpha$ of $\mathcal{A}$ is a compact subinterval of $\mathbb{R}$ that can be expressed as

$$\mathcal{A}_\alpha = [\,\underline{a}_\alpha, \overline{a}_\alpha\,],$$

where $\underline{a}_\alpha$ is the inferior extreme and $\overline{a}_\alpha$ is the superior extreme[1] of the interval $\mathcal{A}_\alpha$. Following this notation, we will also denote the support of $\mathcal{A}$ by $\mathcal{A}_0 = [\,\underline{a}_0, \overline{a}_0\,]$, and $\mathcal{A}_1 = [\,\underline{a}_1, \overline{a}_1\,] = \ker \mathcal{A}$ will stand for the kernel of $\mathcal{A}$.

Given a FN $\mathcal{A}$ and varying α on $\mathbb{I}$, we can consider two functions

$$\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$$

[1]We point out that many researchers also use the notation $\mathcal{A}_\alpha = [L_{\mathcal{A}}(\alpha), R_{\mathcal{A}}(\alpha)]$ to explicitly declare the left and the right extremes of each level set.

where

$$\underline{a}\,(\alpha) = \underline{a}_{\,\alpha} = \min(\mathcal{A}_\alpha) = \min(\{\, x \in \mathbb{R} : \mathcal{A}\,(x) \geq \alpha \,\}) \qquad \text{and}$$
$$\overline{a}\,(\alpha) = \overline{a}_\alpha = \max(\mathcal{A}_\alpha) = \max(\{\, x \in \mathbb{R} : \mathcal{A}\,(x) \geq \alpha \,\})$$

for all $\alpha \in \mathbb{I}$. In particular, notice that

$$\mathcal{A}\,(\underline{a}_{\,\alpha}) \geq \alpha \quad \text{and} \quad \mathcal{A}\,(\overline{a}_\alpha) \geq \alpha \quad \text{for all } \alpha \in \mathbb{I}, \tag{2.2}$$

and

$$\underline{a}\,(\mathcal{A}\,(x)) \leq x \leq \overline{a}\,(\mathcal{A}\,(x)) \quad \text{for all } x \in \mathbb{R}.$$

These functions completely determine the FN $\mathcal{A}$ in the following sense.

Lemma 2.2.5 (Goetschel and Voxman [24]) A fuzzy set $\mathcal{A} : \mathbb{R} \to \mathbb{I}$ is a FN if, and only if, there exist two left continuous mappings $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ such that $\underline{a}$ is non-decreasing, $\overline{a}$ is non-increasing and $\mathcal{A}_\alpha = [\,\underline{a}_{\,\alpha}, \overline{a}_\alpha\,]$ for all $\alpha \in \mathbb{I}$.

Furthermore, if $x_0 \in \ker \mathcal{A}$, then $\underline{a}_{\,\alpha} \leq x_0 \leq \overline{a}_\alpha$ for all $\alpha \in \mathbb{I}$.

With respect to the shape of the plot of a FN, the following results shows some geometric properties that any FN must satisfy.

Lemma 2.2.6 ([24], [47]) If $\mathcal{A} \in \mathcal{F}$ and $x_0 \in \ker \mathcal{A}$, then $\mathcal{A}|_{[x_0,\infty[}$ is a non-increasing and left-continuous mapping and $\mathcal{A}|_{]-\infty,x_0]}$ is a non-decreasing and right-continuous mapping.

Given a FN $\mathcal{A}$, as there exists a point $x_0 \in \mathbb{R}$ such that $\mathcal{A}\,(x_0) = 1$ (as we have commented above, some authors refer this property as *normality*), the 1-level set $\mathcal{A}_1 = [\,\underline{a}_{\,1}, \overline{a}_1\,]$ of $\mathcal{A}$ is a non-empty bounded interval. The *center* and the *radious*[2] (or *central spread*) of the interval $\mathcal{A}_1$, that is, the numbers

$$\mathcal{D}_c\,\mathcal{A} = \frac{\underline{a}_{\,1} + \overline{a}_1}{2} \quad \text{and} \quad \operatorname{spr}\mathcal{A} = \frac{\overline{a}_1 - \underline{a}_{\,1}}{2},$$

[2] We use the term *radious* in the sense that if c is the center and r is the radious, then the interval is $[c-r, c+r]$.

are called, respectively, the *center* and the *central spread* of the FN $\mathcal{A}$ (see [2, 45]). This leads us to consider two functions

$$\mathcal{D}_c : \mathcal{F} \to \mathbb{R} \quad \text{and} \quad \text{spr} : \mathcal{F} \to \mathbb{R}_0^+ \tag{2.3}$$

that assign, to each FN $\mathcal{A}$, its center and its central spreads, respectively.

2.2.2 Some classes of fuzzy numbers

Next, we present some families of fuzzy numbers whose membership functions have very simple geometric shapes but which are, at the same time, widely used in current research mainly because they satisfy the modeling needs that arise in fuzzy problems and, in general, they are easier to set, manage and understand than other more general concepts (see [2]).

As we have just pointed out in (2.1), real numbers $\{ r : r \in \mathbb{R} \}$ can be seen as the FNs $\{ \overline{r} : r \in \mathbb{R} \} \subset \mathcal{F}$. These FNs are usually called *crisp FNs*.

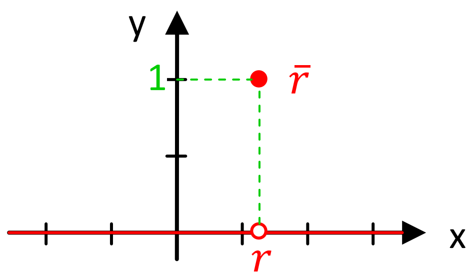

Figure 2.2: A real number r, defined as a fuzzy number.

In general, FNs whose image is the discrete set $\{0, 1\}$ are *rectangular FNs*, which can be expressed as:

$$\mathcal{A}(x) = \begin{cases} 1, & \text{if } a \le x \le b, \\ 0, & \text{in any other case} \end{cases} \tag{2.4}$$

(where $a, b \in \mathbb{R}$ are such that $a \le b$).

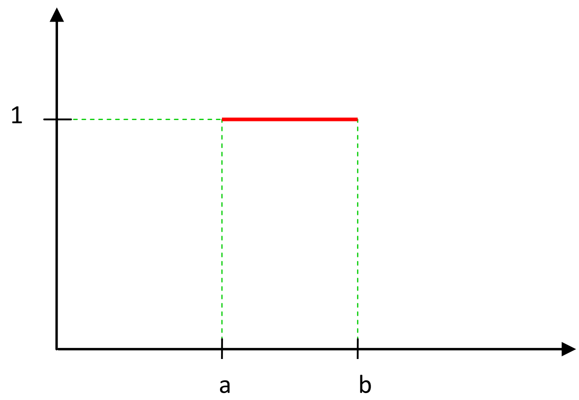

Figure 2.3: Plot of a rectangular fuzzy number.

Definition 2.2.7 *A* triangular fuzzy number *is a fuzzy number denoted by* $\mathcal{A} = (a/b/c)$*, where* $a, b, c \in \mathbb{R}$*,* $a \leq b \leq c$*, defined by:*

$$\mathcal{A}(x) = \begin{cases} \dfrac{x-a}{b-a}, & \text{if } a < x \leq b, \\ \dfrac{c-x}{c-b}, & \text{if } b < x < c, \\ 0, & \text{in any other case.} \end{cases}$$

If $a < b < c$, the plot of a triangular fuzzy number corresponds to a triangle which base is $[a, c]$ and the vertex is located at $x = b$. The kernel of the triangular fuzzy number $\mathcal{A} = (a/b/c)$ is $\{b\}$ while its support is the interval $[a, c]$.

Triangular fuzzy numbers belong to even more general fuzzy number classes that are very useful in practice for modeling uncertainty.

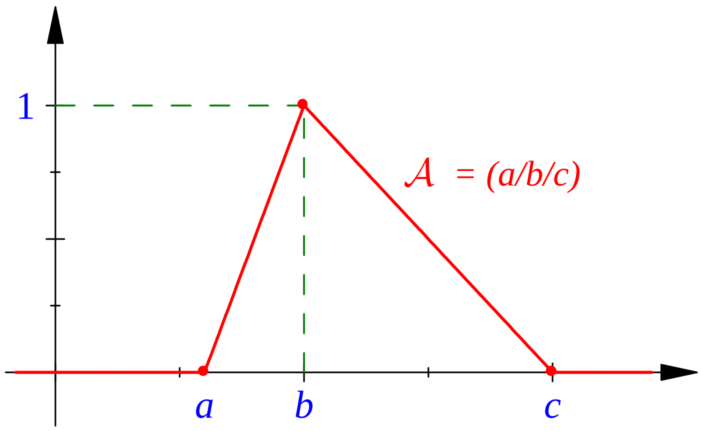

Figure 2.4: Plot of the triangular fuzzy number $(a/b/c)$.

Definition 2.2.8 *A* trapezoidal FN *(for short, a* TFN*) is a FN* $\mathcal{A} = (a/b/c/d)$ *given by*

$$\mathcal{A}(x) = \begin{cases} \dfrac{x-a}{b-a}, & \text{if } a < x < b, \\ 1, & \text{if } b \leq x \leq c, \\ \dfrac{d-x}{d-c}, & \text{if } c < x < d, \\ 0, & \text{in any other case,} \end{cases}$$

where $a, b, c, d \in \mathbb{R}$ *(known as the* corners of the FN*) are such that* $a \leq b \leq c \leq d$. *Let* $\mathcal{T}$ *the family of all trapezoidal FNs on* $\mathbb{R}$.

Clearly, the kernel of $\mathcal{A} = (a/b/c/d)$ is $[b, c]$ and its support is $[a, d]$. Although the classical notion of TFN occurs when $a < b < c < d$, the previous definition, including the possibility of equality of corners, is useful in practice because we capture FNs that are not necessarily continuous as mappings (see Figure 5.3).

TFNs can be uniquely determined by their kernels and their supports, i.e., if $\mathcal{A}$ and $\mathcal{B}$ are TFNs, then $\mathcal{A} = \mathcal{B}$ if, and only if, $\ker \mathcal{A} = \ker \mathcal{B}$ and $\operatorname{supp} \mathcal{A} = \operatorname{supp} \mathcal{B}$.

It is clear that a triangular fuzzy number can be considered as trapezoidal

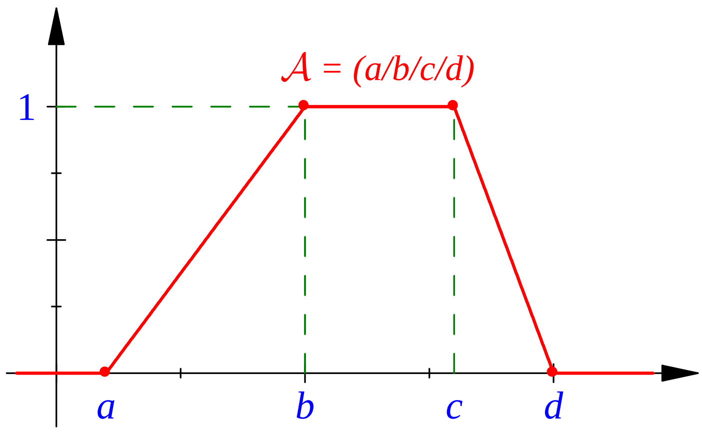

Figure 2.5: Plot of the trapezoidal fuzzy number $(a/b/c/d)$.

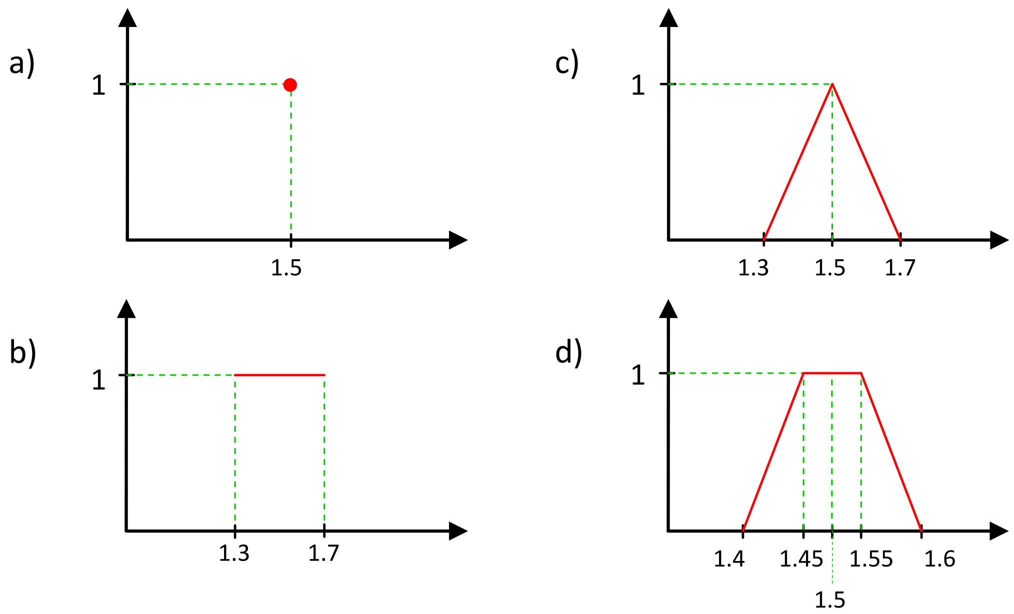

Figure 2.6: Numerical examples of trapezoidal fuzzy numbers.

when $a < b = c < d$. In addition, the real numbers r, seen as fuzzy numbers $\overline{r}$, can also be considered trapezoidal if we choose the case with $r = a = b = c = d$.

The center of the trapezoidal FN $\mathcal{A} = (a/b/c/d)$ is

$$A^c = \mathcal{D}_c\,\mathcal{A} = \frac{b+c}{2},$$

and the three *spreads* of $\mathcal{A}$ are the nonnegative real numbers

$$A^m = \operatorname{spr}\mathcal{A} = \frac{c-b}{2} \geq 0, \quad A^\ell = b - a \geq 0, \quad A^r = d - c \geq 0.$$

The centers and the spreads also completely determine the trapezoidal fuzzy number, since:

$$\begin{aligned} a &= A^c - A^m - A^\ell, \\ b &= A^c - A^m, \\ c &= A^c + A^m, \\ d &= A^c + A^m + A^r. \end{aligned}$$

Thus, a trapezoidal number, as a function of its center and its spreads, will be denoted by

$$\mathcal{A} = \operatorname{Tra}(A^c, A^m, A^\ell, A^r), \tag{2.5}$$

and if it is a triangular number we will use the notation

$$\mathcal{A} = \operatorname{Tri}(A^c, A^\ell, A^r).$$

Using this notation, $\mathbb{R}$ can be embedded in $\mathcal{T}$ in the simple way $r \mapsto \overline{r} = \operatorname{Tra}(r, 0, 0, 0) = \operatorname{Tri}(r, 0, 0) \in \mathcal{T}$.

Fuzzy numbers similar to trapezoidal fuzzy numbers but considering curves on both sides may also be defined as follows.

Definition 2.2.9 (Left right FN) *A left right FN (for short* LRFN*)* of Dubois and Prade *is a FN $\mathcal{A} = (a/b/c/d)_{LR}$, where $a, b, c, d \in \mathbb{R}$ (also known as the*

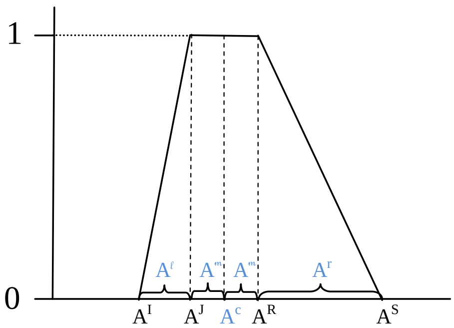

Figure 2.7: Center and spreas of a trapezoidal fuzzy number

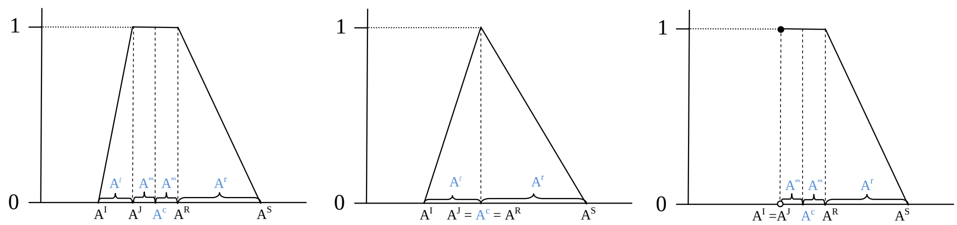

Figure 2.8: Examples of trapezoidal fuzzy numbers

corners of the FN*) are such that* $a \leq b \leq c \leq d$*, defined by:*

$$\mathcal{A}(x) = \begin{cases} L\left(\dfrac{x-a}{b-a}\right), & \textit{if } a < x < b, \\ 1, & \textit{if } b \leq x \leq c, \\ R\left(\dfrac{x-c}{d-c}\right), & \textit{if } c < x < d, \\ 0, & \textit{in any other case.} \end{cases}$$

where $L, R : \mathbb{I} \to \mathbb{I}$ *are strictly monotone, continuous mappings such that* $L(0) = R(1) = 0$ *and* $L(1) = R(0) = 1$*.*

In particular, L is strictly increasing and R is strictly decreasing. Clearly, the kernel of $\mathcal{A}$ is $[b, c]$ and its support is $[a, d]$. Note that a LRFN can be a discontinuous mapping at $x = b$ and $x = c$. Trapezoidal FNs are special cases of LRFNs with $L(x) = x = 1 - R(x)$ for all $x \in \mathbb{I}$.

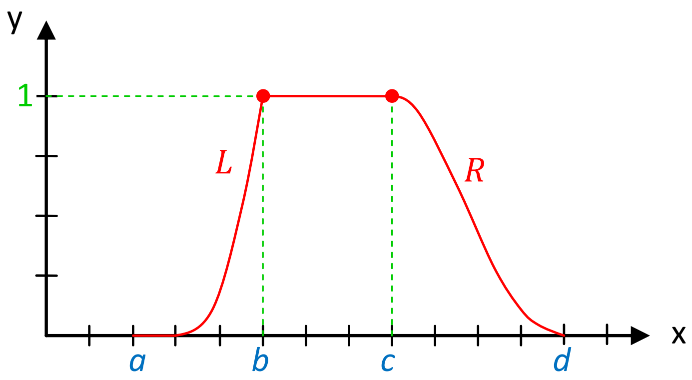

Figure 2.9: Plot of a LR fuzzy number.

2.2.3 Defuzzification of fuzzy numbers

FNs were born in order to generalize the notion of real number to an uncertainty framework. They are used in order to describe natural phenomena in which measure instruments are not precise enough or in which the stocastic nature of experiment produce random results. From this view-point, a FN intrinsically contains more information than a real number. However, in some cases, it is convenient to simplify the representation of a FN and to capture the information that it contains by using a unique real number. With this idea in mind, some parameters were introduced. For our purposes, a *defuzzification* (or a *valuation method*) will be a process in which fuzzy quantities are approximated by real numbers, that is, a mapping $\mathcal{D} : \mathcal{F} \to \mathbb{R}$. Let us introduce some of them (see [5, 60]).

Following [6], a function $S : \mathbb{I} \to \mathbb{I}$ is a *reducing function* if it is increasing and it satisfy the boundary conditions $S(0) = 0$ and $S(1) = 1$. Given a FN $\mathcal{A} \in \mathcal{F}$ and a reducing function S, we consider the following definitions.

- The *value of* $\mathcal{A}$ *(with respect to the reducing function* S*)*, denoted by

$\mathrm{Val}_S(\mathcal{A})$, is the real number

$$\mathrm{Val}_S(\mathcal{A}) = \int_0^1 S(\alpha)\left(\underline{a}_\alpha + \overline{a}_\alpha\right) d\alpha.$$

- The *ambiguity of* $\mathcal{A}$ *w.r.t.* S is the real number:

$$\mathrm{Amb}_S(\mathcal{A}) = \int_0^1 S(\alpha)\left(\overline{a}_\alpha - \underline{a}_\alpha\right) d\alpha.$$

- The *nonspecificity of* $\mathcal{A}$ is:

$$w(\mathcal{A}) = \int_0^1 \left(\overline{a}_\alpha - \underline{a}_\alpha\right) d\alpha.$$

- Given $q \in \mathbb{I}$, the *weighted expected value of* $\mathcal{A}$ is introduced by:

$$EV_q(\mathcal{A}) = (1-q)\int_0^1 \underline{a}_\alpha \, d\alpha + q\int_0^1 \overline{a}_\alpha \, d\alpha.$$

- The *expected value* of the FN $\mathcal{A}$ is $EV(\mathcal{A}) = EV_{1/2}(\mathcal{A})$.

Notice that the value of $\mathcal{A}$ depends on the middle points $\underline{a}_\alpha + \overline{a}_\alpha$ of its level sets $\mathcal{A}_\alpha$. However, other parameters are defined by considering the width $\overline{a}_\alpha - \underline{a}_\alpha$ of each interval $\mathcal{A}_\alpha$.

Finally, the *area under* $\mathcal{A}$, that is,

$$\int_{\mathbb{R}} \mathcal{A}(x)dx,$$

can also be seen as a valuation method of FNs (notice that it is finite because $\mathcal{A}$ has bounded support).

Notice that the previous defuzzifications try to summarize the information that each FN contains by using a unique real number. Rather the defuzzifications $\mathcal{D}_c : \mathcal{F} \to \mathbb{R}$ and $\mathrm{spr} : \mathcal{F} \to \mathbb{R}_0^+$, defined in (2.3), try to describe some geometric characteristic of a FN.

2.2.4 Operations between fuzzy numbers

The fuzzy arithmetic plays an important role when multiple fuzzy variables are involved.

The usual arithmetic defined by Minkowski for non-empty, compact and convex sets with the Euclidean norm $\|\cdot\|$ is based on the following operations:

- Minkowski sum: $A + B = \{\, a + b : a \in A, b \in B \,\}$;
- scalar product: $\lambda A = \{\, \lambda a : a \in A \,\}$

for all $A, B \subset \mathbb{R}^n$ non-empty, compact and convex sets and all $\lambda \in \mathbb{R}$.

This space has a *semi-linear structure* (not a vector space) because there does not exist a symmetrical element respect to the sum (only the neutral element for the sum). To consider a space similar to a linear space, if there exists, the *Hukuhara difference* is introduced.

