Smooth extensions of fuzzy if-then rule bases

Thomas Vetterlein

Department of computer sciences 1
Dortmund University
44227 Dortmund
Thomas.Vetterlein@uni-dortmund.de

Summary. In order to extend fuzzy if-then rules bases, we propose to make use of a method which has been developed for the interpolation of crisp data – the multivariate spline interpolation. Among the various possibilities of how to accomplish the necessary generalisations, we describe here the probably simplest method: We apply spline interpolation to fuzzy data which itself is approximated by vectors of a finite-dimensional real linear space.

1 The problem

We consider in this paper the question how to extend a fuzzy if-then rule base to a total function, requiring that this function is, in some reasonable sense, as smooth as possible. Roughly speaking, we assume to be in a situation of the following kind: We are given two variables both of which may take sharp or unsharp values in a bounded subset of a finite-dimensional real Euclidean space; the first of these variables uniquely determines the second one; and we know about this dependence only from a few special cases. The problem is then how to determine a function which maps the whole domain of the first variable to the domain of the second one, thereby comprising not only the known cases, but also minimising a parameter which measures in some reasonable sense the function's curvature.

Leaving the smoothness requirement aside, the problem is well-known, and various methods have been proposed. Before comparing our approach to already existing ones, we shall first specify the formal background of our considerations, so as to be able to clarify the idea of this paper.

We shall work with what could be called standard fuzzy sets according to [DiKl]. However, rather than working with functions from a base set to the real unit interval, we prefer to have fuzzy sets defined level-wise. Besides, all fuzzy sets will be assumed to be contained in one fixed bounded region of an $\mathbb{R}^p$.

Definition 1.1 *By a domain, we mean a regularly closed, convex, bounded subset of* $\mathbb{R}^p$ *for some* $p \geq 1$. *Let* Ω *be a domain, and let* $\mathcal{K}(\Omega)$ *be the set of non-empty convex and closed subsets of* Ω. *We partially order* $\mathcal{K}(\Omega)$ *by the subset relation* $\subseteq$*; and we endow* $\mathcal{K}(\Omega)$ *with the topology induced by the Hausdorff metric* d_H.

Let $[0,1]$ *be the real unit interval, endowed with the natural ordering and usual topology. A fuzzy vector in* Ω *is then meant to be a decreasing and left-continuous function* $v\colon\ [0,1] \to \mathcal{K}(\Omega)$*; by* $\mathcal{F}(\Omega)$*, we denote the set of all fuzzy vectors in* Ω. *We partially order* $\mathcal{F}(\Omega)$ *pointwise; and we endow* $\mathcal{F}(\Omega)$ *with the metric* $d(v,w) = \sup_{\alpha \in [0,1]} d_H(v(\alpha), w(\alpha)),\ \ v, w \in \mathcal{F}(\Omega)$.

Clearly, $\mathcal{F}(\Omega)$ may be identified with those elements of $\mathcal{E}^n$ from [DiKl] whose support is within the domain Ω [DiKl, Proposition 6.1.6/7].

We will embed $\mathcal{F}(\Omega)$ in the usual way into a function space; see e.g. [DiKl]. In what follows, S^{p-1} denotes the unit sphere of $\mathbb{R}^p$, and $(\cdot,\cdot)$ is the usual scalar product in $\mathbb{R}^p$; $p \geq 1$.

Definition 1.2 *Let* $\Omega \subseteq \mathbb{R}^p$ *a domain, and let* $v\colon\ [0,1] \to \mathcal{K}(\Omega)$ *be a fuzzy vector in* Ω. *We call*

$$s_v\colon\ [0,1] \times S^{n-1} \to \mathbb{R},\quad (\alpha, e) \mapsto \sup\,\{(r,e)\colon\ r \in v(\alpha)\}$$

the support function of v.

Moreover, let $L_\infty([0,1] \times S^{n-1})$ *be the linear space of bounded real-valued functions on* $[0,1] \times S^{n-1}$*, endowed with the supremum norm. Let* $L_\infty([0,1] \times S^{n-1})$ *be pointwise partially ordered.*

Proposition 1.3 *Let* $\Omega \subseteq \mathbb{R}^p$ *be a domain. Then the mapping* $\mathcal{F}(\Omega) \to L_\infty([0,1] \times S^{n-1}),\ \ v \mapsto s_v$ *is injective, isometric, and order preserving.*

Moreover, $s \in L_\infty([0,1] \times S^{n-1})$ *is the support function of some fuzzy vector* $v \in \mathcal{F}(\Omega)$ *iff (i) for all* $\alpha \in [0,1]$, $e, e_1, e_2 \in S^{n-1}$, $\lambda_1, \lambda_2 \geq 0$ *such that* $e = \lambda_1 e_1 + \lambda_2 e_2$*, we have* $s(\alpha, e) \leq \lambda_1 s(\alpha, e_1) + \lambda_2 s(\alpha, e_2)$*, (ii) for all* $e \in S^{n-1}$, $s(\cdot, e)$ *is decreasing and left-continuous, and (iii)* $v \leq s_\Omega$.

