International Journal of Approximate Reasoning 54 (2013) 1217–1242

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Parametric regression analysis of imprecise and uncertain data in the fuzzy belief function framework ☆

Zhi-gang Su [a,*], Yi-fan Wang [b], Pei-hong Wang [a]

[a] Key Laboratory of Energy Thermal Conversion and Control of the Ministry of Education, School of Energy and Environment, Southeast University, Nanjing 210096, China

[b] Information Networking Institute, Carnegie Mellon University, Pittsburgh, PA 15217, USA

ARTICLE INFO

Article history:
Received 17 February 2012
Received in revised form 7 November 2012
Accepted 20 February 2013
Available online 1 March 2013

Keywords:
Evidence theory
Fuzzy belief function
Uncertain data
Fuzzy data
EM algorithm
Regression analysis

ABSTRACT

In this paper, parametric regression analyses including both linear and nonlinear regressions are investigated in the case of imprecise and uncertain data, represented by a fuzzy belief function. The parameters in both the linear and nonlinear regression models are estimated using the *fuzzy evidential EM algorithm*, a straightforward fuzzy version of the evidential EM algorithm. The nonlinear regression model is derived by introducing a kernel function into the proposed linear regression model. An unreliable sensor experiment is designed to evaluate the performance of the proposed linear and nonlinear parametric regression methods, called *parametric evidential regression* (PEVREG) models. The experimental results demonstrate the high prediction accuracy of the PEVREG models in regressions with crisp inputs and a fuzzy belief function as output.

1. Introduction

Supervised learning is concerned with the prediction of a response (i.e., the output variable) based on a learning set $\{(\mathbf{u}_i, x_i)|i = 1 \text{ to } n\}$, where $\mathbf{u}$ is a vector of p input variables and x is the response. This problem is also referred to as regression when the output is a quantitative measurement. Regression analysis is one of the most popular statistical techniques for the identification of a functional relationship between the input variables and response. Many techniques have been proposed in the literature to estimate the regression function, including nearest-neighbor methods, smoothing splines, multi-layer perceptions, radial basis function networks, and projection pursuit methods (see, e.g., [19]). These methods have proven very efficient in a wide range of applications, but they also suffer from certain limitations.

Classical regression techniques [2,19,22] assume perfect knowledge of the value of the response x for a given learning sample. That is to say, the observations are assumed to be both precise and certain. However, in many real-life situations, we cannot obtain such ideal observations. The information about the response is often obtained using measuring devices, or sensors, with limited precision and reliability. The imprecise observations of the response may then be modeled better by real intervals $[x_i^-, x_i^+]$ or (triangular) fuzzy numbers $\widetilde{A}_i$ with core x_i. Several approaches have been proposed for processing such learning data, such as linear fuzzy regression models [3,18,34] and fuzzy [33,38] or neuro-fuzzy inference systems [20]. However, the uncertainty in the observations is not easily accounted for in these approaches. For instance, there is always uncertainty in cases of poor sensor reliability. An observation may be imprecise, uncertain, or both, and each of these situations must be properly represented in a learning system [17,28].

To address this issue, Petit-Renaud and Denoeux [23] first proposed a new approach to regression analysis based on fuzzy belief function theory, which they termed *evidential regression* (EVREG) modeling. The basic idea of EVREG modeling is that

☆ This work was supported by the National Natural Science Foundation of China (No. 51106025).
* Corresponding author. Tel.: +86 25 83794520; fax: +86 25 57714489.
E-mail addresses: suzhigang2008@163.com (Z.-g. Su), ivanrex@gmail.com (Y.-f. Wang), phwang@seu.edu.cn (P.-h. Wang).

0888-613X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.ijar.2013.02.008

each training sample in the neighborhood of the input vector **u** is viewed as a piece of evidence regarding the value of the output x, and the evidence items are discounted as a function of their distance to **u** and combined using the conjunctive rule of combination (the unnormalized Dempster's rule). The result is a fuzzy belief assignment that quantifies the distribution of beliefs concerning the value of x. The EVREG method is therefore a naturally nonparametric regression model, which leads to some desirable properties – for instance, the output of EVREG reflects not only the quality of the training data but also the relevance of these data to the prediction task at hand.