This arithmetic can be extended to the set of FNs following the *Zadeh's extension principle* (1975). Thus the space $\mathcal{F}$ of all FNs can be provided with a semi-linear structure induced naturally by the following sum and product:

- sum: $(\mathcal{A} + \mathcal{B})_\alpha = \{\, a + b : a \in \mathcal{A}_\alpha, b \in \mathcal{B}_\alpha \,\}$;
- scalar product: $\lambda \mathcal{A}_\alpha = \{\, \lambda a : a \in \mathcal{A}_\alpha \,\}$

where $\mathcal{A}, \mathcal{B} \in \mathcal{F}$ and $\lambda \in \mathbb{R}$ whatever $\alpha \in (0, 1]$. Notice that the crisp FN

$$\overline{0} : \mathbb{R} \to \mathbb{I}, \qquad \overline{0}\,(x) = \begin{cases} 1, & \text{if } x = 0, \\ 0, & \text{if } x \neq 0, \end{cases}$$

is a neutral FN for the sum of the semi-linear structure.

If $\mathcal{A} = (a/b/c/d)$ is a trapezoidal FN, its α-cuts are

$$\mathcal{A}_\alpha = [\,(1-\alpha)a + \alpha\, b,\; \alpha\, c + (1-\alpha)d\,]$$

for all $\alpha \in \mathbb{I}$. As a consequence, it is easy to check that:

- $$\mathrm{Tra}(A^c, A^m, A^\ell, A^r) + \mathrm{Tra}(B^c, B^m, B^\ell, B^r) = \mathrm{Tra}(A^c + B^c, A^m + B^m, A^\ell + B^\ell, A^r + B^r);$$
- $$\lambda\,\mathrm{Tra}(A^c, A^m, A^\ell, A^r) = \mathrm{Tra}(\lambda A^c, \lambda A^m, \lambda A^\ell, \lambda A^r)$$

for all $\mathcal{A}, \mathcal{B} \in \mathcal{T}$ and all $\lambda \geq 0$.

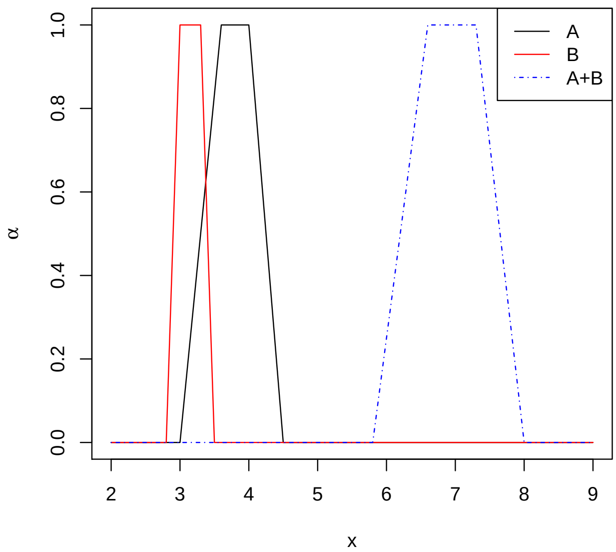

Figure 2.10: Sum of two trapezoidal fuzzy numbers.

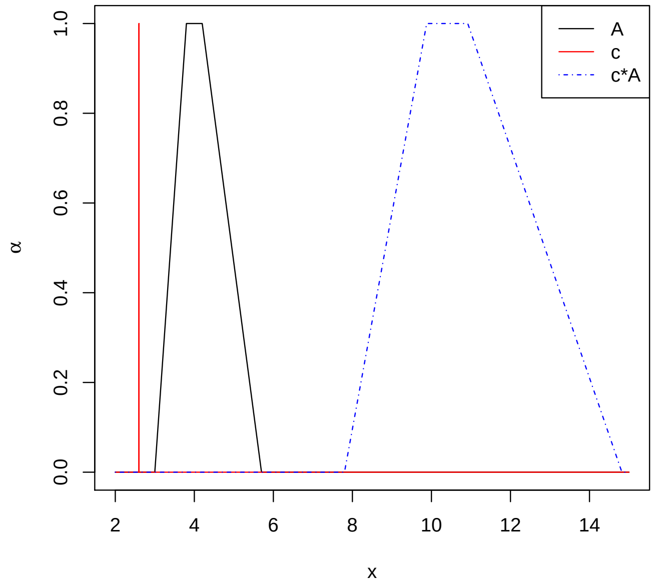

Figure 2.11: Product of a trapezoidal fuzzy number by a scalar.

2.2.5 Fuzzy random variables

Fuzzy random variables have attracted a great interest in recent years due to their interest in probabilistic studies and for practical applications. We first give the notion of random vector.

Definition 2.2.10 (Random vector) *Let* $(\Omega, \mathfrak{A}, P)$ *be a probability space and let* $(\mathbb{R}^N, \mathfrak{B})$ *be a measure space and its Borel* σ*-algebra. A measure function* $X : \Omega \to \mathbb{R}^N$ *is a* random vector*. If* $N = 1$ *we say that* X *is a* random variable*.*

Several notions of fuzzy random variables (for short, FRVs) can be found in the literature. The most widely used are: the definition due to Kwäkernaak [34], who considered a FRV as a fuzzy perception of a crisp random variable, and the one due to Puri and Ralescu [41], who regarded FRVs as random fuzzy sets. Following this last point of view, given be a probability space $(\Omega, \mathfrak{A}, P)$, a mapping $\mathcal{X} : \Omega \to \mathcal{F}$ is a FRV on $\mathbb{R}$ if for all $\alpha \in \mathbb{I}$ the α-level mappings

$\mathcal{X}_\alpha : \Omega \to \mathcal{K}(\mathbb{R})$ (where $\mathcal{K}(\mathbb{R})$ stands for the family of all nonempty compact subintervals of $\mathbb{R}$) defined so that for all $\omega \in \Omega$,

$$\mathcal{X}_\alpha(\omega) = (\mathcal{X}(\omega))_\alpha$$

are convex compact random subsets of $\mathbb{R}$.

Classical real-valued random variables are called *crisp* random variables in the fuzzy context.

In our context, we will use the following notions.

Definition 2.2.11 *(*Trapezoidal fuzzy random variable*) A function $\mathcal{X} : \Omega \to \mathcal{T}$ is a trapezoidal fuzzy random variable (TFRV) if its representation*

$$(x^I/x^J/x^R/x^S) : \Omega \to \mathbb{R}^4$$

is a random vector.

Sometimes it is interesting to summarize the information of a fuzzy random variable into a value that allows us to obtain its expected value to better understand of its behavior, to make comparisons, etc.

Definition 2.2.12 *(*Expected value of a trapezoidal fuzzy random variable*) The expected value of a trapezoidal fuzzy random variable $\mathcal{X}$, is the unique fuzzy set* $\mathrm{E}[\mathcal{X}]$ *in $\mathcal{T}$ whose representation is* $(\mathrm{E}[x^I]/\,\mathrm{E}[x^J]/\,\mathrm{E}[x^R]/\,\mathrm{E}[x^S])$.

2.3 Background on classical regression analysis

Regression analysis is a powerful tool that has a wide variety of applications in areas such as Finance, Psychology, Social Sciences, Biomedicine, Engineering, etc. This methodology is used to find the relationship between two or more quantitative variables (see [11, 63]). For the conventional regression analysis we can find several methods in the literature among which the least squares method is probably the most used procedure in research that is based on experience and experimentation, empirical research. The field of linear regression provides a basic theory for a variety of important statistical techniques, such as Discriminant Analysis, Experimental Designs, etc.

Nowadays regression methods continue to be an area of active research. In recent decades, new methods have been developed for robust regression involving correlated responses such as time series and growth curves, regression in which the predictor variables are curves, images, graphs, or other complex data objects. Regression methods accommodates various types of missing data, nonparametric regression, Bayesian methods for regression, regression in which the predictor variables are measured with error, regression with more predictor variables than observations and casual inference with regression.

Many statistical programs have contributed to the comprehensive understanding and use of this technique providing a large number of procedures for fitting different types of regression models.

This section provides a simple introduction to regression analysis.

2.3.1 Problem setting

The general problem of regression is to find the relationship between a dependent variable (also called output, endogenous or response) Y and a set of independent variables $X = (x_1, ..., x_N)$ (also called input, exogenous, explanatory or predictor variables). In general, if $N = 1$, that is, there is a single explanatory variable, the problem is referred to as simple regression whereas if there is more than one explanatory variables, $N > 1$, is named multiple regression.

In practice, the knowledge of one or more the variables can help us to infer, in a greater or lesser degree, on the value of another, saying then that there is a statistical relationship or stochastic dependence between them. Thus, regression analysis makes use of mathematical models with the objective of analyzing and identifying significant correlations and relationships of dependence between the variables. This model is used to predict the behavior of a dependent variable for given values of the independent variables.

Formally, given a data set $\{X_i, y_i\}_{i=1}^{n}$ obtained from (X, y) where $X_i \in \mathbb{R}^N$ and y_i is the value of the response variable y corresponding to $X_i \in \mathbb{R}^N$ and given a function $f(X, \underline{a})$, the goal is to find the parameter vector $\underline{a}$ such that

$$y_i = f(X_i, \underline{a}) \quad i = 1, 2, ..., n.$$

2.3.2 Loss functions

The solution to the problem is obtained by defining a loss function which measures the prediction errors between y_i and $f(X_i, \underline{a})$. The usual choice of loss function is the norm L_p

$$L_p(y, f(X, \underline{a})) = \| \, y - f(X, \underline{a}) \, \|^p \tag{2.6}$$

where p is a positive number. L_1-norm loss function is also known as least absolute deviations or least absolute errors and L_2-norm loss function is also known as least squares error. The differences of L_1-norm and L_2-norm as a loss function can be summarized as follows [25, 33, 35, 40, 50, 51]:

- **Robustness**: The method of least absolute deviations finds applications in many areas, due to its robustness compared to the least squares method. Least absolute deviations is robust, that is, it is resistant to outliers in the data. Since a L_2-norm squares the error, the model will see a much larger error than the L_1-norm, so the model is much more sensitive to outliers. This may be helpful in studies where outliers may be safely and effectively ignored. If it is important to pay attention to any and all outliers, the method of least squares is a better choice.

- **Stability**: The instability property of the method of least absolute deviations means that, for a small horizontal adjustment of a datum, the regression line may jump a large amount. The method has continuous solutions for some data configurations; however, by moving a datum a small amount, one could "jump past" a configuration which has multiple solutions that span a region. After passing this region of solutions, the least absolute deviations line has a slope that may differ greatly from that of the previous line. In contrast, the least squares solutions are stable in that, for any small adjustment of a data point, the regression line will always move only slightly; that is, the regression parameters are continuous functions of the data.

- **Computational efficiency**. Sometimes L_1-norm does not have an analytical solution, i.e., there is more than one solution or there are infinitely many solutions. However L_2-norm always has one solution. This allows the L_2-norm solutions to be calculated computationally efficiently.

Least Squares Regression	**Least Absolute Deviations Regression**
Not very robust	Robust
Stable solution	Unstable solution
Always one solution	Possible multiple solutions

2.3.3 Least squared method for linear regression

The case in which f is a linear function is likely the most important case. In some instances the model is non-linear but it can be *linearized* using transformations of variables.

The classical linear regression model, for each observation, assumes the following expression:

$$y_i = \beta_0 + \beta_1 x_{1i} + ... + \beta_N x_{N\,i}, + \varepsilon_i \quad i = 1, 2, ..., n$$

where the parameters β_i are unknown and the random residuals ε_i satisfy:

1. The expected value of the residuals is zero: $\mathrm{E}[\varepsilon_i] = 0$, for all $i = 1, 2, ..., n$.
2. The variance of the residuals is a constant value: $\mathrm{Var}[\varepsilon_i] = \sigma^2$ for any $i = 1, 2, ..., n$. This property is known as *homoscedasticity*.
3. There is no autocorrelation between the random terms of the different elements in the sample: $Cov(\varepsilon_i, \varepsilon_j) = 0$, for all $i \neq j$ $(i, j = 1, ..., n)$.

In matrix form the linear model can be written as:

$$y = X\beta + \varepsilon$$

where y is the vector of dimension n with observations of the response variable, X is the matrix of dimension $n \times (N+1)$ with the observations of the explanatory variable and a column of ones and β is the vector which contains $N+1$ unknown parameters we want to estimate, ie:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1N} \\ 1 & x_{21} & \cdots & x_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & \cdots & x_{nN} \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_N \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

To find a solution to the problem we can consider a loss function as seen in equation (2.6). However, the method of least squares estimation is the most widely used. This method minimizes the sum of squared residuals, i.e., minimizes the expression

$$SSR = \sum_{i=1}^{n} \varepsilon_i = \varepsilon^t \varepsilon = (y - X\beta)^t (y - X\beta) = y^t y - 2y^t X\beta + \beta^t \left(X^t X\right) \beta.$$

Differentiating with respect to β and equating to zero we obtain

$$\hat{\beta} = \left(X^t X\right)^{-1} X^t y. \tag{2.7}$$

The least squares estimator verifies to be the minimum variance unbiased estimator (efficient). This result is called Gauss-Markov theorem and it justifies the least squares method is widely used. However, this result does not guarantee that the variance of the estimator is necessarily the smallest. The Gauss-Markov theorem ensures that the least-squares estimators are the best in the class of estimators that are unbiased and linear functions of the observations, but does not guarantee that these estimators are better than other estimators that do not belong to the previous class. This means that there may exist a biased estimator which has lower variance than the estimator calculated by least squares.

The unbiased estimator for the unknown variance σ^2 is

$$\hat{\sigma}^2 = \frac{1}{n + N - 1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$

where the values for $\hat{y}_i$ are derived from (2.7). The square root of $\hat{\sigma}^2$ is sometimes called the standard error of the regression and it has the same units that the endogenous variable. It represents the standard deviation of y compared to the regression equation and can be used as an indicator of the fitness of the model. Given that it has the same units than the endogenous variable it cannot be used to compare models with different endogenous variables. Furthermore, due to $\hat{\sigma}^2$ depends on the residuals of the model, a violation of the previously mentioned hypothesis or a model misspecification could have a very significant impact on the value of $\hat{\sigma}^2$ as an estimator of σ^2. Because of this, $\hat{\sigma}^2$ is a estimator of σ^2 that depends on the model.

Using the linear model equation and the calculated estimation for the vector of unknown parameters, we can obtain predictions of y, y_i, for given values of the independent variable X as follows:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{1i} + ... + \hat{\beta}_N x_{N\,i} = X\hat{\beta}$$

The usual interpretation of a regression coefficient, $\hat{\beta}_i$, is that it provides an estimation of the effect of a one unit change in an independent variable, X_i, holding the other variables constant.

Since each estimation is subject to a margin of error, a confidence interval can be calculated. To study the behavior of the parameters, these intervals and hypothesis tests are solving assuming a certain distribution for the residues. Generally, it is assumed that the distribution is Normal, that is,

$$\varepsilon_i \rightsquigarrow N(0,1).$$

Under these conditions it can be seen that $\hat{\beta}$ leads a multivariate normal distribution with mean vector

$$\left(X^t X\right)^{-1} X^t(X\beta) = \beta$$

and covariance matrix

$$\left(X^t X\right)^{-1} X^t(\sigma^2 I)X\left(X^t X\right)^{-1} = \sigma^2\left(X^t X\right)^{-1}.$$

Hence,

$$\hat{\beta} \rightsquigarrow N\left(\beta, \sigma^2\left(X^t X\right)^{-1}\right).$$

From this distribution we can obtain a confidence interval $(1-\alpha)\,100\%$ for values of β and solve hypothesis testing on this parameter. To solve a regression problem requires a minimum number of observations and when this number increases the degrees of freedom also increases, and then the intervals confidence become more accurate.

2.3.4 The coefficient of determination

The overall evaluation of a linear regression can be made by the residual variance, which is an index of the accuracy of the model. The fitted model is not representative when the residual variance is large (the differences between the fitted and observed values, i.e., the errors, are large). However, this measure is not useful for comparing linear regression models of different variables, because it depends on the units of measurement of the dependent variable.

The variability of the data set is measured through different sums of squares: the total sum of squares, SST, the regression sum of squares SSE (also called the explained sum of squares) and the sum of squares of the residuals SSR (also called the residual sum of squares). Therefore we can define a relative indicator of dispersion which is a more appropriate measure of goodness of fit of a model. The most general definition of the coefficient determinant is

$$R^2 = 1 - \frac{SSR}{SST}$$

In some cases, for example, for linear models, the total sum of squares is equal to the sum of the two other sums of squares and the above definition of R^2 is equivalent to

$$R^2 = \frac{SSE}{SST} = \frac{SSE/n}{SST/n}$$

In this form, R^2 is the ratio of the variance explained by the regression to the total variance and it is often interpreted as the proportion of response variability explained by the regressors in the model.

The coefficient of determination has the property that only takes values between 0 and 1 ($0 \leq R^2 \leq 1$) (since the total variance of the dependent variable is equal to the sum of the explained variance and residual variance) and it is usual to express it as a percentage. It is easy to obtain that if the fitted line pass through all the observed points, then the residual variance is zero and $R^2 = 1$, in this case the regression model explains 100% of the variability of y. Thus, values of R^2 close to 1 mean that the greater part of the variability of response is explained by the regression model. On the other hand, if the variance explained by the regression is zero, then $R^2 = 0$. This happens when the independent variable does not explain any variation in the dependent variable.

If the parameters are estimated by least squares method, R^2 increases as the number of variables in the model increases (R^2 is monotone increasing with the number of variables included, i.e., it will never decrease). This leads to the alternative approach of looking at the adjusted R^2. It will be denoted by $\bar{R}^2$ and defined as

$$\bar{R}^2 = 1 - (1 - R^2)\frac{n-1}{n-N-1}$$

2.3.5 Linear regression diagnostics

An aspect that is often forgotten is that the regression models are based on certain assumptions about the data that are not always fulfilled, so it is necessary to determine whether the basic assumptions of the model are satisfied in our data before performing inference. It is known as *model diagnosis.*

There are several principal assumptions which justify the use of linear regression models for purposes of inference or prediction:

i) Linearity and additivity of the relationship between dependent and independent variables.

ii) Normality of the error distribution.

iii) Homoscedasticity (constant variance) of the errors.

iv) Statistical independence of the errors.

v) The independent variables are not linearly correlated; that is, there is no multicollinearity (this assumption is made if the regression model includes more than one independent variables)

Model diagnostic procedures involve both graphical methods and formal statistical tests. These procedures allow us to explore whether the assumptions of the regression model are valid and decide whether we can trust subsequent inference results. We will need to test the previous assumptions.

Linearity and additivity assumption

The hypothesis of linear relationship between the response variable and the explanatory variables is a basic assumption in the regression analysis. This linear relationship might not exists and hence a linear regression model would not be appropriate. In this case predictions are likely to be seriously in error, especially when you extrapolate beyond the range of the sample data. Nonlinearity is usually evident in a plot of observed versus predicted values. In multiple regression models, nonadditivity may be revealed by systematic patterns in plots of the residuals versus individual independent variables.

Normality assumption

The normality assumption of random errors is necessary to make inferences regarding the parameters as well as for the construction of prediction intervals. This assumption is not absolutely essential if the purpose is the estimation of the parameters of the model. The least squares estimators are optimal, regardless of the probability distribution that random error leads. Moreover, if the random variable is normally distributed, it can be shown that the least squares estimators tend to be also normally distributed, as the sample size increase indefinitely. The diagnosis can be made using a normal probability plot or normal quantile plot of the residuals.

Homoscedasticity assumption

The term σ^2 may not be constant, and may vary between different observations. It is called the problem of *heteroscedasticity*. If there is heteroscedasticity, the minimum squares estimators of the regression coefficients remain unbiased but their variance would not be the minimum possible. Violation of this assumption means that the Gauss-Markov theorem does not apply. This implies that the regression coefficients would have larger standard errors and confidence intervals and tests would be severely affected due to a loss of sensitivity. There are several tests that can be used to detect the presence of heteroscedasticity.

No autocorrelation and independence assumption

The most common meaning of the term autocorrelation refers to the correlation between the elements of a series of observations ordered in time or space. In the context of the regression, the autocorrelation is assumed to not exist in the random errors. Otherwise, this would be the problem of existence of autocorrelation. The impact of autocorrelation on the estimators can be severe with the confidence intervals and hypothesis tests based on the t-student and the F-Snedecor distributions not being appropriate. Typically, the underestimation of the variance causes a false impression of accuracy. Durbin-Watson statistic is a test statistic used to detect the presence of autocorrelation.

No multicollinearity assumption

This hypothesis requires that none of the explanatory variables can be obtained as a linear combination of the others independent variables. If a linear combination of the columns of X is close to 0 (ie two or more predictor variables in the multiple regression model are highly correlated) then the calculation of $(X^tX)^{-1}$ becomes unstable (X^tX may not have an inverse) and this produces a significant increase of the absolute value of the coefficients and the standard errors of the affected coefficients tend to be large. The best regression models are those in which the independent variables correlate minimallywith each other but significantly with the dependence variable. In this case we must be alert to large standard errors of the regression coefficients or to large changes in the estimated regression coefficients when a predictor variable is added or deleted. However some authors have suggested the variance inflation factor is a more formal way to detect multicollinearity.

Most statistical software provide charts and statistics that test whether these assumptions are satisfied for any given model.

2.3.6 Variable selection in regression models

In multiple regression a common goal is to determine which independent variables contribute significantly to explaining the variability in the dependent variable. The model that contains all independent variables will give the maximum value of the determination coeficient (R^2), but not mean that all the independent variables contribute significantly to explaining the variability in the dependent variable. Therefore a first question is whether all the variables should enter the regression model and, if not, we want to know what variables should enter and what variables should exclude. Intuitively it seems good to introduce in the model all significant explanatory variables (according to the individual t-test) to fit the model with all possible variables. But this approach is not appropriate as the variance of the model depends of the number of variables and the variance of the parameters increases as the number of variables increases. There may also be problems of multicollinearity when there are many explanatory variables. Then the selection of important independent variables may be useful in many practical problems, for example in situations we can have a large set of potential explanatory variables.