Note that under (iii), we considered Ω as a crisp element of $\mathcal{F}(\Omega)$ in the usual way.

By Proposition 1.3, we may identify $\mathcal{F}(\Omega)$ with a closed subset of the Banach function space $L_\infty([0,1] \times S^{n-1})$.

We next have to specify which functions between sets of fuzzy sets are taken into account for interpolation.

Definition 1.4 *Let* $\Xi \subseteq \mathbb{R}^m$ *and* $\Upsilon \subseteq \mathbb{R}^n$ *be domains, and let* $\mathcal{X} \subseteq \mathcal{F}(\Xi)$. *Then a function* $f\colon\ \mathcal{X} \to \mathcal{F}(\Upsilon)$ *is called fuzzy if* f *may be extended to a function* $\bar{f}\colon\ \mathcal{F}(\Xi) \to \mathcal{F}(\Upsilon)$ *which preserves the order, that is, for which* $\bar{f}(v) \leq \bar{f}(w)$ *whenever* $v \leq w$ *for* $v, w \in \mathcal{F}(\Xi)$.

In case that $\mathcal{X}$ *is finite, we call* f *a fuzzy if-then rule base.*

It is now possible for formulate the aims towards which we are working. Assume first that our data is crisp, that is, that we are given finitely many pairs from a bounded, closed subset Ω of $\mathbb{R}^m$ and from $\mathbb{R}^n$. Then a degree of smoothness of a function $\Omega \to \mathbb{R}^n$ is given by the integral of the norm of the second derivative over Ω; under certain assumptions, this value exists and we are lead to a uniquely determined interpolating function.

Now, in the fuzzy case, we wish to calculate functions from $\mathcal{F}(\Xi)$, viewed as a subset of the Banach space $L_\infty([0,1] \times S^{m-1})$, to $\mathcal{F}(\Upsilon)$, viewed as a subset of $L_\infty([0,1] \times S^{n-1})$. Proceeding analogously to the crisp case leads apparently to difficult requirements. First of all, an interpolating fuzzy function must possess its second Fréchet derivative; second, its norm must be integrable with respect to some measure on $\mathcal{F}(\Xi)$; third, this measure should be a metrically invariant one.

Unfortunately, this program fails: a metrically invariant measure does not exist on $\mathcal{F}(\Xi)$. A way out of this first difficulty is to restrict $\mathcal{F}(\Xi)$ in a way such that a measure of this kind does exist.

On the other hand, the program simplifies dramatically if we replace the spaces $L_\infty([0,1] \times S^{p-1})$ by finite-dimensional ones – simply by restricting the function domain to a finite subset. This is how we proceed in Section 3.

2 Known approaches

Like in many areas of the theory of fuzzy sets, also in the present one a quite active research is to be noted. We would like to mention three directions – those we know about. It is (i) the logical approach; (ii) the usage of fuzzy relations; (iii) interpolation based on linearity notions. Let a fuzzy if-then rule base $(u_1, v_1), \ldots, (u_k, v_k)$ of pairs from $\mathcal{F}(\Xi) \times \mathcal{F}(\Upsilon)$ for domains Ξ and Υ be given, and let us in this section adopt the usual notion of a fuzzy set as a function from a base set to $[0,1]$.

(i) First of all, an entry (u_i, v_i) of the rule base may be considered as a proposition like "if u_i then v_i" and may be formalised on the base of a logical calculus. For instance, some version of Hájek's Basic Predicate Logic [Haj] may be used. The difference to our setting is easily stated: In the logical framework, we investigate what is expressed by the rule base as it stands, not taking into account what is not derivable. Put into the language of logics, we may say that it is our aim to properly *extend* a given rule base, that is, adding statements which are not part of the information provided by the rules. To this end, we work so-to-say "horizontally" – by using features of the base set –, and not "vertically" – by considering various ways of how to connect truth values. – However, we would like to mention the work of Novák, see e.g. [Nov], where the "horizontal" viewpoint is used also in logics; namely, certain logical connectives are defined which do refer to the structure of the base set.

(ii) Concerning the second point, it is clear that there are in principle lots of possibilities of what to require from a function $f\colon\ \mathcal{F}(\Xi) \to \mathcal{F}(\Upsilon)$ such that

$f(u_1) = v_1, \ldots, f(u_k) = v_k$. A rather popular condition is that f is induced by a fuzzy relation between Ξ and Υ; this means in the simplest case that there is a fuzzy relation $R\colon\ \Xi \times \Upsilon \to [0,1]$ such that $f_R(u_i) = v_i$ for all i, where $f_R(u)(y) = \sup_{x \in \Xi} (u(x) \wedge R(x,y))$ for $u \in \mathcal{F}(\Xi)$ and $y \in \Upsilon$. It is clear that this requirement dramatically restricts the possible choices of f. In particular, f is then already determined by the fuzzy singletons, i.e. those fuzzy sets having a one-point support, because f preserves all suprema. Another point is the fact that a fuzzy relation R such that f_R maps the u_i to v_i may not exist at all; conditions are listed e.g. in [Got].