There has been a recent surge of interest in uncertain data mining (see, e.g., [10,12,13,15,16,31]). A significant contribution to this field was the extension of the Expectation–Maximization (EM) algorithm [9] to uncertain data in the form of the so-called evidential EM (E^{2}M) algorithm proposed by Denoeux [12,13]. This algorithm permits estimation of the parameters of statistical models in cases where the data are uncertain and represented by a (crisp) belief function. Denoeux's work provides an opportunity to study parametric regression, including both linear and nonlinear regressions, in cases where only imprecise and uncertain data are available. In this paper, we apply Denoeux's E^{2}M algorithm to fuzzy belief functions and provide a straightforward presentation of the fuzzy version of the E^{2}M algorithm (the Fuzzy E^{2}M algorithm or FE^{2}M). Through the FE^{2}M algorithm, we introduce a new approach to regression analysis, called the *parametric evidential regression model* (PEVREG). Unlike EVREG, the new PEVREG approach is naturally parametric.

The paper is organized as follows. As background for the remainder of the paper, Section 2 provides a brief overview of fuzzy belief function theory. In Section 3, traditional likelihood methods and the E^{2}M algorithm are applied to fuzzy belief functions. Section 4 investigates the PEVREG model in the fuzzy belief function framework. Section 5 presents an unreliable sensor experiment designed to demonstrate the performance of PEVREG, and the paper is concluded in the final section.

2. Fuzzy belief function theory

In this section, belief function theory and its fuzzy extension [7,8,25,27,30] are briefly introduced as background for the remainder of the paper. Section 2.1 reviews the basic concepts, and Section 2.2 introduces the notion of cognitive independence.

2.1. Basic concepts

Let $\Omega = \{\omega_1, \omega_2, \ldots, \omega_c\}$, the *frame of discernment*, be a collectively exhaustive and mutually exclusive set of c hypotheses or propositions. A basic belief assignment (BBA, or mass function) is a function m: $2^\Omega \to [0, 1]$, satisfying:

$$\sum_{A \subseteq \Omega} m(A) = 1. \tag{1}$$

A BBA is considered to be normal if $m(\phi) = 0$; otherwise it is subnormal. Any subset A of Ω such that $m(A) > 0$ is called a focal element of m. For two given BBAs m_1 and m_2 representing two independent sources of evidence and any binary set operation ∇, the fusion of the two BBAs, denoted by $m = m_1 \circledtriangledown m_2$, may be defined as follows:

$$m(C) = \sum_{A \nabla B = C} m_1(A) m_2(B). \tag{2}$$

The conjunctive rule is obtained when $\nabla = \cap$, and the disjunctive rule is obtained when $\nabla = \cup$. Note that the conjunctive rule may produce a subnormal belief assignment. The Dempster rule, $\oplus$, converts the subnormal BBA into a normal one, m^*, defined for $A \neq \phi$ and $m^*(\phi) = 0$ as follows:

$$m^*(A) = \frac{m(A)}{1 - m(\phi)}. \tag{3}$$

Two evidential functions derived from the BBA are the *belief function, Bel*, and the *plausibility function, Pl*, defined for all A in Ω as:

$$Bel(A) = \sum_{\phi \neq B \subseteq A} m(B), \tag{4}$$

$$Pl(A) = \sum_{A \cap B \neq \phi} m(B). \tag{5}$$

These two functions are connected by the relation $Pl(A) = 1 - Bel(\bar{A})$ for all $A \subseteq \Omega$. The function pl: $\Omega \to [0, 1]$ such that $pl(\omega) = Pl(\{\omega\})$ is called the *contour function* associated with m. If m is Bayesian, we have $pl(\omega) = m(\{\omega\})$. In this case, pl is a probability distribution. A probability distribution P is said to be compatible with m if $Bel(A) \leq P(A) \leq Pl(A)$