To address this problem there are several statistical procedures. A widely used technique is the *Stepwise regression* [28]. Assuming that there is indeed a linear relation between the independent variables and the dependent variable, this technique is a semi-automated process of building a model by successively adding or removing variables. This algorithm has the advantages of a progressive introduction algorithm, but improves it, because he does not keep in the model the variables already entered in one step, avoiding problems of multicollinearity. This is a very frequently algorithm because it tends to give good results when there is a large amount of variables. The main approaches are:

- **Forward selection** The process begins with no variables in the model and tests the addition of each variable using a chosen model comparison criterion, adding the variable (if any) that improves the model the most, and repeating this process until none improves the model.
- **Backward elimination** The process begins with all potential variables and tests the deletion of each variable using a chosen model comparison criterion, deleting the variable (if any) that improves the model the most by being deleted, and repeating this process until no further improvement is possible.
- **Bidirectional elimination** This process is a combination of the above, testing at each step for variables to be included or excluded.

2.3.7 Logistic regression

Finally remark that when we have dichotomous response variable we want to predict, or we want to evaluate the association or relationship with other (more than one) independent and control variables, the procedure is known as *multivariate (binary) logistic regression* [16, 30]. Logistic regression is probably the type most widely used multivariate analysis in Life Sciences. There is no doubt that this type of regression is one of the best statistical tools with the ability to analyze data in Clinical and Epidemiological research, hence its wide use. The logistic regression is used for predicting binary outcomes of the dependent variable (treating the dependent variable as the outcome of a Bernoulli trial) rather than continuous outcomes. To solve this problem, the following transformation the standard model is made

$$\pi(x) = \frac{e^{X\beta}}{1 + e^{X\beta}}$$

The regression coefficients are usually estimated using *maximum likelihood estimation.*

2.4 Fuzzy multiple regression based on real distances

Recently the theory of fuzzy sets has been used to integrate variability and imprecision in the analysis of statistical data. In particular, the problem of regression analysis in a fuzzy environment has been discussed in the literature from different points of view and considering a variety of input and output. The conventional methods for parameters estimation are: a) linear programming methods; b) least-squares methods; and c) support vector machines (SVMs).

Tanaka et al. [53] first introduced the linear regression analysis. They proposed a formulation of possibilistic linear regression analysis and determined the fuzzy parameters by applying linear programming models [54]. However it is known that this method has several drawbacks [17,31,62]. It is very sensitive to outliers and some coefficients tend to become crisp due to the characteristics of linear programming. Another observed problem is, when the coefficients are FNs the spread of the estimated response becomes wider as the magnitudes of the explanatory variables increase, even though the spreads of the observed responses decrease, or as more observations are included in the model. This contradicts intuition. To overcome this problem, Diamond [17], Wu and Tseng [62] and Kao and Chyu [31] considered numeric coefficients to describe the fuzzy relationship between the fuzzy response variable and fuzzy (or numeric) exploratory variables. All of them, and the methodologies proposed in Chen and Hsueh [14] and in Roldán et al. [46], used the concept of least squares to determine the regression coefficients. Another method is SVR [29], that is, a nonparametric, regularized, and nonlinear regression tool used for classification and regression analysis. See Näther [38], Chen y Hsueh [14] y Roldán et al. [46] for an extensive review of the main approaches to fuzzy regression.

Under these conditions, we highlight the following facts:

- From the statistical point of view, the use of least squares method in the analysis of traditional regression has numerous advantages: software for solving linear and nonlinear cases are available in many statistical packages, it is easy to use, handle and understand, has advantages in computation, is more suitable for learning, etc.
- Many practical applications also show the importance of the case in which some variables can be observed accurately, but others present vagueness/ imprecision, and they described in an approximate way rather than by exact values.
- To represent this imprecision many researchers prefer to use simple, such as triangular or trapezoidal FNs that satisfy the need to model and fuzzy problems are easier to set and manage forms.

Taking into account the previous considerations, this report focuses on the observational situation where the response variable is fuzzy and exploratory variables are crisp. In this context, Roldán et al. [46] proposed a procedure based on the results of traditional least squares regression verifying:

- it is easy to compute in practice,
- it may be applied in different contexts,
- it allows us to solve linear and non-linear fuzzy regression problems, and

- a comparative study with the proposals of different authors shows that the total error obtained is similar or lower to other complex techniques.

This report extends the fuzzy regression proposed in that paper to multiple regression analysis. A fuzzy regression model obtained by least squares method is proposed and discussed considering the most common membership functions. In order to illustrate the regression method, some results obtained by analyzing economic data of China are given in the following chapter.

2.4.1 Some real distances measures between trapezoidal fuzzy numbers

In this section we will give some notions about distance measures that we will use for trapezoidal fuzzy regression models.

Consider the set of trapezoidal numbers, $\mathcal{T}$ and the partial order $\mathcal{A} \leq \mathcal{B}$ if, and only if, $A^c \leq B^c$, $A^m \leq B^m$, $A^\ell \leq B^\ell$ and $A^s \leq B^s$, where $\mathcal{A} = \text{Tra}(A^c, A^m, A^\ell, a^s)$ and $\mathcal{B} = \text{Tra}(B^c, B^m, B^\ell, B^s)$. Let $P = (p_0, p_1, p_2, p_3)$, $K = (k_0, k_1, k_2, k_3)$ and $Q = (q_0, q_1, q_2, q_3)$ with $p_j, k_j, q_j > 0$ for $j = 0, 1, 2, 3$. We define $d = d_{PKQ} : \mathcal{T} \times \mathcal{T} \to \mathbb{R}_0^+$ for all $\mathcal{A}, \mathcal{B} \in \mathcal{T}$,

$$\begin{aligned} D_{RRM}(\mathcal{A}, \mathcal{B}) = \big(q_0 \left| (A^c)^{k_0} - (B^c)^{k_0} \right|^{p_0} + q_1 \left| (A^m)^{k_1} - (B^m)^{k_1} \right|^{p_1} + \\ q_2 \left| (A^\ell)^{k_2} - (B^\ell)^{k_2} \right|^{p_2} + q_3 \left| (A^s)^{k_3} - (B^s)^{k_3} \right|^{p_3}\big) . \end{aligned} \tag{2.8}$$

Then $(\mathcal{T}, D_{RRM})$ is a semimetric space on $\mathbb{R}_0^+$ and if $0 < p_j \leq 1$ for $j = 0, 1, 2, 3$, then D_{RRM} is a metric space on $\mathbb{R}_0^+$.

For example, the center and the spreads of the distance measure defined between two TFNs plotted in in Figure 5.4.a, are shown as blue areas in Fig. 5.4.b for the particular case $p_j = k_j = 1$, $j = 0, 1, 2, 3$, $q_1 = q_2 = 1$ and $q_0 = q_3 = \frac{1}{2}$.

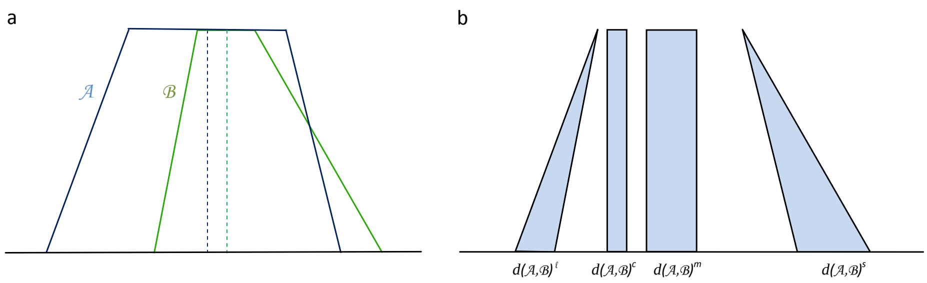

Figure 2.12: (a) Two trapezoidal fuzzy numbers and (b) the center and the spreads of the distance measure, D_{RRM}, between them for the particular case $p_j = k_j = 1$, $j = 0, 1, 2, 3$, $q_1 = q_2 = 1$ y $q_0 = q_3 = \frac{1}{2}$.

Another concept of distance between TFNs introduced by Yang and Ko [64] is as follows,

$$\begin{aligned} D^2_{YK}(\mathcal{A},\mathcal{B}) &= (A^c - B^c)^2 + ((A^c - A^m) - (B^c - \lambda_1 B^m))^2 + \\ &\left((A^c - A^\ell) - \left(B^c - \lambda_2 B^\ell\right)\right)^2 + ((A^c - A^s) - (B^c - \lambda_3 B^s))^2 . \end{aligned} \tag{2.9}$$

where $\lambda_i = \frac{1}{2}$ with $i = 1, 2, 3$. $(\mathcal{T}, D_{YK})$ is a metric space on $\mathbb{R}^+_0$.

2.4.2 Fuzzy multiple linear regression

In this section we analyze a linear regression model with a fuzzy response variable and N crisp explanatory variables. The response variable is described by TFNs.

Trapezoidal fuzzy random variables are introduced to model those random experiments in which the characteristic observed for each output may be described with a particular class of FNs determined by four random values: the center value, the center spread and the left and right spreads.

Let us consider a random experiment in which a trapezoidal fuzzy random variable $\mathcal{Y}$ and N real random variables $X = (x_1, ..., x_N)$ are observed in n statistical units, $\{X_i, \mathcal{Y}_i\}_{i=1}^n$ where $X_i = (x_{1i}, ..., x_{N\,i})^t$. Since $\mathcal{Y}$ is determined by Tra(y^c, y^m, y^ℓ, y^s), the linear regression model refers to these four real values which define a TFN.

The center y^c can be related to variables $x_1, ..., x_N$ through the classical regression model. However, the estimated values of variables y^m, y^ℓ and y^s must be positive so we should be careful when performing the corresponding regression analysis. We propose to model the spreads using the classical regression method finding a model that only takes non-negative values, it is well defined in the domain of X and in the set of values that will make the predictions the problem would be solved, D.

Under these conditions, the regression problem that we will solve in this chapter can be expressed as

$$\begin{cases} y^c = X\underline{a}_c + \varepsilon_c, \\ f\left(y^m\right) = X\underline{a}_m + \varepsilon_m, \\ g\left(y^\ell\right) = X\underline{a}_l + \varepsilon_\ell, \\ h\left(y^s\right) = X\underline{a}_s + \varepsilon_s, \end{cases} \tag{2.10}$$

where $\varepsilon^c, \varepsilon^m, \varepsilon^\ell$ y ε^s are random variables with crisp values such that $\mathrm{E}[\varepsilon^c|X] = \mathrm{E}[\varepsilon^m|X] = \mathrm{E}[\varepsilon^\ell|X] = \mathrm{E}[\varepsilon^s|X] = 0$ and whose variance $\sigma^2_{\varepsilon_c}$, $\sigma^2_{\varepsilon_m}$, $\sigma^2_{\varepsilon_l}$ and $\sigma^2_{\varepsilon_s}$, are finite and strictly positive, $\underline{a}_c = (a_{c0,}a_{c1}, \ldots, a_{cN})'$, $\underline{a}_m = (a_{m0}, a_{m1}, \ldots, a_{mN})'$, $\underline{a}_\ell = (a_{\ell 0}, a_{\ell 1}, \ldots, a_{\ell N})'$ y $\underline{a}_s = (a_{s0}, a_{s1}, \ldots, a_{sN})'$ are $(N+1)\times 1$-vectors of the parameters related to X. The covariance matrix of the vector of explanatory variables X will be denoted by Σ_X and the covariance matrix of the vector $(\varepsilon_c, \varepsilon_m, \varepsilon_l, \varepsilon_s)$ by Σ. Considering the above expected values, we conclude that $\varepsilon^c, \varepsilon^m, \varepsilon^\ell$ y ε^s are uncorrelated random variables with the explanatory variables.

To solve the above regression problem we will consider the distance measure defined in (2.8) with $p_j = 2$ and $q_j = k_j = 1$ for $j = 0, 1, 2, 3$.

Proposition 2.4.1 Let $\mathcal{Y}$ be a trapezoidal fuzzy random variable and X a vector of N random variables verifying the model (2.10), then we have

$$\begin{aligned}\underline{\hat{a}}_c &= \left(X^t X\right)^{-1} X^t y^c,\\ \underline{\hat{a}}_m &= \left(X^t X\right)^{-1} X^t f\left(y^m\right),\\ \underline{\hat{a}}_l &= \left(X^t X\right)^{-1} X^t g\left(y^l\right),\\ \underline{\hat{a}}_s &= \left(X^t X\right)^{-1} X^t h\left(y^c\right).\end{aligned}$$

Proof. According to the definition of the distance measure between the TFNs in (2.8) with $p_j = 2$ and $q_j = k_j = 1$ for $j = 0, 1, 2, 3$, the objective function is:

$$\Phi = \|y^c - \widehat{y^c}\|^2 + \left\|f\left(y^m\right) - \hat{f}\left(y^m\right)\right\|^2 + \left\|f\left(y^l\right) - \hat{f}\left(y^l\right)\right\|^2 + \left\|f\left(y^s\right) - \hat{f}\left(y^s\right)\right\|^2$$

where $\|.\|^2$ denotes the Euclidean square norm. Therefore, the problem can be divided into four independent least square problems that would lead us to the result.

Proposition 2.4.2 Under the previous conditions, the distance measure D_{RRM} defined in equation (2.8) for $p_i = 2$ y $k_i = q_i = 1$ for $i = 0, 1, 2, 3,$ produces the same estimators $\underline{\hat{a}}_c$, $\underline{\hat{a}}_m$, $\underline{\hat{a}}_l$, $\underline{\hat{a}}_s$ that the distance introduced by Yang and Ko D_{YK} (see equation (2.9)).

This proof can be found in [22].

Proposition 2.4.3 Under the above conditions, the estimators $\underline{\hat{a}}_c$, $\underline{\hat{a}}_m$, $\underline{\hat{a}}_l$, $\underline{\hat{a}}_s$, are unbiased, strongly consistent and efficient.

These properties are direct consequences that the fuzzy regression problem has been reduced to four problems of classical regression.

Therefore confidence regions and hypothesis tests on the regression parameters can be solved in the classical way.

2.4.3 The determination coefficient

Let $\mathcal{Y}$ be a trapezoidal fuzzy random variable and let X be a vector of N crisp random variables satisfying model (2.10). Then once obtained optimal regression models taking into account the results of the previous sections for the center and the spreads, the fuzzy linear regression model is obtained as:

$$\hat{\mathcal{Y}} = \mathrm{Tra}(\hat{y}^c, \hat{y}^m, \hat{y}^\ell, \hat{y}^s).$$

To assess the goodness of fit of the model we will use a coefficient of determination that we define below. Let us note by SSE_i the sums of squares explained, by SSR_i the residual sums of squares corresponding to the fittings $\hat{y}^i$, and by SST_i the total sums where $i = c, m, \ell, s$. It is cleat that for $i = c, m, \ell, s$, $\mathrm{SST}_i = \mathrm{SSE}_i+ \mathrm{SSR}_i$. Therefore, considering the definition of the distance function given in equation (2.8), the sum of squares explained, and the total residual, denoted by SSE, SSR and SST, respectively, of the regression model are obtained as follows:

$$\begin{aligned}
SSE &= D^2_{RRM}(\hat{\mathcal{Y}}, \bar{\mathcal{Y}}) = SSE_c + SSE_m + SSE_\ell + SSE_s,\\
SSR &= D^2_{RRM}(\mathcal{Y}, \hat{\mathcal{Y}}) = SSR_c + SSR_m + SSR_\ell + SSR_s,\\
SST &= D^2_{RRM}(\mathcal{Y}, \bar{\mathcal{Y}}) = SST_c + SST_m + SST_\ell + SST_s.
\end{aligned}$$

Definition 2.4.4 (Determination coefficient) *Let $\mathcal{Y}$ be a trapezoidal fuzzy random variable and X a vector of N crisp random variables satisfying model (2.10). The determination coefficient $\mathcal{R}^2$ is the part of the sum of squares explained by the regression model, that is defined,*

$$\mathcal{R}^2 = 1 - \frac{SSR}{SST} = \frac{SSE}{SST} = \frac{D^2_{RRM}(\hat{\mathcal{Y}}, \bar{\mathcal{Y}})}{D^2_{RRM}(\mathcal{Y}, \bar{\mathcal{Y}})} = \frac{SSE_c + SSE_m + SSE_\ell + SSE_s}{SST_c + SST_m + SST_\ell + SST_s}.$$

Since the distance function allows the decomposition SST = SSE + SSR, this coefficient represents the the proportion of the total variation in the fuzzy response variable that is explained by the fuzzy regression model.

2.4.4 Fuzzy nonlinear regression

Nonlinear regression is a form of regression analysis in which observational data are modeled by a function which is a nonlinear combination of the model parameters and depends on one or more independent variables. While the linear model describes some of the relationship between the variables, a nonlinear model is likely to do a much better fit.

Let us consider the fuzzy simple regression problem, that is, we consider only one independent variable. This can be useful for multiple regression analysis in the case that stepwise regression only introduces one explanatory variable in the model (this case is happening in our case study).

Formally we are interested in a random experiment in which a trapezoidal fuzzy random variable $\mathcal{Y}$ and a crisp random variable x are observed. If we consider $p_j = 2$ and $k_j = q_j = 1$ in the distance measure defined in (2.8), for $j = 0, 1, 2, 3$, we obtain the usual mean square error, so the minimization process is reduced to apply the widely used and well known method of least squares to estimate the functions y_x^c, y_x^m, y_x^i or/and y_x^s. The use of this method is justified by its important properties and because it allows to use different types of settings (linear, parabolic, polynomial, exponential, etc.) that are already implemented in the programs and tools (some of them are free software) used for the analysis of statistical data. Our goal is not to improve this method but use it to make our method more powerful and useful.

If we plot the data and observe a curvilinear trend, we can then proceed to fit a model using two types of models:

- Nonlinear models such as exponentials, growth curves, and other types of functions, and

- Polynomial models involving powers of x.

Statistical packages, such as *Statgraphics* (distributed by *Statpoint Technologies*, Inc., *www.Statgraphics.com*), fit a variety of functional forms, listing the models in decreasing order of R-squared. Another approach to fitting a nonlinear equation is to consider polynomial functions of x. Polynomials have the attractive property of being able to approximate many kinds of functions.

In this conditions, we may fit a nonlinear model for y_x^c, y_x^m, y_x^i or/and y_x^s taking into account a least squares criterion. However we must take into account that the models spreads, y_x^m, y_x^i, y_x^s, does not take negative values on the domain of x. Therefore between the different functions that could be obtained using a statistical package, we should be careful and choose that model that remaining significant, it only takes non-negative values on the domain of x.

2.4.5 Fuzzy regression when the expected values of the spreads do not depend on the explanatory variables

If we cannot find a model that is significative and does not take negative values for the center, left or right spread, we can estimate them by positive constants considering an appropriate error function as proposed below.

Henceforth we will assume that the expected values of the center spread, $M_1, ..., M_n$, do not depend on the explanatory variables and it has not been possible to fit a significant model for the center spread (similarly for the left or the right spread) and we are interested in estimating it by a constant. Next theorem gives the estimator $\hat{M}$. According to the distance measure defined in (2.8), the error can be expressed as: $e = \sum_{i=1}^{n} \left| x^{k_1} - M_i^{k_1} \right|^{p_1}$ where $e : (0, \infty) \to (0, \infty)$ for all $x \in (0, \infty)$. For simplicity we will denote $k_1 = k$ and $p_1 = p$.

Proposition 2.4.5 In the previous conditions, the estimator $\hat{M}$ for M, can be obtained as follows:

a) Si $p > 1$, $\hat{M}$ is the only positive solution of the equation system:

$$\hat{M} \rightarrow \left\{ \begin{array}{c} \sum_{i=1}^{r} \left(x^k - M_i^k\right)^{p-1} = \sum_{i=r+1}^{n} \left(M_i^k - x^k\right)^{p-1} \\ \text{in the interval } \left]M_r, M_{r+1}\right] \, {}_{r=1}^{r=n-1} \end{array} \right\}. \tag{2.11}$$

b) If $p = 1$, $\hat{M}$ is the median of the spreads ($\hat{M} \rightarrow \text{median}\, M_i$).

c) If $0 < p < 1$, $\hat{M}$ is obtained by searching the following values:

$$\hat{M} \rightarrow M_j \quad \text{such that} \quad \left\{ \sum_{i=1}^{n} \left| M_j^k - M_i^k \right|^p \right\}_{j=1}^{n} \quad \text{is minimum.} \tag{2.12}$$

Furthermore, $\hat{M}$ is uniquely determined if $p > 1$, or $p = 1$ and the median of the corresponding spreads is unique or $0 < p < 1$ and corresponding values in (2.12) have a unique absolute minimum.

In the case $p = 3$, the system (2.11) is equivalent to the biquadratic system:

$$\left\{ \begin{array}{c} (2r-n)x^{2k} - 2x^k \left(\sum_{i=1}^{r} M_i^k - \sum_{i=r+1}^{n} M_i^k \right) \\ + \left(\sum_{i=1}^{r} M_i^{2k} - \sum_{i=r+1}^{n} M_i^{2k} \right) = 0 \\ \text{in the interval } \left]M_r, M_{r+1}\right] \end{array} \right\}_{r=1}^{r=n-1}.$$

In practical situations, $\hat{\text{M}}$, $\hat{\text{I}}$, or $\hat{\text{S}}$ can be calculated as the solution of a minimization problem, directly programming and minimizing the distance function (using, for example, Mathematica).

Note that all the proposed estimators are unbiased and consistent as the experiment simulation shows in [46]. The aim of previous proposition was to propose alternative estimators that can give a solution to particular real cases and that can be easily obtained.

2.4.6 An extension of the previous methodology

The approach considered in the previous sections can be generalized to the case in which the explanatory variables are also fuzzy.

Let us consider a random experiment in which a trapezoidal fuzzy random variable $\mathcal{Y}$ and N trapezoidal fuzzy random variables $\mathcal{X} = (\mathcal{X}_1, ..., \mathcal{X}_N)$ are observed in n statistical units. If we are interested in analyzing the relationship between $\mathcal{Y}$ and $\mathcal{X}$, this problem can be reduced to the previous approach considering the following explanatory crisp random vector: $X = (x_1^c, x_1^m, x_1^i, x_1^s, x_2^c, ..., x_N^c, x_N^m, x_N^i, x_N^s)$.

CHAPTER 3

Finite fuzzy numbers and finite fuzzy random variables

In [47], Roldán *et al.* did a complete study of the image of a FN. In their manuscript, they considered FNs $\mathcal{A}$ whose image (or range) $\mathcal{A}(\mathbb{R})$ is a countable (or finite) subset of $\mathbb{I}$, and they called them *discrete* FNs. Among other properties, they succeeded in proving that such family of FNs is closed under the usual operations between FNs, that is, if $\mathcal{A}$ and $\mathcal{B}$ are discrete FNs, them $\mathcal{A}+\mathcal{B}$, $\mathcal{A}-\mathcal{B}$, $\mathcal{A}\cdot\mathcal{B}$ and $\mathcal{A}/\mathcal{B}$ (if this last FN is well defined) also are discrete FNs.

One of the main aims of this report is to study a more restrictive, but great enough, subclass of FNs: the family of FNs whose image is a finite subset of $\mathbb{I}$. In this chapter we characterize each finite FN using its center and its spreads. We also show that they are dense in the set $\mathcal{F}$ of all FNs, so they can be used to approximate any fuzzy quantity.

3.1 Finite fuzzy numbers

In this section we study FNs whose image is finite. In general, given two nonempty sets X and Y, we will say that a function $f : X \to Y$ is *finite* if the image of f is a finite subset of Y. The following notion is a key piece of the current study.

Definition 3.1.1 *(Roldán* et al. *[47]) A FN $\mathcal{A}$ is* finite *if its image, $\mathcal{A}(\mathbb{R})$, is a finite subset of $\mathbb{I}$. If Λ is a finite subset of $\mathbb{I}$, we will denote by $\mathcal{F}_\Lambda$ the family of all finite FNs $\mathcal{A} \in \mathcal{F}$ such that $\mathcal{A}(\mathbb{R}) \subseteq \Lambda$.*

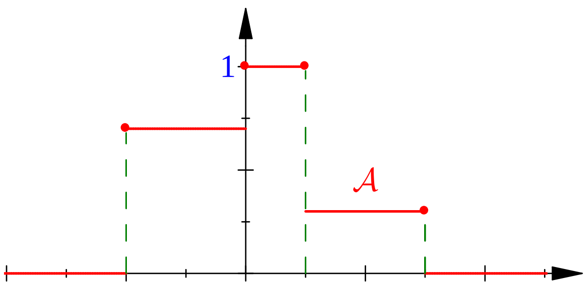

Figure 3.1: A finite fuzzy number.

As the support of any FN $\mathcal{A} \in \mathcal{F}$ is bounded, then $\mathcal{A}(x) = 0$ for all $x \in \mathbb{R} \setminus \operatorname{supp}(\mathcal{A})$. Furthermore, the normality condition implies that there exists $x_0 \in \mathbb{R}$ such that $\mathcal{A}(x_0) = 1$. This means that the points 0 and 1 are always included in the image $\mathcal{A}(\mathbb{R})$ of any FN. Thus, in order to guarantee that the set $\mathcal{F}_\Lambda$ is nonempty, we will always assume that the points 0 and 1 are included in the finite subsets Λ that we will consider inside the interval $\mathbb{I}$.

The converse property characterizes a subclass of finite FN.

Proposition 3.1.2 A FN $\mathcal{A}$ is rectangular if, and only if, its image is $\mathcal{A}(\mathbb{R}) = \{0, 1\}$.

PROOF: Taking into account (2.4), it is clear that the image of any rectangular FN is $\{0,1\}$. Conversely, suppose that $\mathcal{A} \in \mathcal{F}$ is a FN such that $\mathcal{A}(\mathbb{R}) = \{0,1\}$ and let $[a,b] = \ker \mathcal{A}$. Then a and b are real numbers such that $a \leq b$. Furthermore, $\mathcal{A}(x) = 1$ if, and only if, $a \leq x \leq b$. As $\mathcal{A}(\mathbb{R}) = \{0,1\}$, then necessarily $\mathcal{A}(x) = 0$ for all $x \in \mathbb{R} \setminus [a,b]$. Therefore $\mathcal{A}$ is the rectangular FN whose corners are a and b. ■

Lemma 3.1.3 If $\Lambda_1$and Λ_2 are finite subsets of $\mathbb{I}$ such that $\Lambda_1 \subseteq \Lambda_2$, then $\mathcal{F}_{\Lambda_1} \subseteq \mathcal{F}_{\Lambda_2}$. And if $\Lambda_1 \subset \Lambda_2$, then $\mathcal{F}_{\Lambda_1} \subset \mathcal{F}_{\Lambda_2}$.

PROOF: Given $\mathcal{A} \in \mathcal{F}_{\Lambda_1}$, we know that $\mathcal{A}(\mathbb{R}) \subseteq \Lambda_1$, and as $\Lambda_1 \subseteq \Lambda_2$, then $\mathcal{A}(\mathbb{R}) \subseteq \Lambda_2$, so $\mathcal{A} \in \mathcal{F}_{\Lambda_2}$. In particular, $\mathcal{F}_{\Lambda_1} \subseteq \mathcal{F}_{\Lambda_2}$.

If $\Lambda_1 \subset \Lambda_2$, then there is $\alpha_0 \in \Lambda_2$ such that $\alpha_0 \notin \Lambda_1$. If we consider the FN $\mathcal{A}$ given by:

$$\mathcal{A}(x) = \begin{cases} \alpha_0\ , & \text{if } 1 \leq x < 2, \\ 1, & \text{if } 2 \leq x \leq 3, \\ 0, & \text{in any other case,} \end{cases}$$

then $\mathcal{A}(\mathbb{R}) = \{0, \alpha_0, 1\} \subseteq \Lambda_2$, so $\mathcal{A} \in \mathcal{F}_{\Lambda_2}$. But as $\alpha_0 \notin \Lambda_1$, then $\mathcal{A} \notin \mathcal{F}_{\Lambda_1}$. ■

Corollary 3.1.4 If $\Lambda_1, \Lambda_2, \ldots, \Lambda_m$ are finite subsets of $\mathbb{I}$ containing the points 0 and 1, then $\mathcal{F}_{\Lambda_1} \cup \mathcal{F}_{\Lambda_2} \cup \ldots \cup \mathcal{F}_{\Lambda_m} \subseteq \mathcal{F}_{\Lambda_1 \cup \Lambda_2 \cup \ldots \cup \Lambda_m}$.

PROOF: If we denote $\Lambda_1 \cup \Lambda_2 \cup \ldots \cup \Lambda_m$ by Λ, then Λ is a finite subset of $\mathbb{I}$ such that $\Lambda_i \subseteq \Lambda$ for all $i \in \{1, 2, \ldots, m\}$. Then Lemma 3.1.3 ensures us that $\mathcal{F}_{\Lambda_i} \subseteq \mathcal{F}_{\Lambda}$ for all $i \in \{1, 2, \ldots, m\}$, so $\mathcal{F}_{\Lambda_1} \cup \mathcal{F}_{\Lambda_2} \cup \ldots \cup \mathcal{F}_{\Lambda_m} \subseteq \mathcal{F}_{\Lambda_1 \cup \Lambda_2 \cup \ldots \cup \Lambda_m}$. ■

In [47], the authors did a study of finite (and countable) FNs. They proved that if $\mathcal{A}$ and $\mathcal{B}$ are finite FNs, then $\mathcal{A} + \mathcal{B}$, $\mathcal{A} - \mathcal{B}$ and $\mathcal{A} \cdot \mathcal{B}$ are also finite FNs. Later, in [2], the authors introduced a family of distance measures $D : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ between FNs that, in many cases, is closed on the family of all finite FNs, that is, if $\mathcal{A}$ and $\mathcal{B}$ are finite FNs, then $D(\mathcal{A}, \mathcal{B})$ is also a finite FN.

3.2 A characterization of finite fuzzy numbers

In (2.5) we described an easy-to-handle method to characterize trapezoidal FNs by using their centers and spreads in the following way. If $\mathcal{A} = (a/b/c/d)$ is a trapezoidal FN, then the real number

$$A^c = \mathcal{D}_c \mathcal{A} = \frac{b+c}{2}$$

is its center, and the three spreads of $\mathcal{A}$ are the nonnegative real numbers

$$A^m = \operatorname{spr} \mathcal{A} = \frac{c-b}{2} \geq 0, \quad A^\ell = b - a \geq 0, \quad A^r = d - c \geq 0.$$

Center and spreads completely characterize the trapezoidal FN because:

$$\begin{aligned} a &= A^c - A^m - A^\ell, \\ b &= A^c - A^m, \\ c &= A^c + A^m, \\ d &= A^c + A^m + A^r. \end{aligned}$$

Hence, $\mathcal{A}$ can be written as follows:

$$\mathcal{A} = \operatorname{Tra}(A^c, A^m, A^\ell, A^r). \tag{3.1}$$

Representation (3.1) of the FN $\mathcal{A}$ has several advantages when we study a fuzzy regression methodology by involving trapezoidal FNs.

In this section, inspired by the previous alternative expression, we describe an equivalent way to represent finite FNs by using a real number (that can be interpreted as its *center*) and a finite set of nonnegative real numbers (that we will call its *spreads*).

Theorem 3.2.1 A FN $\mathcal{A}$ is finite if and only if there exist $n \in \mathbb{N}$, $\{\alpha_i\}_{i=0}^n \subset \mathbb{I}$,

$\{a_i\}_{i=1}^n \subset \mathbb{R}$, $\{b_i\}_{i=1}^n \subset \mathbb{R}$, such that

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_{n-1} < \alpha_n = 1, \tag{3.2}$$

$$a_1 \leq a_2 \leq \ldots \leq a_{n-1} \leq a_n \leq b_n \leq b_{n-1} \leq \ldots \leq b_2 \leq b_1, \tag{3.3}$$

$$\mathcal{A}(x) = \begin{cases} \alpha_n = 1, & \text{if } x \in [a_n, b_n], \\ \alpha_i, & \text{if } x \in [a_i, b_i] \setminus [a_{i+1}, b_{i+1}] \quad i \in \{1, \ldots, n-1\}, \\ 0, & \text{otherwise.} \end{cases} \tag{3.4}$$

In such a case, $\ker \mathcal{A} = [a_n, b_n]$ and $\operatorname{supp} \mathcal{A} = [a_1, b_1]$. Furthermore,

$$a_i = \underline{a}_{\alpha_i} \quad \text{and} \quad b_i = \overline{a}_{\alpha_i} \quad \text{for all } i \in \{1, \ldots, n\}. \tag{3.5}$$

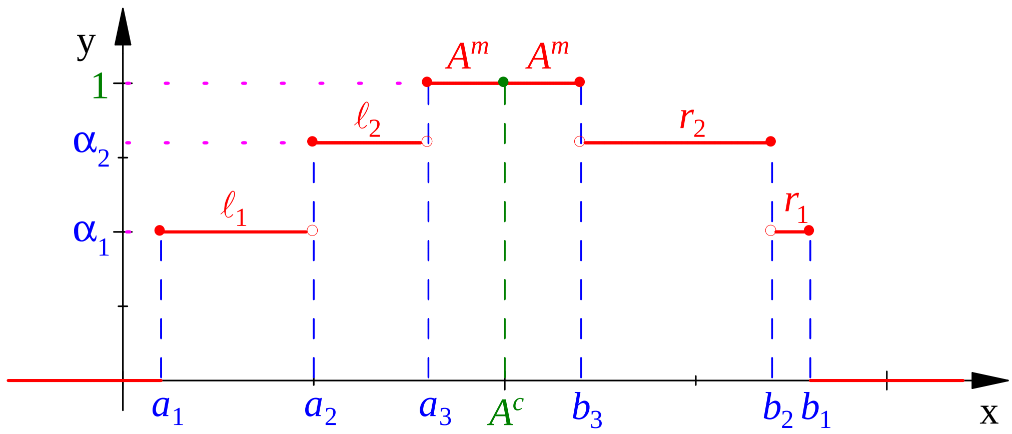

Figure 3.2: A finite FN whose image is included in $\{\alpha_0 = 0 < \alpha_1 < \alpha_2 < \alpha_3 = 1\}$.

PROOF: Assume that $\mathcal{A}$ is a finite FN and let $\mathcal{A}(\mathbb{R}) = \{0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_{n-1} < \alpha_n = 1\}$ be its finite image. Let $a_i = \underline{a}_{\alpha_i}$ and $b_i = \overline{a}_{\alpha_i}$ for all $i \in \{1, \ldots, n\}$. Hence $[a_i, b_i] = \mathcal{A}_{\alpha_i}$ for all i. Since $\alpha_i < \alpha_{i+1}$, then $[a_{i+1}, b_{i+1}] = \mathcal{A}_{\alpha_{i+1}} \subseteq \mathcal{A}_{\alpha_i} = [a_i, b_i]$, so $a_i \leq a_{i+1} \leq b_{i+1} \leq b_i$ for all $i \in \{0, 1, \ldots, n-1\}$. In particular, (3.3) holds. It is clear that if $t \in [a_n, b_n] = \mathcal{A}_{\alpha_n} = \mathcal{A}_1 = \ker \mathcal{A}$, then $\mathcal{A}(t) = 1$, and if $t \notin [a_1, b_1] = \mathcal{A}_{\alpha_1}$, then $\mathcal{A}(t) < \alpha_1$, so $\mathcal{A}(t) = \alpha_0 = 0$.

Finally, if $t \in [a_i, b_i] \setminus [a_{i+1}, b_{i+1}] = \mathcal{A}_{\alpha_i} \setminus \mathcal{A}_{\alpha_{i+1}}$ for some $i \in \{1, \ldots, n-1\}$, then $\alpha_i \leq \mathcal{A}(t) < \alpha_{i+1}$, which means that $\mathcal{A}(t) = \alpha_i$ because the image of $\mathcal{A}$ is the finite set $\{\alpha_i\}_{i=0}^n$. Hence (3.4) also holds.

Conversely, assume that there exist $n \in \mathbb{N}$, $\{\alpha_i\}_{i=0}^n \subset \mathbb{I}$, $\{a_i\}_{i=1}^n \subset \mathbb{R}$ and $\{b_i\}_{i=1}^n \subset \mathbb{R}$ such that (3.2)-(3.4) holds. In particular, by (3.4), the image of the membership function $\mathcal{A}$ is included in the finite set $\{0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_{n-1} < \alpha_n = 1\}$, so $\mathcal{A}$ is a finite FN. ■

In other words, a finite FNs whose image contains exactly $n+1$ points depends on $3n-1$ variables (that is, $\{\alpha_i\}_{i=1}^{n-1} \subset \mathbb{I}$, $\{a_i\}_{i=1}^n \subset \mathbb{R}$ and $\{b_i\}_{i=1}^n \subset \mathbb{R}$ because $\alpha_0 = 0$ and $\alpha_n = 1$ always take the same values) that satisfy the relationships (3.2)-(3.4).

We can characterize the finiteness of a FN by the finiteness of the functions that determine its level sets.

Theorem 3.2.2 A FN $\mathcal{A}$ is finite if, and only if, its corresponding functions $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ (such that $\mathcal{A}_\alpha = [\underline{a}_\alpha, \overline{a}_\alpha]$ for all $\alpha \in \mathbb{I}$) are also finite.

Proof: Suppose that a FN $\mathcal{A}$ is finite and let $\{\alpha_i\}_{i=0}^n \subset \mathbb{I}$, $\{a_i\}_{i=1}^n \subset \mathbb{R}$, $\{b_i\}_{i=1}^n \subset \mathbb{R}$ be constants as in Theorem 3.2.1. By (3.5),

$$\{a_1, a_2, \ldots, a_n\} = \left\{\underline{a}_{\alpha_1}, \underline{a}_{\alpha_2}, \ldots, \underline{a}_{\alpha_n}\right\} \subseteq \underline{a}(\mathbb{I}) \qquad \text{and}$$
$$\{b_1, b_2, \ldots, b_n\} = \{\overline{a}_{\alpha_1}, \overline{a}_{\alpha_2}, \ldots, \overline{a}_{\alpha_n}\} \subseteq \overline{a}(\mathbb{I}).$$

Let us show the contrary inclusions, that is,

$$\underline{a}(\mathbb{I}) \subseteq \{a_1, a_2, \ldots, a_n\} \qquad \text{and}$$
$$\overline{a}(\mathbb{I}) \subseteq \{b_1, b_2, \ldots, b_n\}.$$

Indeed, let $r \in \underline{a}(\mathbb{I})$ be arbitrary. Then there is $\beta_0 \in \mathbb{I}$ such that $\underline{a}(\beta_0) = r$. If $\beta_0 = 0$, then

$$r = \underline{a}(\beta_0) = \underline{a}(0) = \min(\operatorname{supp}\mathcal{A}) = \min([a_1, b_1]) = a_1 \in \{a_1, a_2, \ldots, a_n\}.$$

Next assume that $\beta_0 > 0$. Since $0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_{n-1} < \alpha_n = 1$, there is a unique $i_0 \in \{1, \ldots, n\}$ such that $\alpha_{i_0 - 1} < \beta_0 \leq \alpha_{i_0}$. Since $\mathcal{A}(\mathbb{R}) \subseteq \{\alpha_i\}_{i=0}^{n}$, then, for $t \in \mathbb{R}$,

$$\mathcal{A}(t) \geq \beta_0 \quad \Leftrightarrow \quad \mathcal{A}(t) \geq \alpha_{i_0}.$$

Therefore

$$\begin{aligned} r = \underline{a}(\beta_0) &= \min(\{t \in \mathbb{R} : \mathcal{A}(x) \geq \beta_0\}) \\ &= \min(\{t \in \mathbb{R} : \mathcal{A}(x) \geq \alpha_{i_0}\}) = \underline{a}(\alpha_{i_0}) = a_{i_0} \in \{a_1, a_2, \ldots, a_n\}. \end{aligned}$$

In any case, $r \in \{a_1, a_2, \ldots, a_n\}$, so $\underline{a}(\mathbb{I}) \subseteq \{a_1, a_2, \ldots, a_n\}$, which means that $\underline{a}(\mathbb{I}) = \{a_1, a_2, \ldots, a_n\}$. Similarly, it can be proved that $\overline{a}(\mathbb{I}) \subseteq \{b_1, b_2, \ldots, b_n\}$ and we deduce that $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ are finite functions.

Conversely, suppose that $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ are finite functions. Let $\underline{a}(\mathbb{I}) = \{a_1, a_2, \ldots, a_n\}$ and $\overline{a}(\mathbb{I}) = \{b_1, b_2, \ldots, b_m\}$ be such that

$$a_1 \leq a_2 \leq \ldots \leq a_{n-1} \leq a_n \leq x_0 \leq b_m \leq b_{m-1} \leq \ldots \leq b_2 \leq b_1,$$

where $x_0 \in \mathbb{R}$ is a point such that $\mathcal{A}(x_0) = 1$ (that is, $x_0 \in \ker \mathcal{A}$). Let $\alpha_i = \mathcal{A}(a_i)$ for all $i \in \{1, \ldots, n\}$ and let $\beta_j = \mathcal{A}(b_j)$ for all $j \in \{1, \ldots, m\}$. Hence $\Lambda = \{\alpha_i\}_{i=1}^{n} \cup \{\beta_j\}_{i=1}^{m} \cup \{0\}$ is a finite subset of $\mathbb{I}$. Let us show that $\mathcal{A}(\mathbb{R}) \subseteq \Lambda$. Let $t \in \mathbb{R}$ be arbitrary. If $\mathcal{A}(t) = 0$, then $\mathcal{A}(t) \in \Lambda$. Next, suppose that $\mathcal{A}(t) > 0$. We consider the cases $t \leq x_0$ and $t \geq x_0$.

- Suppose that $t \leq x_0$ and let $\alpha = \mathcal{A}(t) > 0$. Since $\underline{a}(t) \in \underline{a}(\mathbb{I}) = \{a_1, a_2, \ldots, a_n\}$ and $\overline{a}(t) \in \overline{a}(\mathbb{I}) = \{b_1, b_2, \ldots, b_m\}$, there are $i_0 \in \{1, \ldots, n\}$ and $j_0 \in \{1, \ldots, m\}$ such that

  $$\mathcal{A}_\alpha = [\,\underline{a}_\alpha, \overline{a}_\alpha\,] = [\,a_{i_0}, b_{j_0}\,].$$

  Since $\mathcal{A}(t) = \alpha$, then $\mathcal{A}(t) \in \mathcal{A}_\alpha = [\,\underline{a}_\alpha, \overline{a}_\alpha\,] = [\,a_{i_0}, b_{j_0}\,]$. In particular, $\underline{a}_\alpha = a_{i_0} \leq t \leq x_0$. By Lemma 2.2.6, the restriction $\mathcal{A}|_{]-\infty, x_0]}$ is a non-decreasing function. Therefore $\mathcal{A}(\underline{a}_\alpha) = \mathcal{A}(a_{i_0}) \leq \mathcal{A}(t)$. But as $\mathcal{A}(t) = \alpha$

and (2.2) implies that $\mathcal{A}\left(\underline{a}_{\alpha}\right) \geq \alpha$, then

$$\alpha \leq \mathcal{A}\left(\underline{a}_{\alpha}\right) = \mathcal{A}\left(a_{i_0}\right) \leq \mathcal{A}\left(t\right) = \alpha.$$

As a consequence, $\mathcal{A}\left(t\right) = \alpha = \mathcal{A}\left(a_{i_0}\right) = \alpha_{i_0} \in \Lambda$, so $\mathcal{A}\left(t\right) \in \Lambda$.

- If $t \geq x_0$, we can repeat the previous arguments but, in this case, $\mathcal{A}\left(t\right) = \mathcal{A}\left(b_{j_0}\right) = \beta_{j_0} \in \Lambda$.

In any case, as $\mathcal{A}\left(\mathbb{R}\right) \subseteq \Lambda$, we conclude that $\mathcal{A}$ is a finite FN. ∎

Numbers a_i and b_i in Theorem 3.2.1 can be positive or negative, but there is an order between them. For our purposes, we present a canonical representation of finite FNs involving nonnegative real numbers. To do this, given a finite FN $\mathcal{A}$, let

$$A^c = \mathcal{D}_c\,\mathcal{A} = \frac{\underline{a}_1 + \overline{a}_1}{2} = \frac{a_n + b_n}{2} \in \mathbb{R} \qquad \text{and}$$
$$A^m = \operatorname{spr}\mathcal{A} = \frac{\overline{a}_1 - \underline{a}_1}{2} = \frac{b_n - a_n}{2} \geq 0$$

be the *center* and the *central spread* of a $\mathcal{A}$, respectively. Hence

$$\ker\mathcal{A} = \left[a_n, b_n\right] = \left[A^c - A^m, A^c + A^m\right].$$

Next, let us consider the numbers

$$A^{\ell,i} = \underline{a}_{\alpha_{i+1}} - \underline{a}_{\alpha_i} = a_{i+1} - a_i \geq 0 \quad \text{and} \tag{3.6}$$
$$A^{r,i} = \overline{a}_{\alpha_i} - \overline{a}_{\alpha_{i+1}} = b_i - b_{i+1} \geq 0 \quad \text{for all } i \in \{1, 2, \ldots, n-1\}. \tag{3.7}$$

Since $A^{\ell,i}$ and $A^{r,i}$ are nonnegative numbers, we can also call $\{A^{\ell,i}\}_{i=1}^{n-1}$ the *left spreads of* $\mathcal{A}$, and $\{A^{r,i}\}_{i=1}^{n-1}$ will be called the *right spreads of* $\mathcal{A}$. Thus, a finite FN can equivalently be characterized in the following way.

Theorem 3.2.3 Given $n \in \mathbb{N}$, $\{\alpha_i\}_{i=0}^{n} \subset \mathbb{I}$ verifying $0 = \alpha_0 < \alpha_1 < \alpha_2 < \ldots < \alpha_{n-1} < \alpha_n = 1$, $A^c \in \mathbb{R}$ and $A^m, A^{\ell,1}, A^{\ell,2}, \ldots, A^{\ell,n-1}, A^{r,1}, A^{r,2}, \ldots, A^{r,n-1} \in$

$[0, \infty)$, there exists a unique finite FN $\mathcal{A}$ such that

$$\underline{a}_\alpha = \begin{cases} A^c - A^m, & \text{if } \alpha_{n-1} < \alpha \leq 1, \\ A^c - A^m - A^{\ell,n-1} - A^{\ell,n-2} - \ldots - A^{\ell,i+1}, & \text{if } \alpha_i < \alpha \leq \alpha_{i+1} \\ \qquad \text{(for some } i \in \{1, 2, \ldots, n-2\}\text{)}, & \\ A^c - A^m - A^{\ell,n-1} - A^{\ell,n-2} - \ldots - A^{\ell,1}, & \text{if } 0 \leq \alpha \leq \alpha_1; \end{cases} \tag{3.8}$$

$$\overline{a}_\alpha = \begin{cases} A^c + A^m, & \text{if } \alpha_{n-1} < \alpha \leq 1, \\ A^c + A^m + A^{r,n-1} + A^{r,n-2} + \ldots + A^{r,i+1}, & \text{if } \alpha_i < \alpha \leq \alpha_{i+1} \\ \qquad \text{(for some } i \in \{1, 2, \ldots, n-2\}\text{)}, & \\ A^c + A^m + A^{r,n-1} + A^{r,n-2} + \ldots + A^{r,1}, & \text{if } 0 \leq \alpha \leq \alpha_1. \end{cases} \tag{3.9}$$

The FN $\mathcal{A}$ is the unique finite FN whose center is A^c, whose central spread is A^m, whose left spreads are $\{A^{\ell,i}\}_{i=1}^{n-1}$, whose right spreads are $\{A^{r,i}\}_{i=1}^{n-1}$ and whose image is included in $\{\alpha_i\}_{i=0}^{n}$.

Proof: If we consider the funstions $\underline{a}, \overline{a} : \mathbb{I} \to \mathbb{R}$ given by equations (3.8) and (3.9), and taking into account that $A^{\ell,i}, A^{r,i} \geq 0$ for all $i \in \{1, 2, \ldots, n-1\}$, we deduce that $\underline{a}$ is non-decreasing, $\overline{a}$ is non-increasing and both of them are left continuous functions. Furthermore, if $\alpha, \beta \in \mathbb{I}$ are such that $\alpha \leq \beta$, then

$$\underline{a}_\alpha \leq \underline{a}_\beta \leq A^c - A^m \leq A^c \leq A^c + A^m \leq \overline{a}_\beta \leq \overline{a}_\alpha.$$

Lemma 2.2.5 guarantees that there is a unique FN (let us call it $\mathcal{B}$) whose level sets are determined by the functions $\underline{b} = \underline{a}$ and $\overline{b} = \overline{a}$. As the functions $\underline{a}$ and $\overline{a}$ are finite, Theorem 3.2.2 guarantees that such FN is also finite. Since $\ker \mathcal{B} = [\underline{a}_1, \overline{a}_1] = [A^c - A^m, A^c + A^m]$, then the center of $\mathcal{A}$ is A^c and its central spread is A^m. Furthermore, for all $i \in \{1, 2, \ldots, n-1\}$,

$$\begin{aligned} B^{\ell,i} &= \underline{b}_{\alpha_{i+1}} - \underline{b}_{\alpha_i} = \underline{a}_{\alpha_{i+1}} - \underline{a}_{\alpha_i} \\ &= \left(A^c - A^m - A^{\ell,n-1} - A^{\ell,n-2} - \ldots - A^{\ell,i+1}\right) \\ &\quad - \left(A^c - A^m - A^{\ell,n-1} - A^{\ell,n-2} - \ldots - A^{\ell,i+1} - A^{\ell,i}\right) \\ &= A^{\ell,i} \end{aligned}$$

and, similarly, $B^{r,i} = A^{r,i}$. Finally, we check that the image of $\mathcal{B}$ is included in $\Lambda = \{\alpha_i\}_{i=0}^n$. To prove that $\mathcal{B}(\mathbb{R}) \subseteq \Lambda$, we reason by contradiction. Suppose that there is some $t \in \mathbb{R}$ such that $\alpha = \mathcal{B}(t)$ satisfies $\alpha \in \mathbb{I} \diagdown \Lambda$. As $0, 1 \in \Lambda$, then $0 < \alpha < 1$. Hence, there is a unique $i \in \{1, 2, \ldots, n-1\}$ such that $\alpha_{i_0} < \alpha < \alpha_{i_0+1}$. By (3.8)-(3.9),

$$\mathcal{B}_\alpha = [\,\underline{a}_\alpha\,,\,\overline{a}_\alpha\,] = \left[\,\underline{a}_{\alpha_{i_0+1}}\,,\,\overline{a}_{\alpha_{i_0+1}}\right] = \mathcal{B}_{\alpha_{i_0+1}},$$

so

$$\{\, x \in \mathbb{R} : \mathcal{B}(x) \geq \alpha \,\} = \mathcal{B}_\alpha = \mathcal{B}_{\alpha_{i_0+1}} = \{\, x \in \mathbb{R} : \mathcal{B}(x) \geq \alpha_{i_0+1} \,\}.$$

As a consequence, since $\mathcal{B}(t) = \alpha$, then $t \in \mathcal{B}_\alpha = \mathcal{B}_{\alpha_{i_0+1}}$, which means that $\alpha = \mathcal{B}(t) \geq \alpha_{i_0+1}$. This contradicts the fact that $\alpha < \alpha_{i_0+1}$. Therefore, it is impossible to find a point $\alpha \in \mathcal{B}(\mathbb{R}) \diagdown \Lambda$, which concludes that $\mathcal{B}(\mathbb{R}) \subseteq \Lambda$.

This proves that the FN $\mathcal{B}$ given by functions (3.8)-(3.9) is the unique finite FN whose center is A^c, whose central spread is A^m, whose left spreads are $\{A^{\ell,i}\}_{i=1}^{n-1}$, whose right spreads are $\{A^{r,i}\}_{i=1}^{n-1}$ and whose image is included in $\{\alpha_i\}_{i=0}^n$. ■

Following the previous characterization, a finite FN can be equivalently determined in the way

$$\mathcal{A} = \mathrm{FN}\left(\{\alpha_i\}_{i=0}^n, A^c, A^m, \left\{A^{\ell,i}\right\}_{i=1}^{n-1}, \left\{A^{r,i}\right\}_{i=1}^{n-1}\right),$$

which has the advantage that all spreads are nonnegative, that is,

$$A^c \in \mathbb{R} \quad \text{but} \quad A^m, A^{\ell,i}, A^{r,i} \geq 0.$$

If we set

$$\Lambda = \{\alpha_0 = 0, \alpha_1, \ldots, \alpha_{n-1}, \alpha_n = 1\} \subset \mathbb{I},$$

then we will use the notation

$$\mathcal{A} = \mathrm{FFN}_\Lambda\left(A^c, A^m, \left\{A^{\ell,i}\right\}_{i=1}^{n-1}, \left\{A^{r,i}\right\}_{i=1}^{n-1}\right) \tag{3.10}$$

to describe any FN $\mathcal{F}$ in $\mathcal{F}_\Lambda$, where A^c is its center, and A^m, $\left\{A^{\ell,i}\right\}_{i=1}^{n-1}$ and $\left\{A^{r,i}\right\}_{i=1}^{n-1}$ are its (nonnegative) spreads.

3.3 Finite fuzzy random variables

In order to describe a regression procedure involving fuzzy variables that only take finite FNs, the following notion is necessary.

Definition 3.3.1 *Given a finite set $\Lambda \subset \mathbb{I}$, we will say that a fuzzy random variable $\mathcal{X} : \Omega \to \mathcal{F}$ is a* Λ-fuzzy random variable *if its image is included in $\mathcal{F}_\Lambda$. For short, we will say that $\mathcal{X}$ is a* Λ-FRV*, and we will denote it by $\mathcal{X} : \Omega \to \mathcal{F}_\Lambda$.*

Using the canonical representation (3.10), any Λ-FRV $\mathcal{X}$ can be written as

$$\mathcal{X} = \mathrm{FFN}_\Lambda \left(X^c, X^m, \left\{ X^{\ell,i} \right\}_{i=1}^{n-1}, \left\{ X^{r,i} \right\}_{i=1}^{n-1} \right),$$

where $X^c : \Omega \to \mathbb{R}$ is a real random variable and

$$X^m, X^{\ell,1}, X^{\ell,2}, \ldots, X^{\ell,n-1}, X^{r,1}, X^{r,2}, \ldots, X^{r,n-1} : \Omega \to [0, \infty)$$

are nonnegative real random variables. Hence, $\mathcal{X}$ can be equivalently seen as the real random vector

$$(X^c, X^m, X^{\ell,1}, X^{\ell,2}, \ldots, X^{\ell,n-1}, X^{r,1}, X^{r,2}, \ldots, X^{r,n-1}),$$

whose first component is valued in $\mathbb{R}$ but the other ones only take nonnegative values.

Sometimes it is interesting to summarize the information of a fuzzy random variable into a unique value that allows us to obtain its expected value to better understand of its behavior, to make comparisons, etc.

Definition 3.3.2 *The* expected value *of a Λ-FRV $\mathcal{X}$ is the unique finite FN* $\mathrm{E}[\mathcal{X}]$ *in $\mathcal{F}_\Lambda$ whose representation is*

$$\mathrm{E}[\mathcal{X}] = \mathrm{FFN}_\Lambda \left(\mathrm{E}[X^c], \mathrm{E}[X^m], \left\{ \mathrm{E}[X^{\ell,i}] \right\}_{i=1}^{n-1}, \left\{ \mathrm{E}[X^{r,i}] \right\}_{i=1}^{n-1} \right).$$

3.4 Density in $\mathcal{F}$ of finite fuzzy numbers

The following result shows that the set of all finite FNs is dense in the set of all FNs.

Theorem 3.4.1 Given a FN $\mathcal{A} \in \mathcal{F}$ and $\varepsilon > 0$, there is a finite subset $\Lambda \subset \mathbb{I}$ and a finite FN $\mathcal{A}' \in \mathcal{F}_\Lambda$ such that

$$\sup\left(\{\,|\,\mathcal{A}(x) - \mathcal{A}'(x)\,| : x \in \mathbb{R}\,\}\right) \leq \varepsilon. \tag{3.11}$$

Proof: Let consider a point $x_0 \in \ker \mathcal{A}$ and the funcions $F_1, F_2 : \mathbb{R} \to \mathbb{I}$ given, for all $x \in \mathbb{R}$, by:

$$F_1(x) = \begin{cases} \mathcal{A}(x), & \text{if } x \leq x_0, \\ 1, & \text{if } x > x_0; \end{cases}$$

$$F_2(x) = \begin{cases} 0, & \text{if } x \leq x_0, \\ 1 - \mathcal{A}(x), & \text{if } x > x_0. \end{cases}$$

In [44], the authors proved that F_1 and F_2 are distribution functions representing some unique real random variables X_1 and X_2. In particular, $X_1 \leq x_0$ and $X_2 \geq x_0$. Furthermore, as the support of $\mathcal{A}$ is bounded, then X_1 and X_2 are also bounded. As the set of all finite random variables is dense in the set of all bounded, real random variables, there are finite, bounded, real random variables X_1' and X_2' such that

$$\sup\left(\{\,|\,X_1(x) - X_1'(x)\,| : x \in \mathbb{R}\,\}\right) \leq \varepsilon,$$
$$\sup\left(\{\,|\,X_2(x) - X_2'(x)\,| : x \in \mathbb{R}\,\}\right) \leq \varepsilon,$$
$$X_1' \leq x_0 \qquad \text{and} \qquad X_2' \geq x_0.$$

If we define

$$\mathcal{A}'(x) = \begin{cases} X_1'(x), & \text{if } x \leq x_0, \\ 1 - X_2'(x), & \text{if } x > x_0, \end{cases}$$

then $\mathcal{A}'$ is a finite FN such that (3.11) holds. ∎

CHAPTER 4

The fuzzy regression procedure based on finite fuzzy numbers

In this chapter we introduce a fuzzy regression methodology considering that the response variable is a Λ-FRV. This procedure is especially interesting when the explained variable can only take a finite amount of values, or when the response variable is approximated in such way that we can only access to a finite number of its α-cuts. The first part of this chapter is dedicated to introduce the tools that are involved in many regression procedures

4.1 Introduction to fuzzy distance measures between finite fuzzy numbers

One of the main difficulties that we found when we handle FNs is the fact that the set $\mathcal{F}$ is not canonically endowed with a partial order that could extend the total order of real numbers. Lots of partial orders on $\mathcal{F}$ can be considered but none of them is coherent with human intuition. As a consequence, it is not easy to generalize the notion of *metric* to the fuzzy setting because the *triangle* *inequality* has not a clear interpretation without a consistent partial order.

In the first part of this chapter we explore the necessity of having a genuine fuzzy distance in order to measure how similar of distinct are two FNs. Then, we present what could be, from our point of view, the most coherent definition of fuzzy distance in the set of all FNs. After that, we introduce a family of fuzzy distance measures which is especially interesting when working with finite FNs and that we will employ to describe a fuzzy regression methodology in the next chapter.

4.1.1 About the necessity of handling genuine fuzzy distance measures

Due to its possible applications in several areas, the notion of *metric* plays a key role in many fields of study that interpret the distance measure between two points as the difference between them (for instance, it is basic for carrying out a regression process). Traditionally, the distance between two points has always been a real number. Even in a fuzzy context, the distance between two FNs is usually interpreted as a crisp number considering, for instance, the area between them (see [14, 46]). However, this conception does not capture both imprecision and uncertainty and, consequently, is not consistent with factors such as vagueness and ambiguity which affect the behavior of the phenomenon studied in the fuzzy setting. Therefore, it would be more reasonable to use a FN rather than a real number to measure the distance between two FNs.

4.1.2 Real-valued metrics on $\mathcal{F}$

Let us recall the notion of metric in an arbitrary nonempty set X. Following [49], a *metric on a set* X is a mapping $d : X \times X \to \mathbb{R}_0^+$ verifying, for all

$x, y, z \in X$,

(i) $d(x,x) = 0$;

(ii) if $d(x,y) = 0$, then $x = y$;

(iii) *(symmetry)* $d(x,y) = d(y,x)$;

(iv) *(triangle inequality)* $d(x,z) \leq d(x,y) + d(y,z)$.

A mapping $d : X \times X \to \mathbb{R}_0^+$ is a *pseudometric* (respectively, a *semimetric*; *pseudosemimetric*) *on* X if it satisfies (i), (iii) and (iv) (respectively, (i), (ii) and (iii); (i) and (iii)).

Having in mind the previous definition, it is not difficult to consider metrics (or semimetrics) on $\mathcal{F}$ as follows. Given real numbers $p \in (0, \infty)$ and $q \in (0,1)$, let $\delta_{p,q} : \mathcal{F} \times \mathcal{F} \to \mathbb{R}_0^+$ be the function given, for all $\mathcal{A}, \mathcal{B} \in \mathcal{F}$, by:

$$\delta_{p,q}(\mathcal{A}, \mathcal{B}) = \sqrt[p]{(1-q) \int_0^1 |\,\underline{a}_\alpha - \underline{b}_\alpha\,|^p \, d\alpha + q \int_0^1 \left|\,\overline{a}_\alpha - \overline{b}_\alpha\,\right|^p d\alpha}\,.$$

Then $\delta_{p,q}$ is a semimetric and, in some cases, it is a metric on $\mathcal{F}$. It is of interests the particular case $p = 2$ in which:

$$\delta_{2,q}(\mathcal{A}, \mathcal{B}) = \sqrt{(1-q) \int_0^1 (\underline{a}_\alpha - \underline{b}_\alpha)^2 \, d\alpha + q \int_0^1 \left(\overline{a}_\alpha + \overline{b}_\alpha\right)^2 d\alpha}\,.$$

If $p = \infty$, we can also consider:

$$\delta_{\infty,q}(\mathcal{A}, \mathcal{B}) = (1-q) \sup_{0<\alpha\leq 1} |\,\underline{a}_\alpha - \underline{b}_\alpha\,| + q \sup_{0<\alpha\leq 1} \left|\,\overline{a}_\alpha - \overline{b}_\alpha\,\right|.$$

As we have just commented, from our point of view, crisp distances are not coherent in this framework because, intrinsically, they do not capture the most important characteristic of FNs: their ability to model situations in which uncertainty appears. If FNs represent vague and unknown quantities, how it is possible to measure the distance between two of them by a unique, precise real number? It seems more reasonable to use a FN in order to measure the distance between two FNs, that is, to consider a mapping

$$D : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$$

acting as a metric.

Nevertheless, we realize that, in order to generalize the triangle inequality, the set $\mathcal{F}$ of all FNs on $\mathbb{R}$ carries a drawback of importance: it is not canonically ordered. The usual order between real-valued functions (that is, $f \leq g$ if $f(x) \leq g(x)$ for all $x \in \mathrm{dom}(f) = \mathrm{dom}(g)$) is not consistent with human interpretation of FNs. There exist several notions of possible partial orders on $\mathcal{F}$ but none of them is widely accepted. This obstacle leads to a subsequent consequence: the family $\mathcal{F}$ is not endowed with a universally accepted metric by all researchers.

In this context, it is more reasonable to consider *distance measures* rather than *metrics* because, usually, we must measure how similar or different are two FNs. Thus, it is of interest to extend the notion of metric to a more general context.

4.1.3 Distance measures on an arbitrary set endowed with a binary relation

In the next definition we establish what properties we can require to a mapping acting as a distance measure on a partially ordered set when it is valued on the own set, that is, we generalize the axioms that a metric must satisfy to the general setting of sets endowed with binary relations.

Definition 4.1.1 *Let S be a nonempty set endowed with a binary relation $\sqsubseteq$ and let $0_S \in S$ be a selected point such that $0_S \sqsubseteq 0_S$. Consider the set*

$$S^+_{0,\sqsubseteq} = \{\, x \in S : 0_S \sqsubseteq x \,\}$$

and let

$$s : S^+_{0,\sqsubseteq} \times S^+_{0,\sqsubseteq} \to S^+_{0,\sqsubseteq}$$

be a mapping. A distance function on $(S, 0_S, \sqsubseteq, s)$ *(or a* metric*) is a mapping*

$$d : S \times S \to S^+_{0,\sqsubseteq}$$

verifying, for all $x, y, z \in S$,

(i) $d(x,x) = 0_S$;

(ii) *if* $d(x,y) = 0_S$, *then* $x = y$;

(iii) (symmetry) $d(x,y) = d(y,x)$;

(iv) (triangle inequality) $d(x,z) \sqsubseteq s(d(x,y), d(y,z))$.

We also say that $(S, 0_S, s)$ *is a* metric space w.r.t. the partial order $\sqsubseteq$. *The function* d *is:*

- *a* pseudometric *if it satisfies (i), (iii) and (iv);*
- *a* semimetric *(*on $(S, 0_S)$*) if it satisfies (i), (ii) and (iii);*
- *a* pseudosemimetric *(*on $(S, 0_S)$*) if it satisfies (i) and (iii).*

4.2 Fuzzy distance measures on $\mathcal{F}$

Taking into account the previous definition, we can consider the following notion as a genuine fuzzy metric on $\mathcal{F}$. Definitively, it has sense to use the crisp FN $\bar{0}$ as the selected point 0_S and the sum of FNs as mapping $s : S^+_{0,\sqsubseteq} \times S^+_{0,\sqsubseteq} \to S^+_{0,\sqsubseteq}$. However, it would be depend on a binary relation.

Definition 4.2.1 *Let* $\preccurlyeq$ *be a binary relation on* $\mathcal{F}$ *such that* $\bar{0} \preccurlyeq \bar{0}$ *and*

$$\bar{0} \preccurlyeq \mathcal{A} + \mathcal{B} \qquad \text{for all } \mathcal{A}, \mathcal{B} \in \mathcal{F} \text{ such that } \bar{0} \preccurlyeq \mathcal{A} \text{ and } \bar{0} \preccurlyeq \mathcal{B}.$$

Let $\mathcal{F}^+_0 = \{\mathcal{A} \in \mathcal{F} : \bar{0} \preccurlyeq \mathcal{A}\}$. *A* distance function *(or a* metric*)* on $\mathcal{F}$ *is a mapping*

$$D : \mathcal{F} \times \mathcal{F} \to \mathcal{F}^+_0$$

verifying, for all $\mathcal{A}, \mathcal{B}, \mathcal{C} \in \mathcal{F}$,

(i) $D(\mathcal{A}, \mathcal{A}) = \bar{0}$;

(ii) *if* $D(\mathcal{A}, \mathcal{B}) = \bar{0}$, *then* $\mathcal{A} = \mathcal{B}$;

(iii) (symmetry) $D(\mathcal{B}, \mathcal{A}) = D(\mathcal{A}, \mathcal{B})$;

(iv) (triangle inequality) $D(\mathcal{A}, \mathcal{C}) \preccurlyeq D(\mathcal{A}, \mathcal{B}) + D(\mathcal{B}, \mathcal{C})$.

We also say that $(\mathcal{F}, D)$ *is a* fuzzy metric space w.r.t. the partial order $\preccurlyeq$. *The function* D *is:*

- *a* fuzzy pseudometric *if it satisfies (i), (iii) and (iv);*
- *a* fuzzy semimetric *if it satisfies (i), (ii) and (iii);*
- *a* fuzzy pseudosemimetric *if it satisfies (i) and (iii).*

In [45], the authors presented a first approach to the problem of introducing, at the same time,

▶ binary relations on $\mathcal{F}$ that are partial orders, at least, on the subclass of trapezoidal FNs;

▶ fuzzy distance measures on $\mathcal{F}$ that are true metrics, at least, on the subclass of trapezoidal FNs.

Their families depended on a large list of some geometric characteristics of FNs. Nevertheless, for technical reasons, although such families are appropriate to handle a large list of FNs, they are not the best possibilities for working with finite FNs. A slight change, that generalize the family into an even more general one, must be done. To explain such modification, we recall here the key definitions in such paper.

A *partition* of the interval $\mathbb{I}$ is a finite set $\mathcal{P} = \{\delta_0, \delta_1, \ldots, \delta_n\}$ such that $0 = \delta_0 < \delta_1 < \ldots < \delta_n = 1$. If $\mathcal{P} = \{\delta_i\}_{i=0}^n$ is a partition of $\mathbb{I}$ and $f : S \to \mathbb{R}$ is a mapping defined on $S \supseteq \mathcal{P}$, we will denote by Δf_i, for all $i \in \{1, 2, \ldots, n\}$, to the average rate of change of f on the interval $[\delta_{i-1}, \delta_i]$, that is,

$$\Delta f_i = \Delta f_{[\delta_{i-1}, \delta_i]} = \frac{f(\delta_i) - f(\delta_{i-1})}{\delta_i - \delta_{i-1}}.$$

Let $\mathcal{D} : \mathcal{F} \to \mathbb{R}$ be a defuzzification and let $\mathcal{P} = \{\delta_i\}_{i=0}^n$ be a partition of $\mathbb{I}$.

Definition 4.2.2 *(Roldán-López de Hierro* et al. *[45], Definition 2) Given two FNs $\mathcal{A}, \mathcal{B} \in \mathcal{F}$, we will write $\mathcal{A} \preccurlyeq \mathcal{B}$* w.r.t. *$(\mathcal{D}, \mathcal{P})$ if the following conditions are fulfilled:*

$$\begin{cases} \mathcal{D}\mathcal{A} \leq \mathcal{D}\mathcal{B}, \\ \operatorname{spr}\mathcal{A} \leq \operatorname{spr}\mathcal{B}, \\ \left.\begin{array}{l} \Delta \underline{a}_i \leq \Delta \underline{b}_i \\ \Delta \overline{a}_i \geq \Delta \overline{b}_i \end{array}\right\} \quad \textit{for all } i \in \{1, 2, \ldots, n\}. \end{cases}$$

Definition 4.2.3 *(Roldán-López de Hierro* et al. *[45], Definition 5) Let $q, q_0 \geq 0$, let $q_1, q_2 : \mathbb{I} \to [0, \infty[$ be two left continuous, non-increasing mappings on $\mathbb{I}$ and let $\phi_0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_0^+$, $\psi : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$, $\{\phi_i : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+\}_{i=1}^n$ and $\{\varphi_i : \mathbb{R}_0^- \times \mathbb{R}_0^- \to \mathbb{R}_0^+\}_{i=1}^n$ be pseudosemimetrics on their respective domains. For all FNs $\mathcal{A}, \mathcal{B} \in \mathcal{F}$ and all $\alpha \in \mathbb{I}$, define:*

$$\begin{aligned} \underline{D(\mathcal{A}, \mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) \\ &\qquad - q_1(\alpha) \sum_{i=1}^n \phi_i\left(\Delta \underline{a}_i, \Delta \underline{b}_i\right), \qquad (4.1) \\ \overline{D(\mathcal{A}, \mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) + q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) \\ &\qquad + q_2(\alpha) \sum_{i=1}^n \varphi_i\left(\Delta \overline{a}_i, \Delta \overline{b}_i\right). \qquad (4.2) \end{aligned}$$

Let $D(\mathcal{A}, \mathcal{B})$ be the only FN determined by its α-cuts (4.1)-(4.2).

In some sense, functions q_1 and q_2 control the class of FN that contains the FN $D(\mathcal{A}, \mathcal{B})$ (rectangular, triangular, trapezoidal, etc.) It is easy to show that if q_1 and q_2 are finite, then $D(\mathcal{A}, \mathcal{B})$ is also a finite FN whatever $\mathcal{A}$ and $\mathcal{B}$. However, under general finite functions, we have not a control about the resulting left and right spreads of the obtained finite FN. Thus, we need to refine the previous definition.

4.3 A partial order on $\mathcal{F}_\Lambda$

In this section we study the binary relation $\mathcal{A} \preccurlyeq \mathcal{B}$ *w.r.t.* $(\mathcal{D}, \Lambda)$ given in Definition 4.2.2 when $\mathcal{A}$ and $\mathcal{B}$ are finite FNs in $\mathcal{F}_\Lambda$ (notice that the partition $\mathcal{P}$ coincides with the finite set Λ). For practical reasons, we are only interested in the case in which the defuzzification $\mathcal{D}$ is the most important one, that is, $\mathcal{D}_c$. We prove that, when this binary relation is particularized to $\mathcal{F}_\Lambda$, then it becomes a partial order.

Lemma 4.3.1 If we consider the defuzzification $\mathcal{D}_c$ and the partition $\Lambda = \{ \alpha_0 = 0 < \alpha_1 < \ldots < \alpha_{n-1} < \alpha_n = 1 \}$, then two finite FNs in $\mathcal{F}_\Lambda$,

$$\mathcal{A} = \mathrm{FFN}_\Lambda \left(A^c, A^m, \left\{ A^{\ell,i} \right\}_{i=1}^{n-1}, \left\{ A^{r,i} \right\}_{i=1}^{n-1} \right) \quad \text{and}$$
$$\mathcal{B} = \mathrm{FFN}_\Lambda \left(B^c, B^m, \left\{ B^{\ell,i} \right\}_{i=1}^{n-1}, \left\{ B^{r,i} \right\}_{i=1}^{n-1} \right), \tag{4.3}$$

satisfy $\mathcal{A} \preccurlyeq \mathcal{B}$ w.r.t. $(\mathcal{D}_c, \Lambda)$ if and only if

$$A^c \leq B^c, \quad A^m \leq B^m, \quad A^{\ell,i-1} \leq B^{\ell,i-1} \quad \text{and}$$
$$A^{r,i-1} \leq B^{r,i-1} \ \text{ for all } i \in \{2, 3, \ldots, n-1\}. \tag{4.4}$$

Furthermore, the binary relation $\preccurlyeq$ w.r.t. $(\mathcal{D}_c, \Lambda)$ is a partial order on $\mathcal{F}_\Lambda$.

Proof: If

$$\mathcal{A} = \mathrm{FFN}_\Lambda \left(A^c, A^m, \left\{ A^{\ell,i} \right\}_{i=1}^{n-1}, \left\{ A^{r,i} \right\}_{i=1}^{n-1} \right) \quad \text{and}$$
$$\mathcal{B} = \mathrm{FFN}_\Lambda \left(B^c, B^m, \left\{ B^{\ell,i} \right\}_{i=1}^{n-1}, \left\{ B^{r,i} \right\}_{i=1}^{n-1} \right)$$

are two finite FNs in $\mathcal{F}_\Lambda$, given by using the representation (3.10), we deduce from (3.6)-(3.7) that, for all $i \in \{2, 3, \ldots, n\}$,

$$\Delta\underline{a}_i = \Delta\underline{a}_{[\alpha_{i-1},\alpha_i]} = \frac{\underline{a}(\alpha_i) - \underline{a}(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} = \frac{\underline{a}_{\alpha_i} - \underline{a}_{\alpha_{i-1}}}{\alpha_i - \alpha_{i-1}} = \frac{A^{\ell,i-1}}{\alpha_i - \alpha_{i-1}}, \tag{4.5}$$

$$\Delta\overline{a}_i = \Delta\overline{a}_{[\alpha_{i-1},\alpha_i]} = \frac{\overline{a}(\alpha_i) - \overline{a}(\alpha_{i-1})}{\alpha_i - \alpha_{i-1}} = \frac{\overline{a}_{\alpha_i} - \overline{a}_{\alpha_{i-1}}}{\alpha_i - \alpha_{i-1}} = \frac{-A^{r,i-1}}{\alpha_i - \alpha_{i-1}}. \tag{4.6}$$

It follows from (3.8) and (3.9) that

$$\underline{a}(\alpha_1) = \underline{a}(\alpha_0) = A^c - A^m - A^{\ell,n-1} - A^{\ell,n-2} - \ldots - A^{\ell,1} \quad \text{and}$$
$$\overline{a}(\alpha_1) = \overline{a}(\alpha_0) = A^c + A^m + A^{r,n-1} + A^{r,n-2} + \ldots + A^{r,1},$$

so $\Delta\underline{a}_1 = \Delta\overline{a}_1 = \Delta\underline{b}_1 = \Delta\overline{b}_1 = 0$. Therefore, for all $i \in \{2, 3, \ldots, n\}$,

$$\Delta\underline{a}_i \leq \Delta\underline{b}_i \quad \Leftrightarrow \quad \frac{A^{\ell,i-1}}{\alpha_i - \alpha_{i-1}} \leq \frac{B^{\ell,i-1}}{\alpha_i - \alpha_{i-1}} \quad \Leftrightarrow \quad A^{\ell,i-1} \leq B^{\ell,i-1},$$
$$\Delta\overline{a}_i \geq \Delta\overline{b}_i \quad \Leftrightarrow \quad \frac{-A^{r,i-1}}{\alpha_i - \alpha_{i-1}} \geq \frac{-B^{r,i-1}}{\alpha_i - \alpha_{i-1}} \quad \Leftrightarrow \quad A^{r,i-1} \leq B^{r,i-1}.$$

This means that, following Definition 4.2.2, $\mathcal{A} \preccurlyeq \mathcal{B}$ w.r.t. $(\mathcal{D}_c, \Lambda)$ if and only if the following conditions hold:

- $\mathcal{D}_c\mathcal{A} \leq \mathcal{D}_c\mathcal{B} \quad \Leftrightarrow \quad A^c \leq B^c$;
- $\operatorname{spr}\mathcal{A} \leq \operatorname{spr}\mathcal{B} \quad \Leftrightarrow \quad A^m \leq B^m$;
- $\Delta\underline{a}_i \leq \Delta\underline{b}_i$ for all $i \in \{1, 2, \ldots, n\}$

  $\Leftrightarrow \quad A^{\ell,i-1} \leq B^{\ell,i-1}$ for all $i \in \{2, 3, \ldots, n\}$;
- $\Delta\overline{a}_i \geq \Delta\overline{b}_i$ for all $i \in \{1, 2, \ldots, n\}$

  $\Leftrightarrow \quad A^{r,i-1} \leq B^{r,i-1}$ for all $i \in \{2, 3, \ldots, n\}$.

Furthermore, the binary relation $\preccurlyeq$ w.r.t. $(\mathcal{D}_c, \Lambda)$ is reflexive, transitive and antisymmetric, so it is a partial order on $\mathcal{F}_\Lambda$. This completes the proof. ∎

4.4 A family of fuzzy distance measures on $\mathcal{F}_\Lambda$

In this section we take advantage of the introduced partial order $\preccurlyeq$ w.r.t. $(\mathcal{D}_c, \Lambda)$ to endow the set $\mathcal{F}_\Lambda$ with a family of distinct fuzzy distance measures that will be useful to describe a fuzzy regression methodology involving finite fuzzy random variables. In the following result, we slightly modify the family of distance measures given in Definition 4.2.3 by replacing the functions q_1 and q_2 by respective families of functions $\{p_i : \mathbb{I} \to \mathbb{R}_0^+\}_{i=1}^n$ and $\{h_i : \mathbb{I} \to \mathbb{R}_0^+\}_{i=1}^n$ satisfying the same properties than q_1 and q_2.

Theorem 4.4.1 Let $\mathcal{D} : \mathcal{F} \to \mathbb{R}$ be a defuzzification, let $\Lambda = \{\alpha_0 = 0 < \alpha_1 < \ldots < \alpha_{n-1} < \alpha_n = 1\}$ be a partition of $\mathbb{I}$, let $q, q_0 \geq 0$, let $\{p_i, h_i : \mathbb{I} \to \mathbb{R}_0^+\}_{i=1}^n$ be a family of left continuous, non-increasing mappings on $\mathbb{I}$, and let $\phi_0 : \mathbb{R} \times \mathbb{R} \to \mathbb{R}_0^+$, $\psi : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+$, $\{\phi_i : \mathbb{R}_0^+ \times \mathbb{R}_0^+ \to \mathbb{R}_0^+\}_{i=1}^n$ and $\{\varphi_i : \mathbb{R}_0^- \times \mathbb{R}_0^- \to \mathbb{R}_0^+\}_{i=1}^n$ be pseudosemimetrics on their respective domains. For all FNs $\mathcal{A}, \mathcal{B} \in \mathcal{F}$ and all $\alpha \in \mathbb{I}$, define:

$$\begin{aligned}\underline{D(\mathcal{A},\mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) \\ &\qquad - \sum_{i=1}^n p_i(\alpha)\, \phi_i\left(\Delta\underline{a}_i, \Delta\underline{b}_i\right), \qquad (4.7)\\ \overline{D(\mathcal{A},\mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) + q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) \\ &\qquad + \sum_{i=1}^n h_i(\alpha)\, \varphi_i\left(\Delta\overline{a}_i, \Delta\overline{b}_i\right). \qquad (4.8)\end{aligned}$$

Then there exists a unique FN $D(\mathcal{A}, \mathcal{B})$ determined by its α-cuts (4.7)-(4.8).

PROOF: It follows from the fact that the function $\alpha \mapsto \underline{D(\mathcal{A},\mathcal{B})}_\alpha$ is left continuous and nondecreasing on $\mathbb{I}$, and the function $\alpha \mapsto \overline{D(\mathcal{A},\mathcal{B})}_\alpha$ is left continuous and non-increasing on $\mathbb{I}$. Since $\underline{D(\mathcal{A},\mathcal{B})}_\alpha \leq q_0\phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) \leq \overline{D(\mathcal{A},\mathcal{B})}_\alpha$ for all $\alpha \in \mathbb{I}$, then there exists a unique FN $D(\mathcal{A}, \mathcal{B})$ determined by the given α-cuts (see Lemma 2.2.5). ■

Observación 4.4.2 *Clearly, if $p_i(\alpha) = q_1(\alpha)$ and $h_i(\alpha) = q_2(\alpha)$ for all $i \in \{1, 2, \ldots, n\}$ and all $\alpha \in \mathbb{I}$, then equations (4.7)-(4.8) generalize equations (4.1)-(4.2). As a result, similarity measures obtained by Theorem 4.4.1 are more general than obtained by Definition 4.2.3.*

We are interested in the following particular case of the previous theorem. Given $\beta \in \mathbb{I}$, let denote by $q_\beta : \mathbb{I} \to \mathbb{I}$ the function given, for all $\alpha \in \mathbb{I}$, by

$$q_\beta(\alpha) = \begin{cases} 1, & \text{if } \alpha \leq \beta, \\ 0, & \text{if } \beta > \alpha. \end{cases}$$

Clearly, each function q_β is non-increasing and left-continuous on $\mathbb{I}$.

Theorem 4.4.3 Assume that, in Theorem 4.4.1, we choose $\mathcal{D} = \mathcal{D}_c$, the partition $\Lambda = \{\alpha_0 = 0 < \alpha_1 < \ldots < \alpha_{n-1} < \alpha_n = 1\}$, $q = q_0 = 1$, $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha)$ for all $i \in \{2, 3, \ldots, n\}$,

$$\phi_0(x, y) = \psi(x, y) = (x - y)^2 \quad \text{and} \quad \phi_i(x, y) = \varphi_i(x, y) = (\alpha_i - \alpha_{i-1})^2 (x - y)^2$$

in their respective domains. Then, for all finite FNs $\mathcal{A}, \mathcal{B} \in \mathcal{F}_\Lambda$ given as in (4.3):

$$D(\mathcal{A}, \mathcal{B}) = \text{FFN}_\Lambda \Bigg((A^c - B^c)^2, \ (A^m - B^m)^2, \\ \left\{ \left(A^{\ell,i} - B^{\ell,i}\right)^2 \right\}_{i=1}^{n-1}, \ \left\{ \left(A^{r,i} - B^{r,i}\right)^2 \right\}_{i=1}^{n-1} \Bigg). \tag{4.9}$$

In particular, $D(\mathcal{A}, \mathcal{B}) \in \mathcal{F}_\Lambda$.

Notice that in the previous statement we have not considered neither p_1 nor h_1 because, as we have shown in the proof of Lemma 4.3.1, in the case of finite FNs of $\mathcal{F}_\Lambda$, we have that $\Delta\underline{a}_1 = \Delta\overline{a}_1 = \Delta\underline{b}_1 = \Delta\overline{b}_1 = 0$.

PROOF: Let $\mathcal{C} = D(\mathcal{A}, \mathcal{B}) \in \mathcal{F}$. First of all, notice that

$$q_0\, \phi_0(\mathcal{D}_c\mathcal{A}, \mathcal{D}_c\mathcal{B}) = (\mathcal{D}_c\mathcal{A} - \mathcal{D}_c\mathcal{B})^2 = (A^c - B^c)^2 \quad \text{and}$$
$$q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) = (\operatorname{spr}\mathcal{A} - \operatorname{spr}\mathcal{B})^2 = (A^m - B^m)^2.$$

On the other hand, by (4.5) and (4.6), for all $i \in \{2, 3, \ldots, n\}$,

$$\phi_i\left(\Delta\underline{a}_i, \Delta\underline{b}_i\right) = (\alpha_i - \alpha_{i-1})^2 \left(\frac{A^{\ell,i-1}}{\alpha_i - \alpha_{i-1}} - \frac{B^{\ell,i-1}}{\alpha_i - \alpha_{i-1}}\right)^2 = \left(A^{\ell,i} - B^{\ell,i}\right)^2,$$

$$\varphi_i\left(\Delta\overline{a}_i, \Delta\overline{b}_i\right) = (\alpha_i - \alpha_{i-1})^2 \left(\frac{-A^{r,i-1}}{\alpha_i - \alpha_{i-1}} - \frac{-B^{r,i-1}}{\alpha_i - \alpha_{i-1}}\right)^2 = \left(A^{r,i} - B^{r,i}\right)^2.$$

In order to prove that $D(\mathcal{A}, \mathcal{B}) \in \mathcal{F}_\Lambda$, we use Theorem 3.2.3. Indeed, if $\alpha \in (\alpha_{n-1}, \alpha_n] = (\alpha_{n-1}, 1]$, then $\alpha > \alpha_{i-1}$ for all $i \in \{2, 3, \ldots, n\}$, so $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha) = 0$. Then

$$\begin{aligned}\underline{D(\mathcal{A}, \mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{DA}, \mathcal{DB}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) - \sum_{i=2}^{n} q_{\alpha_{i-1}}(\alpha)\, \phi_i\left(\Delta\underline{a}_i, \Delta\underline{b}_i\right) \\ &= (A^c - B^c)^2 - (A^m - B^m)^2.\end{aligned}$$

If $\alpha \in (\alpha_{n-2}, \alpha_{n-1}]$, then $\alpha > \alpha_{i-1}$ for all $i \in \{2, 3, \ldots, n-1\}$, but $\alpha \leq \alpha_{n-1}$, so $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha) = 0$ for all $i \in \{2, 3, \ldots, n-1\}$ but $p_n(\alpha) = h_n(\alpha) = q_{\alpha_{n-1}}(\alpha) = 1$. Then

$$\begin{aligned}\underline{D(\mathcal{A}, \mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{DA}, \mathcal{DB}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) - \sum_{i=2}^{n} q_{\alpha_{i-1}}(\alpha)\, \phi_i\left(\Delta\underline{a}_i, \Delta\underline{b}_i\right) \\ &= (A^c - B^c)^2 - (A^m - B^m)^2 - q_{\alpha_{n-1}}(\alpha)\, \phi_n\left(\Delta\underline{a}_n, \Delta\underline{b}_n\right) \\ &= (A^c - B^c)^2 - (A^m - B^m)^2 - \left(A^{\ell,n-1} - B^{\ell,n-1}\right)^2.\end{aligned}$$

Similarly, if $\alpha \in (\alpha_j, \alpha_{j+1}]$, for some $j \in \{1, 2, \ldots, n-2\}$, then $\alpha > \alpha_{i-1}$ for all $i \in \{1, 2, \ldots, j+1\}$, but $\alpha \leq \alpha_{i-1}$ for all $i \in \{j+2, j+3, \ldots, n\}$. Therefore $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha) = 0$ for all $i \in \{2, 3, \ldots, j+1\}$ but $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha) = 1$ for all $i \in \{j+2, j+3, \ldots, n\}$. Then

$$\begin{aligned}\underline{D(\mathcal{A}, \mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{DA}, \mathcal{DB}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) - \sum_{i=2}^{n} q_{\alpha_{i-1}}(\alpha)\, \phi_i\left(\Delta\underline{a}_i, \Delta\underline{b}_i\right) \\ &= (A^c - B^c)^2 - (A^m - B^m)^2 - q_{\alpha_{n-1}}(\alpha)\, \phi_n\left(\Delta\underline{a}_n, \Delta\underline{b}_n\right) \\ &\qquad - q_{\alpha_{n-2}}(\alpha)\, \phi_{n-1}\left(\Delta\underline{a}_{n-1}, \Delta\underline{b}_{n-1}\right) - \ldots \\ &\qquad - q_{\alpha_{j+1}}(\alpha)\, \phi_{j+2}\left(\Delta\underline{a}_{j+2}, \Delta\underline{b}_{j+2}\right) \\ &= (A^c - B^c)^2 - (A^m - B^m)^2 - \left(A^{\ell,n-1} - B^{\ell,n-1}\right)^2 \\ &\qquad - \left(A^{\ell,n-2} - B^{\ell,n-2}\right)^2 - \ldots - \left(A^{\ell,j+1} - B^{\ell,j+1}\right)^2.\end{aligned}$$

Finally, if $\alpha \in [\alpha_0, \alpha_1] = [0, \alpha_1]$, then $\alpha \leq \alpha_{i-1}$ for all $i \in \{2, 3, \ldots, n-1\}$, so $p_i(\alpha) = h_i(\alpha) = q_{\alpha_{i-1}}(\alpha) = 1$ for all $i \in \{2, 3, \ldots, n\}$ and

$$\begin{aligned}
\underline{D(\mathcal{A},\mathcal{B})}_\alpha &= q_0\, \phi_0\left(\mathcal{D}\mathcal{A}, \mathcal{D}\mathcal{B}\right) - q\, \psi(\operatorname{spr}\mathcal{A}, \operatorname{spr}\mathcal{B}) - \sum_{i=2}^{n} q_{\alpha_{i-1}}(\alpha)\, \phi_i\left(\Delta \underline{a}_i, \Delta \underline{b}_i\right) \\
&= \left(A^c - B^c\right)^2 - \left(A^m - B^m\right)^2 - q_{\alpha_{n-1}}(\alpha)\, \phi_n\left(\Delta \underline{a}_n, \Delta \underline{b}_n\right) \\
&\qquad - q_{\alpha_{n-2}}(\alpha)\, \phi_{n-1}\left(\Delta \underline{a}_{n-1}, \Delta \underline{b}_{n-1}\right) - \ldots \\
&\qquad - q_{\alpha_1}(\alpha)\, \phi_2\left(\Delta \underline{a}_2, \Delta \underline{b}_2\right) \\
&= \left(A^c - B^c\right)^2 - \left(A^m - B^m\right)^2 - \left(A^{\ell,n-1} - B^{\ell,n-1}\right)^2 \\
&\qquad - \left(A^{\ell,n-2} - B^{\ell,n-2}\right)^2 - \ldots - \left(A^{\ell,1} - B^{\ell,1}\right)^2 .
\end{aligned}$$

The same reasoning is useful to work out the values of $\overline{D(\mathcal{A},\mathcal{B})}_\alpha$, which are given by

$$\overline{D(\mathcal{A},\mathcal{B})}_\alpha = \begin{cases}
\left(A^c - B^c\right)^2 + \left(A^m - B^m\right)^2, \\
\qquad \text{if } \alpha_{n-1} < \alpha \leq 1, \\
\left(A^c - B^c\right)^2 + \left(A^m - B^m\right)^2 + \sum_{i=j+1}^{n-1} \left(A^{r,i} - B^{r,i}\right)^2, \\
\qquad \text{if } \alpha_j < \alpha \leq \alpha_{j+1} \quad (\text{for some } j \in \{1, 2, \ldots, n-2\}), \\
\left(A^c - B^c\right)^2 + \left(A^m - B^m\right)^2 + \sum_{i=1}^{n-1} \left(A^{r,i} - B^{r,i}\right)^2, \\
\qquad \text{if } 0 \leq \alpha \leq \alpha_1.
\end{cases}$$

In particular, $\mathcal{C} = D\left(\mathcal{A}, \mathcal{B}\right)$ is the unique finite FN in $\mathcal{F}_\Lambda$ whose center is $C^c = \left(A^c - B^c\right)^2$, whose central spread is $C^m = \left(A^m - B^m\right)^2$, whose left spreads are

$$\left\{ \; C^{\ell,i} = \left(A^{\ell,i} - B^{\ell,i}\right)^2 \; \right\}_{i=1}^{n-1}$$

and whose right spreads are

$$\left\{ \; C^{r,i} = \left(A^{r,i} - B^{r,i}\right)^2 \; \right\}_{i=1}^{n-1} .$$

This proves that (4.9) holds. ∎

In particular, the restriction to $\mathcal{F}_\Lambda \times \mathcal{F}_\Lambda$ of the distance measure $D : \mathcal{F} \times \mathcal{F} \rightarrow \mathcal{F}$ provides us an inner distance measure, that is, it is a mapping $D : \mathcal{F}_\Lambda \times \mathcal{F}_\Lambda \rightarrow$

$\mathcal{F}_\Lambda$ that we will denote by

$$D_\Lambda : \mathcal{F}_\Lambda \times \mathcal{F}_\Lambda \to \mathcal{F}_\Lambda. \tag{4.10}$$

4.5 Fuzzy regression procedure

In this section we introduce a fuzzy regression methodology based on finite fuzzy numbers. To do this, we will use Theorem 4.4.3, which permit us to performance a *least square methodology* for getting an appropriate fuzzy model on each component, and Lemma 4.3.1 to contrast which is the model with least (fuzzy) error.

Let $\underline{X}, \mathcal{Y}$ be two variables where $\underline{X} = (X_1, ..., X_N)'$ is a random vector and $\mathcal{Y}$ is a finite fuzzy random variable taking values in $\mathcal{F}_\Lambda$ (where Λ is a finite subset of $\mathbb{I}$ containing 0 and 1). We are interested in analyzing the relationship between $\mathcal{Y}$ and $\underline{X}$. The regression model we consider can be formalized as:

$$\mathcal{Y} = \mathrm{FFN}_\Lambda\left(Y^c_{\underline{X},\underline{a}_c} + \varepsilon^c, Y^m_{\underline{X},\underline{a}_m} + \varepsilon^m, \left\{Y^{\ell,i}_{\underline{X},\underline{a}_m} + \varepsilon^\ell_i\right\}_{i=1}^{n-1}, \left\{Y^{\ell,i}_{\underline{X},\underline{a}_m} + \varepsilon^\ell_i\right\}_{i=1}^{n-1}\right), \tag{4.11}$$

where $\varepsilon^c, \varepsilon^m, \varepsilon^\ell_i$ and ε^r_i are the residuals (i.e., real-valued random variables such that $\mathrm{E}[\varepsilon^c|\underline{X}] = \mathrm{E}[\varepsilon^m|\underline{X}] = \mathrm{E}[\varepsilon^\ell_i|\underline{X}] = \mathrm{E}[\varepsilon^r_i|\underline{X}] = 0$ if we consider linear models and whose variances are finite) and $\underline{a}_c = (a_{c1}, \ldots, a_{cN})'$, $\underline{a}_m = (a_{m1}, \ldots, a_{mN})'$, $\underline{a}^i_\ell = (a^i_{\ell 1}, \ldots, a^i_{\ell N})'$ and $\underline{a}^i_r = (a^i_{r1}, \ldots, a^i_{rN})'$ are the $(N \times 1)$-vectors of the parameters related to the vector $\underline{X}$.

Therefore, the conditional expectation, that is, the population regression function, is:

$$\mathrm{E}[\mathcal{Y}|\underline{X}] = \mathrm{FFN}_\Lambda\left(Y^c_{\underline{X},\underline{a}_c}, Y^m_{\underline{X},\underline{a}_m}, \left\{Y^{\ell,i}_{\underline{X},\underline{a}_m}\right\}_{i=1}^{n-1}, \left\{Y^{r,i}_{\underline{X},\underline{a}_m}\right\}_{i=1}^{n-1}\right).$$

The function $\mathcal{Z} : \mathrm{dom}\,\underline{X} \to \mathcal{F}_\Lambda$ that we are interested in obtaining to predict $\mathcal{Y}$ from $\underline{X}$ must be defined as

$$\mathcal{Z}_{\underline{X}} = \mathrm{FFN}_\Lambda\left(Y^c_{\underline{X}}, Y^m_{\underline{X}}, \left\{Y^{\ell,i}_{\underline{X}}\right\}_{i=1}^{n-1}, \left\{Y^{r,i}_{\underline{X}}\right\}_{i=1}^{n-1}\right), \tag{4.12}$$

where Y^c : dom $\underline{X} \to \mathbb{R}$ and $Y^m_{\underline{X}}, Y^{\ell,i}_{\underline{X}}, Y^{r,i}_{\underline{X}}$: dom $\underline{X} \to [0, \infty)$ are arbitrary (nonnegative) functions.

Consider a random experiment in which we observe the variable $(\underline{X}, \mathcal{Y})$ on n statistical units, i.e., suppose that we have a random sample

$$\left\{ \left(\underline{X_s}, \mathcal{Y}_s = \mathrm{FFN}_\Lambda \left(Y^c_s, Y^m_s, \left\{ Y^{\ell,i}_s \right\}_{i=1}^{n-1}, \left\{ Y^{r,i}_s \right\}_{i=1}^{n-1} \right) \right) \right\}_{s=1}^{k}$$

obtained from $(\underline{X}, \mathcal{Y})$. If we consider the distance measure D_Λ defined in (4.10), the objective function in terms of the parameters $\underline{a}_c$, $\underline{a}_m$, $\underline{a}^i_\ell$ and $\underline{a}^i_r$ of model (4.11) is:

$$\mathcal{E} = \sum_{s=1}^{k} D_\Lambda(\mathcal{Y}_s, \mathcal{Z}_{\underline{X_s}}), \tag{4.13}$$

that is, we are looking for a function $\mathcal{Z}_{\underline{X}} = \mathrm{E}[\mathcal{Y}|\underline{X}]$ such that the fuzzy total error (4.13) is as small as possible by considering the partial order $\preccurlyeq$ w.r.t. $(\mathcal{D}_c, \Lambda)$ on $\mathcal{F}_\Lambda$. If the objective regression function $\mathcal{Z}_{\underline{X}}$ is given by (4.12), then the fuzzy total error (4.13) is

$$\begin{aligned}
\mathcal{E} = & \sum_{s=1}^{k} \mathrm{FFN}_\Lambda \Bigg(\left(Y^c_{\underline{X_s}} - Y^c_s \right)^2, \left(Y^m_{\underline{X_s}} - Y^m_s \right)^2, \\
& \qquad \left\{ \left(Y^{\ell,i}_{\underline{X_s}} - Y^{\ell,i}_s \right)^2 \right\}_{i=1}^{n-1}, \left\{ \left(Y^{r,i}_{\underline{X_s}} - Y^{r,i}_s \right)^2 \right\}_{i=1}^{n-1} \Bigg) \\
= & \, \mathrm{FFN}_\Lambda \Bigg(\sum_{s=1}^{k} \left(Y^c_{\underline{X_s}} - Y^c_s \right)^2, \sum_{s=1}^{k} \left(Y^m_{\underline{X_s}} - Y^m_s \right)^2, \\
& \qquad \left\{ \sum_{s=1}^{k} \left(Y^{\ell,i}_{\underline{X_s}} - Y^{\ell,i}_s \right)^2 \right\}_{i=1}^{n-1}, \left\{ \sum_{s=1}^{k} \left(Y^{r,i}_{\underline{X_s}} - Y^{r,i}_s \right)^2 \right\}_{i=1}^{n-1} \Bigg).
\end{aligned}$$

As a consequence, the total fuzzy error can be considered a fuzzy sum of squares of residuals and we must minimize each component to obtain the optimal solution. This means that the estimation problem of the regression parameters is solved by means of the least squares criterion through a classical regression problem. Software for linear and non-linear models is available in many statistical packages and different models can be fitted for the center and the spreads. From those models obtained for the spreads, we choose the positive models, which is always possible to perform, since we can consider a logarithmic transformation of the data or fit a positive constant (see [46]).

To sum up, taking into account classical regression techniques, the center Y^c can be related to $(X_1, ..., X_N)'$, the spread for the center Y^m or some transformation of it can be related to $(X_1, ..., X_N)'$, every left spread $\{Y^{\ell,i}\}$ for $i \in \{1, ..., n-1\}$, or some transformation of it, can be related to $(X_1, ..., X_N)'$, and every right spread $\{Y^{s,i}\}$ for $i \in \{1, ..., n-1\}$, or some transformation of it, can be related to $(X_1, ..., X_N)'$.

It is important to remark that when we consider finite FNs, the spreads can take small values. Therefore it is easy to obtain fitted models that could take negative values. Only for spreads in which this problem could arise, we will overcome it replacing the spread $\hat{Y}^r$ by $\max\{\, \hat{Y}^r_{\underline{X},\underline{a}_r} + \varepsilon^r, 0\,\}$ in the fitted regression model ($r \in \{m, \ell, s\}$). This replacement let us guarantee that the response is always a FN.

For each of the possible combinations of a center and the corresponding spreads we calculate a fuzzy error using the previous equation. Finally we sort the fuzzy models using the partial order $\preccurlyeq$ w.r.t. $(\mathcal{D}_c, \Lambda)$ and, for the optimal solution of the fuzzy regression problem, we choose the fuzzy model with the lowest fuzzy error.

Note that the main aim of this memory is to develop a fuzzy methodology, using finite FNs, that can be considered easy to understand and powerful. In this sense, many people are used to consider stepwise linear regression methods to solve problems. This is a method for regressing multiple variables which essentially does multiple regression a number of times, each time removing the variables that are not important for the explanation, and leaving the variables that better explain the distribution, considering the forward selection approach (see Section 2.3.6). We propose to consider models obtained by this method in the solution of the fuzzy problem.

CHAPTER 5

A case study: Chinese convertible bonds

In this chapter we illustrate how the fuzzy regression methodology introduced in Chapter 4 can be applied in a real-world context: the case of Chinese convertible bonds, which can be interpreted as a fuzzy random variable.

First of all, we approximate its values as an appropriate finite fuzzy random variable and later we compute a fuzzy model in order to predict the values of such varia ble depending on three real random variables: the price of A-share, the price of H-share and SHIBOR rate.

5.1 Introduction to convertible bonds

In this section, in order to illustrate the theoretical results described in the previous sections, we use data obtained from Bloomberg database, which is one of the most used financial information system worldwide. Firstly some financial concepts are introduced.

A convertible bond is a financial instrument composed of two parts; a bond and an option. A bond is a fixed income security used by corporations, countries and supranational institutions to raise debt. Investors in bonds lend money to these institutions and in return they receive interest (typically every quarter or every year) and at the maturity date (when the bond expiries) it receives its last interest payment plus the notional (the amount that the investor gave initially to the company). The previously described bond is the most basic type of bond and it should be noted that there are numerous variations and more sophisticated bonds such as discount bonds or hybrid perpetual bonds.

For the purposes of this analysis we focus on convertibles bonds in which the underlying bond is a straight "plain vanilla" bond. The other component of the convertible bond is the option. The option gives the holder (the investor) the right to exchange the bond for stock (equity) of the issuer company if the stock price reaches a certain (predetermined) level.

Convertible bonds are a popular financial instrument because they gives investors protection on the downfall (fixed income investors have priority in the case of company liquidation) while at the same time gives the investors upside. The return on a fixed income (if hold to maturity) is the interest received plus the discount (if there was any) between the buying price and the par value. In other words, the returns on fixed income securities are capped and for instance a massive improvement in the results of the company would have very little impact on the returns received by the investor.

On the other hand, an equity investor could benefit greatly if the share price goes up, which tends to be linked with strong business performance of the company but have the risk that in the case that the company goes bankrupt they are likely to receive no return at all for their investment (in fact, in the event of a bankruptcy equity investors tend to lose virtually all its initial investment). In the case of bankruptcy fixed income investors have priority of claims over the assets of the company compared to equity investors.

In the case of a convertible bond its price is clearly impacted by changes in the price of the underlying stock (as there is an equity option embedded in the bond) but it is also impacted, as a fixed income security, by the market interest level. Interest rates and bond prices are inversely related with a decrease in interest rate typically causing an increase in price. This can be easily seen with an example.

Let us consider the case of a straight fixed income security (no option embedded) that is paying a 4 percent annual interest. Now, if the market interest rate declines to for example 3 percent and there is no change in price in the bond that we are analyzing then rational investors would sale other bonds and buy the bond paying 4 percent. In reality want happens is that the price of the 4 percent bond increases until the level in which the extra money that an investor has to pay for buying the 4 percent bond makes the yield on this bond and the market level the same.

5.1.1 Chinese convertible bonds

Every country has their peculiarities given to historical reasons, the country organization, legal systems and many other factors. In the case of the Chinese case there are two markets, one in mainland China and one in Hong Kong. These two markets operates independently and are regulated by different institutions. There are two stock exchanges in mainland China: the Shanghai Stock Exchange and the Shenzhen Stock Exchange.

In Hong Kong there is only one stock exchange, The Hong Kong Stock Exchange. The stocks listed in mainland China are usually called A-shares while the stock listed in Hong Kong are usually called H-shares. There are a significant amount of dual listed companies i.e., the same companies with stocks in the mainland and Hong Kong. For instance, China Construction Bank (CCB), which is a large national bank, is listed in Hong Kong as well as in Shanghai.

In this situation we have two different stocks (trading at two different prices) that represent ownership in the same company. In a perfect, frictionless world the prices of those two stocks should be the same but a large amount of factors causes that in fact the prices are different. It should be noted that there are no dual listed companies between the two mainland stock exchanges, i.e., there are no companies listed at the same time in the Shanghai and Shenzhen stocks exchanges. In the case of a dual listed company, it would seem reasonable to assume that the prices in the two exchanges could have an impact on the price of the convertible bond.

As previously mentioned, another factor that can impact the price of a convertible bond is the market rate. SHIBOR (Shanghai Interbank Borrowing Rate) is the rate at which the large domestic banks lend to each other in China (and it is equivalent to the LIBOR rate for western countries). SHIBOR is a very good proxy for the market interest rate and hence it was included in this analysis.

5.2 Fuzzy regression analysis

In this section, to illustrate the theoretical results we use data obtained from Bloomberg (one of the most used financial information system that is widely used by professionals worldwide).

The convertible bond chosen for the analysis was issued by the Industrial and Commercial Bank of China (ICBC) in 2010, its maturity date is 2016 and its Bloomberg identifer is EI3894694 (a Bloomberg screen shot of this bond can be found in Figure 5.1). ICBC is one of the largest Chinese banks (commonly referred to as the big 4). ICBC is listed in The Shanghai Stock Exchange (ticker 601398 CH) and in the Hong Kong Stock Exchange (ticker 1398 HK).

Figure 5.1: Bloomberg screen shot of the convertible bond chosen

All the data were obtained from Bloomberg (Bloomberg Professional), which is one the most commonly professional data providers used in the industry. The data set consists of the price for the convertible bond, the A-share, the H-share and SHIBOR for the period between September 2010 and May 2014, this gives 45 data for each series.

In this context it is clear that the price of a convertible bond is related to the value of the action in which can be converted (in this case A-shares) but as the share price of type A may be influenced by the price of the action of H type (and viceversa) seems logical to include the price of these two types of shares in the analysis. Another factor that seems interesting to consider is the Shanghai interbank offered rate (SHIBOR).

To avoid the loss of information, experts in financial analysis described the variables in terms of finite FNs. We assume that the central spreads are zero. Therefore, the purpose of the following example is to study the statistical relationship between the price of a convertible bond ($\mathcal{Y}$) and the price of A-share ($\underline{\mathcal{X}_1}$), H-share ($\underline{\mathcal{X}_2}$), and SHIBOR rate ($\underline{\mathcal{X}_3}$). We have considered a set of 45 data, from September 2010 to May 2014 (see Tables 5.1 and 5.2). To avoid the loss of information, experts in financial analysis described the variables in terms of finite FNs. We assume that the central spreads are zero.

Table 5.1: Data.

Date	Response variable, $\mathcal{Y}$	Explanatory variable 1, $\underline{\mathcal{X}_1}$
t	$\text{FFN}_\Lambda(Y^c, Y^m, Y^{l,1}, Y^{l,2}, Y^{r,1}, Y^{r,2})$	$\text{FFN}_\Lambda(X_1^c, X_1^m, X_1^{l,1}, X_1^{l,2}, X_1^{r,1}, X_1^{r,2})$
Sep-10	$(110.86, 0, 0.30, 0.46, 0.75, 0.47)$	$(3.98, 0, 0.03, 0.05, 0.04, 0.02)$,
Oct-10	$(123.63, 0, 2.92, 2.99, 0.99, 1.06)$	$(4.38, 0, 0.13, 0.13, 0.09, 0.03)$
Nov-10	$(124.11, 0, 1.00, 3.39, 1.90, 0.93)$	$(4.52, 0, 0.06, 0.16, 0.17, 0.03)$
⋮	⋮	⋮
May-14	$(101.81, 0, 0.31, 0.81, 0.60, 0.65)$	$(3.48, 0, 0.03, 0.03, 0.04, 0.05)$

The Figures 5.3-5.2 display the plot of the fuzzy variables considered (images are in order from left to right).

We use these data to regress the fuzzy response variable $\mathcal{Y}$ about the three fuzzy exploratory variables $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{X}_2$. This problem can be reduced to one that can be solved with the methodology described in the previous sections and considering the crisp random vector $X = (x_1^i, x_1^c, x_1^s, x_2^i, x_2^c, x_2^s, x_3^i, x_3^c, x_3^s)$ (in this case, $x_1^m = x_2^m = x_3^m = 0$) as the exploratory random variable.

We follow the following steps:

Step 1. Generate a *Table of fits* with a statistical package listing the mo-

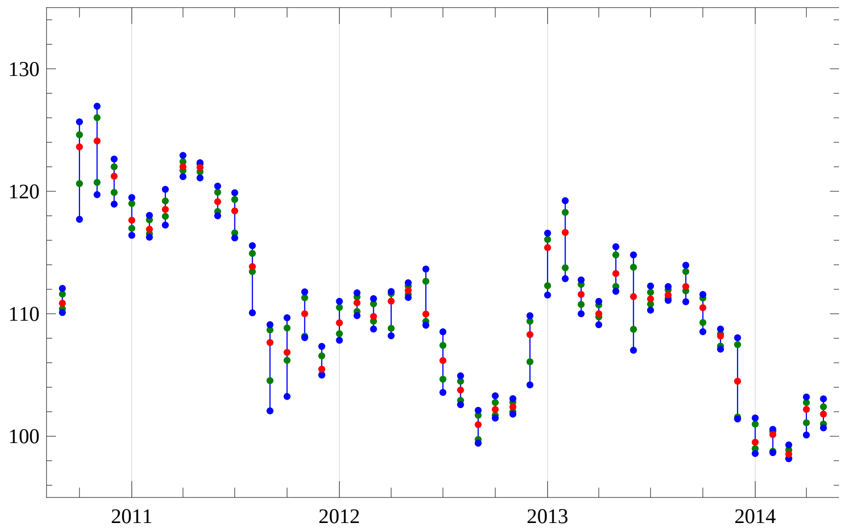

Figure 5.2: Price of the convertible bond ($\mathcal{Y}$) (from Table 5.1)

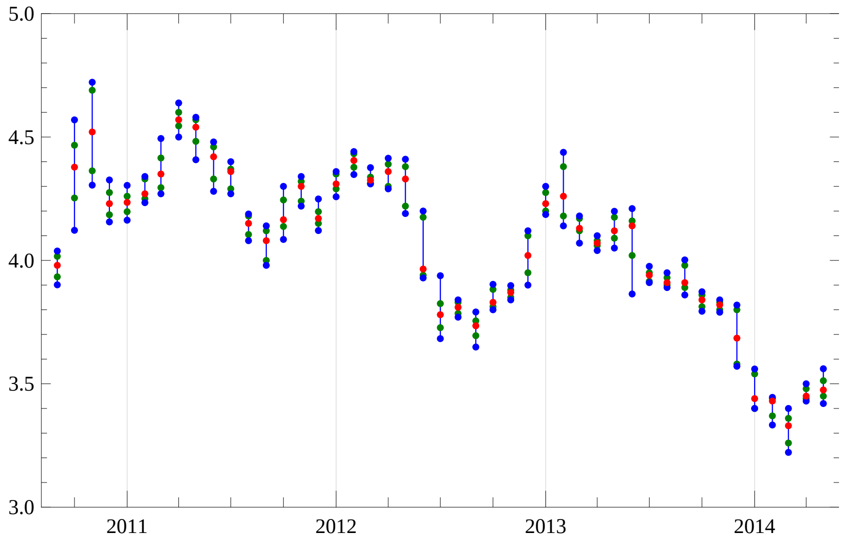

Figure 5.3: Prices of A-share ($\mathcal{X}_1$) (from Table 5.1)

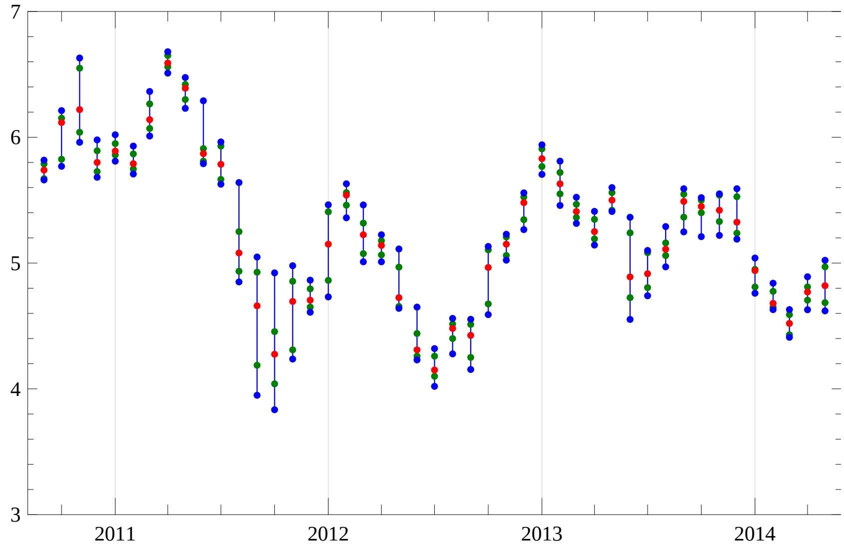

Figure 5.4: H-share values ($\mathcal{X}_2$) (from Table 5.2)

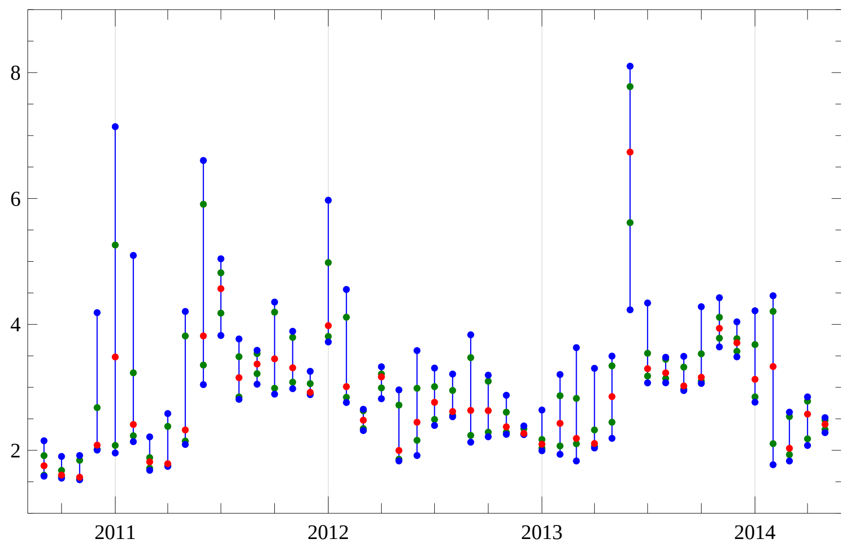

Figure 5.5: SHIBOR rate ($\mathcal{X}_3$) (from Table 5.2)

Table 5.2: Data.

Date	Explanatory variable 2, $\underset{\sim}{\mathcal{X}_2}$	Explanatory variable 3, $\underset{\sim}{\mathcal{X}_3}$
t	$\text{FFN}_\Lambda(X_2^c, X_2^m, X_2^{l,1}, X_2^{l,2}, X_2^{r,1}, X_2^{r,2})$	$\text{FFN}_\Lambda(X_3^c, X_3^m, X_3^{l,1}, X_3^{l,2}, X_3^{r,1}, X_3^{r,2})$
Sep-10	$(5.74, 0, 0.01, 0.07, 0.05, 0.03)$,	$(1.75, 0, 0.04, 0.15, 0.16, 0.24)$
Oct-10	$(6.11, 0, 0.06, 0.29, 0.04, 0.06)$	$(1.61, 0, 0.03, 0.02, 0.07, 0.22)$
Nov-10	$(6.22, 0, 0.08, 0.18, 0.33, 0.08)$	$(1.57, 0, 0.02, 0.02, 0.27, 0.08)$
⋮	⋮	⋮
May-14	$(4.82, 0, 0.07, 0.14, 0.15, 0.05)$	$(2.41, 0, 0.05, 0.08, 0.06, 0.04)$

dels for the center and the spreads in decreasing order of some goodness-of-fit statistics. The stepwise method was used in this step to select the explanatory variables included in the the classical linear models which are needed to build the fuzzy regression model. In the case that only a single predictor variable contribute significantly to explaining the variability in the dependent variable, statistical packages can fit a variety of functional forms and support different goodness-of-fit statistics for parametric models. These functionals forms have been considered in this case.

Step 2. Use the distance measure selected by the researcher from the beginning and calculate the fuzzy errors for each of the possible for each of the possible fuzzy regression models.

Step 3. Sort the fuzzy models with the partial order $\preccurlyeq$ introduced in Definition 4.2.2 taking into account the fuzzy errors obtained in Step 2 for each one. In the examples, these results are shown in a table named *Table of fuzzy fits*. Note that the order established between the models using real numbers in step 1, which would seem logical to keep it, do not necessarily coincide with the ranking obtained when models are arranged in this step according to the fuzzy order above introduced.

Step 4. To find the optimal solution of the fuzzy regression problem we choose the functions corresponding to the lowest fuzzy error.

5.3 Results

In order to search for a suitable fuzzy regression model capable to express the statistical relationship between the price of a convertible bond ($\mathcal{Y}$) and the price of A-share ($\mathcal{X}_1$), H-share ($\mathcal{X}_2$), and SHIBOR rate ($\mathcal{X}_3$), we consider the methodology explained in the previous chapters.

The fitted model for the center is given by:

$$\hat{y}^c_{\mathcal{X}} = 11.268\, x^c_2 + 5.049\, x^c_3 + 17.237\, x^s_2 + 36.472.$$

The center spreads are zero, this means that

$$\hat{y}^m_x = 0.$$

Only one variable results significant through stepwise regression for the left spread model. Then non-linear models were considered to find the best fit. The fitted model for the left spread is given by:

$$\hat{y}^i_{\mathcal{X}} = \sqrt{1.3545 + 436.290\,(x^i_2)^2}$$

The fitted model for the right spread is given by:

$$\hat{y}^s_{\mathcal{X}} = 12.935\, x^s_2 + 2.838\, x^i_3.$$

Therefore the fuzzy regression model with the lowest error is given by

$$\begin{aligned}\hat{\mathcal{Y}} = \mathrm{FFN}_\Lambda \Big(& 35.002 + 11.231 X_1^c + 5.480 X_2^c + 44.545 X_1^{r,2}, 0, \\ & \exp\{-5.576 + 10.683 X_2^{l,1} + 0.657 X_2^c + 2.905 X_2^{r,2} + 11.045 X_1^{r,2}\}, \\ & \max\{0.191 + 18.938 X_1^{l,2} + 4.872 X_2^{l,1} - 0.590 X_3^{s,1}, 0\}, \\ & \exp\{-1.555 + 10.630 X_1^{r,1} + 0.300 X_3^c\}, \\ & \exp\{-1.801 + 12.791 X_1^{r,2} + 5.501 X_1^{r,1} + 0.145 X_3^c\} \Big).\end{aligned}$$

The corresponding fuzzy sum of squares of residuals is

$$\mathrm{FFN}_\Lambda(276.20, 0, 14.35, 18.99, 6.47, 0.64)$$

and approximately 86.22% of the total variation of price of the convertible bond is explained by the previous fuzzy model.

CHAPTER 6

Conclusions and future work

This memory has introduced a fuzzy regression model when the response variable is a finite fuzzy random variables. We have considered a fuzzy distance measure that allows to solve the problem considering classical multiple regression models and stepwise methods with crisp variables. The goodness of fit have been measured by the fuzzy total error. However taking into account that a real-life study can consider at the same time, fuzzy and crisp variables, a fuzzy regression methodology should be able to deal, in a unified manner both cases. Therefore similar to the classical regression models, the goodness of fit can be measured using a real R-squared. Notice that it is not necessary to impose non-negativity conditions from the beginning and the problem has been easily solved in an optimal fuzzy way.

Summarizing, the main advantages of this methodology are:

- It provides the researchers an easy approach for the problem of analyzing regression relationships when the observed data can be affected by different sources of uncertainty.
- This study does not consider non-negativity restrictions from the beginning.

- Researchers can model the statistical relationships between fuzzy variables with a method that may be applied in different contexts.

- The technique is not limited to consider linear models.

- The regression proposed in this paper is not limited to numeric intervals or triangular fuzzy variables (as in other papers).

- The methodology is not limited to crisp explanatory variables.

Another important objective of this analysis was to give market participants (investors) a quantitative tool to help them make investment decisions and the obtained fuzzy regression equation seems to add value to the investment process. This example was used to illustrate the proposed methodology and demonstrate the efficiency of this method. It should be noted that the scope was not to create an algorithm to tell an investor when to buy or sell the convertible bond but a quantitative tool that would help him/her doing so by himself/herself according to his/her own expectations on the independent variables (A-share price, H-share price and SHIBOR level).

The results obtained in the case study have some interesting features. The model obtained is not too complex and hence allows the practitioner to analyze from economic and finance fundamentals the regression equation obtained. From a financial point of view is surprising that the fitted model explains approximately 85% of the variability of the price of the convertible bond without including the A-share price in the regression equation. This is perhaps due to having two share prices (one in Hong Kong) and one in Shanghai. It should be noted that share prices of dual listed Chinese companies move in the same direction only roughly 60% of the time, i.e., the share price (of the same company) goes up in the Shanghai Stock Exchange and the Shenzhen Stock Exchange only in roughly 60% of the days. There are multiple factors that cause these price discrepancies such as different currencies, holidays, trading hours or short-selling regulation.

Some open problems that deserve our attention in the near future are as follows:

- Throughout this report has arisen the need to introduce a partial order in the set of finite data. This partial order can be considered to establish a ranking. It would be interesting to study the corresponding properties and its possible applications to an economic context.
- The study of the problem of the nearest parametric approximation of a finite fuzzy number. Some properties (translation invariance, scale invariance, additivity, preservation of value, ambiguity, expected value) of the nearest parametric approximation can also be studied.

List of Figures

Bibliography

[1] M. Adabitabar Firozja, G.H. Fath-Tabar, Z. Eslampia, The similarity measure of generalized fuzzy numbers based in interval distance, Appl. Math. Lett. **25** (2012) 1528–1534.

[2] C. Aguilar-Peña, A.F. Roldán-López-de-Hierro, C. Roldán-López-de-Hierro, J. Martínez-Moreno, A family of fuzzy distance measures of fuzzy numbers, Soft Comput. **20** (2016) (1) 237–250.

[3] T. Allahviranloo, M.A. Firozja, Ranking of fuzzy numbers by a new metric, Soft Comput. **14** (2010) 773–783.

[4] T. Allahviranloo, R. Nuraei, M. Ghanbari, E. Haghi, A.A. Hossein-zadeh, A new metric for L-R fuzzy numbers and its application in fuzzy linear systems, Soft Comput **16** (2012) 1743–1754.

[5] A.I. Ban, L. Coroianu, Nearest interval, triangular and trapezoidal approximation of a fuzzy number preserving ambiguity, Int. J. Approx. Reasoning **53** (2012) 805–836.

[6] A.I. Ban, On the nearest parametric approximation of a fuzzy number – Revisited, Fuzzy Sets Syst. **160** (2009) 3027–3047.

[7] J.J. Buckley, L.J. Jowers, Monte Carlo methods in fuzzy optimization, Studies in Fuzziness and Soft Computing, Volume **222**, Springer, 2008.

[8] J. Casasnovas, J.V. Riera, Discrete fuzzy numbers defined on a subset of natural numbers, Theoretical Advances and Applications of Fuzzy Logic and Soft Computing: Advances in Soft Computing **42** (2007) 573–582.

[9] C.T. Chang, An approximation approach for representing S-shaped membership functions, IEEE Trans. Fuzzy Syst. **18** (2) (2010) 412–424.

[10] S.S.L. Chang, L.A. Zadeh, On fuzzy mapping and control, IEEE Trans. Syst. Man. Cybern. **2** (1) (1972) 30–34.

[11] S.H. Chatterjee, S. Ali, B. Price, Regression Analysis by Example. 5th Edition. Wiley (2012).

[12] S.M. Chen, New methods for subjective mental workload assessment and fuzzy risk analysis, Cybernetics and Systems **27** (1996) 449–472.

[13] S.J. Chen, S.M. Chen, Fuzzy risk analysis based on similarity measures of generalized fuzzy numbers, IEEE Trans. Fuzzy Syst. **11** (2003) 45–56.

[14] L.H. Chen, C.C. Hsueh, Fuzzy regression models using the least-squares method based on the concept of distance, IEEE Trans. Fuzzy Syst. **17** (2009) 1259–1272.

[15] C.H. Cheng, A new approach for ranking fuzzy numbers by distance method, Fuzzy Sets Syst. **95** (1988) 307–317.

[16] R. Christensen, Log-linear models and logistic regression. Springer-Verlag (1997).

[17] P. Diamond, Fuzzy least squares. Inform. Sci. **46** (1988) 141–157.

[18] D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz, Soft methods for handling variability and imprecision. Advances in soft computing, Springer, Berlin (2008).

[19] D. Dubois, H. Prade, Operations on fuzzy numbers, International Journal of System Sciences **9** (1978) 613–626.

[20] D. Dubois, H. Prade, Possibility Theory, New York, Plenum Press, 1988.

[21] D. Dubois, H. Prade, Fuzzy elements in a fuzzy set. Proc. 10th Internat. Fuzzy Systems Assoc. (IFSA) Congr. 2005, 55-60, Beijing, Springer.

[22] M.B. Ferraro, P. Giordani, A multiple linear regression model for imprecise information, Metrika **75** (2012) 1049–1068.

[23] T. Ganesan, P. Vasant, I. Elamvazuthi, Hybrid PSO approach for solving non-convex optimization problems, Archives of Control Sciences **22** (1) (2012) 5–23.

[24] R. Goetschel, W. Voxman, Elementary fuzzy calculus, Fuzzy Sets Syst. **18** (1986) 31–43.

[25] S. Gorard, Revisiting a 90-year-old debate: the advantages of the mean deviation, British Educational Research Association Annual Conference, University of Manchester (2004) 16–18.

[26] P. Grzegorzewski, Trapezoidal approximations of fuzzy numbers preserving the expected interval - Algorithms and properties, Fuzzy Sets Syst. **159** (2008) 1354–1364.

[27] D. Guha, D. Chakraborty, A new approach to fuzzy distance measure and similarity measure between two generalized fuzzy numbers, Appl. Soft Comput. **10** (2010) 90–99.

[28] F.E. Harrell, Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis, Springer-Verlag, New York (2001).

[29] D.H. Hong, C. Hwang, Support vector fuzzy regression machines. Fuzzy Set Syst. **138** (2002) 271–281.

[30] D.W. Hosmer, S. Lemeshow, R.X. Sturdivant, Applied logistic regression. 3rd Edition, Wiley (2013).

[31] C. Kao, C.L. Chyu, Least-squares estimates in fuzzy regression analysis, Eur. J. Oper. Res. **148** (2003) 426–435.

[32] V. Krätschmer, Limit distribution of least squares estimators in linear regression models with vague concepts, J. Multivariate Anal. **97** (2006) 1044–1069.

[33] R. Koenker, K.F. Hallock, Quantile Regression, Journal of Economic Perspectives **15** (2001) 143–156.

[34] H. Kwäkernaak, Fuzzy random variables–I. Definitions and theorems, Inform. Sci. **15** (1978) 1–29.

[35] Y. Li, G.R. Arce, A maximum likelihood approach to least absolute deviation regression, EURASIP Journal on Applied Signal Processing **12** (2004) 1762–1769.

[36] J.G. Lin, Q.Y. Zhuang, C. Huang, Fuzzy statistical analysis of multiple regression with crisp and fuzzy covariates and applications in analyzing economic Data of China, Computational Economics **39** (2012) 29–49.

[37] M. Mizumoto, J. Tanaka, Some properties of fuzzy numbers, in: M.M. Gupta et al. (Eds.), Advances in Fuzzy Set Theory and Applications 153–164, North-Holland, NewYork (1979).

[38] W. Näther, Regression with fuzzy random data, Comp. Stat. Data Anal. **51** (2006) 235–252.

[39] A.M. Nejad, M. Mashinchi, Ranking fuzzy numbers based on the areas on the left and the right sides of fuzzy number, Comput. Math. Appl. **61** (2011) 431–442.

[40] R.F. Phillips, Least absolute deviations estimation via the EM algorithm, Statistics and Computing **12** (2002) 281–285.

[41] M. L. Puri, D. A. Ralescu, Fuzzy random variables, J. Math. Anal. Appl. **114** (1986) 409–422.

[42] S. E. Rodabaugh, Categorical foundations of variable-basis fuzzy topology, Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory (U. Höhle and S. E. Rodabaugh, eds.), The Handbooks of Fuzzy Sets Series, vol. 3, Dordrecht: Kluwer Academic Publishers, 273–388 (1999).

[43] S. E. Rodabaugh, Fuzzy real lines and dual real lines as poslat topological, uniform, and metric ordered semirings with unity, Mathematics of Fuzzy Sets: Logic, Topology and Measure Theory (U. Höhle and S. E. Rodabaugh, eds.), The Handbooks of Fuzzy Sets Series, vol. 3, Dordrecht: Kluwer Academic Publishers, 607–631 (1999).

[44] A. Roldán, J. Martínez-Moreno, C. Roldán, On interrelationships between fuzzy metric structures, Iran. J. Fuzzy Syst. **10** (2013) 133–150.

[45] A.F. Roldán López de Hierro, C. Aguilar, J. Martínez-Moreno, C. Roldán, Estimation of a fuzzy regression model using fuzzy distances, IEEE Trans. Fuzzy Syst. **24** (2) (2016) 344–359.

[46] A. Roldán, J. Martínez-Moreno, C. Roldán, A fuzzy regression model based on distances and random variables with crisp input and fuzzy output data: a case study in biomass production, Soft Comput. **16** (2012) 785–795.

[47] A. Roldán, J. Martínez-Moreno, C. Roldán, Some applications of the study of the image of a fuzzy number: Countable fuzzy numbers, operations,

regression and a specificity-type ordering, Fuzzy Sets Syst. **257** (2014) 204-216.

[48] A. Roldán, M. de la Sen, J. Martínez-Moreno, C. Roldán: An approach version of fuzzy metric spaces including an ad hoc fixed point theorem, Fixed Point Theory Appl. **2015**, Article ID 2015:33, 23 pages (2015).

[49] B. Schweizer, A. Sklar, Probabilistic Metric Spaces, New York, Dover Publications (2005).

[50] M. Shi, M.A. Lukas, An L_1 estimation algorithm with degeneracy and linear constraints, Comput. Stat. Data Anal. **39** (2002) 35–55.

[51] E. Siemsen, K.A. Bollen, Least absolute deviation estimation in structural equation modeling, Sociological Methods and Research **36** (2007) 227–265.

[52] S.M. Taheri, M. Kelkinnama, Fuzzy linear regression based on least absolutes deviations, Iran. J. Fuzzy Syst. **9** (2012) 121–140.

[53] H. Tanaka, I. Hayashi, J. Watada, Linear regression analysis with fuzzy model, IEEE Trans. Syst. Man. Cybern. SMC-12 (1982) 903–907.

[54] H. Tanaka, I. Hayashi, J. Watada, Possibilistic linear regression analysis for fuzzy data, Eur. J. Oper. Res. **40** (1989) 389–396.

[55] L. Tran, L. Duckstein, Comparison of fuzzy numbers using a fuzzy distance measure, Fuzzy Sets Syst. **130** (2002) 331–341.

[56] P. Vasant, N. Barsoum, Hybrid pattern search and simulated annealing for fuzzy production planning problems, Comput. Math. Appl. **60** (4) (2010) 1058–1067.

[57] P. Vasant, Meta-heuristics optimization algorithms in Engineering, Business, Economics and Finance (2013).

[58] P. Vasant, T. Ganesan, I. Elamvazuthi, Hybrid tabu search Hopfield recurrent ANN fuzzy technique to the production planning problems: a case study of crude oil in refinery industry. International Journal of Manufacturing, Materials, and Mechanical Engineering **2** (1) (2012) 47–65.

[59] J. Vicente Riera, J. Torrens, Aggregation of subjective evaluations based on discrete fuzzy numbers, Fuzzy Sets Syst. **191** (2012) 21–40.

[60] W. Voxman, Canonical representations of discrete fuzzy numbers, Fuzzy Sets Syst. **118** (2001) 457–466.

[61] C.X. Wu, M. Ma, The basic of fuzzy analysis, National Defence Industry Press, Beijing (1991).

[62] B. Wu, N.F. Tseng, A new approach to fuzzy regression models with application to business cycle analysis, Fuzzy Sets Syst. **130** (2002) 33–42.

[63] S. Weisberg, Applied linear regression. 4th Edition. Wiley (2014).

[64] M.S. Yang, C.H. Ko, On a class of fuzzy c-numbers clustering procedures for fuzzy data, Fuzzy Sets and Syst., **84** (1996) 49–60.

[65] L.A. Zadeh, Fuzzy set, Inform. Control **8** (1965) 338–353.