(iii) In connection to the third line of research, we should mention the well-known work of Kóczy and Hirota at the first place; see e.g. [KoHi]. His and several others' work comes closest to what we have in mind. For a review of methods developed so far, the article [Jen] can be recommended. Let us mention the idea behind the method which also is contained in [Jen]; it deals with rule bases $(u_1, v_1), \ldots, (u_k, v_k)$ of pairs of convex fuzzy sets over $\mathbb{R}$. To calculate the image $f(u)$ of some value u in the domain under the interpolating function f, one has first to determine two "neighbouring" entries u_1 and u_2; then, roughly spoken, $f(u)$ is constructed in the analogous way from v_1 and v_2 as u may be constructed from u_1 and u_2. – This approach as well as comparable ones are based on a clear geometric intuition, and their common advantage is that their technical realization is straightforward. A disadvantage is that their applicability is usually restricted; it seems that ill-behaved cases can mostly not be excluded. Besides, the transition to fuzzy sets of higher dimensions, if possible at all, requires often new ideas, like also in [JeKlKo], the paper subsequent to [Jen]. Finally, when we have a look what happens in the case that our data are crisp, we realize that we have to do with the easiest possible method; in the one-dimensional case, the methods reduce to the linear interpolation between neighbouring points.

3 Interpolation of crisp data – splines

When trying to make use of methods which have been developed for the interpolation of crisp data, we have to overcome the problem that we no longer have to do with finite-dimensional spaces: even $\mathcal{F}(\Omega)$, where $\Omega \subseteq \mathbb{R}$, embeds into an infinite-dimensional space. So there are in principle two possible ways: either, we generalise the existing methods to the infinite-dimensional case; or we reduce our spaces to finite-dimensional ones. In this paper, we shall describe a method according to the second way.

We first have to review the crisp case; so assume that we are given pairs $(x_1, y_1) \ldots, (x_k, y_k) \in \mathbb{R}^m \times \mathbb{R}^n$ and that we want to determine the smoothest possible function $f\colon\ \mathbb{R}^m \to \mathbb{R}^n$ which interpolates the given data. This problem is made precise, and has its unique solution, according to the following method, developed in the multidimensional case first by Atteia [Att]. We shall use the following variant of it.

Definition 3.1 *Let $m, n \geq 1$, and let $\Omega \subseteq \mathbb{R}^m$ be bounded and regularly closed. Then set*

$$\begin{aligned} F \stackrel{\text{def}}{=} \{f\colon\ \Omega \to \mathbb{R}^n :\ & f \textit{ is continuous and of class } \mathcal{C}^1, \\ & Df \textit{ is bounded and Lipschitz continuous,} \\ & D^2 f \textit{ is } L^1\}; \end{aligned}$$

let $||f||_\infty = \sup_{x \in \Omega} f(x)$, $||Df||_\infty = \sup_{x \in \Omega} ||Df(x)||$, *and* $||D^2 f||_1 = \int_\Omega ||D^2 f(x)|| dx$*; endow F with the norm*

$$||f|| = ||f||_\infty + ||Df||_\infty + ||D^2 f||_1, \ \ f \in F.$$

Here, the derivatives Df, $D^2 f$ are understood to have $\overset{o}{\Omega}$ as their domains.

F, together with $|| \cdot ||$, is a Banach space. The key fact which we need is now the following [Att, Hol].

Theorem 3.2 *Let $m, n \geq 1$, and let $\Omega \subseteq \mathbb{R}^m$ be bounded and regularly closed. Let K be a non-empty closed, convex subset of F; and for $f \in F$, set*

$$m(f) = \int_\Omega ||D^2 f(x)|| dx.$$

Let $N = \{f \in F\colon\ m(f) = 0\}$, and $C_K = \{f \in F\colon\ f + K = K\}$. Assume that N is finite-dimensional and that $N \cap C_K = \{0\}$. Then there is a unique $f \in K$ minimising $m(f)$.

With respect to the notation of this theorem, we see that if $x_1, \ldots, x_k \in \Omega$ such that there is no non-trivial affine function mapping all x_i to 0, and $y_1, \ldots y_k \in \mathbb{R}^n$, we may set $K = \{f \in F\colon\ f(x_1) = y_1, \ldots, f(x_k) = y_k\}$ to conclude that there is a uniquely determined function $f \in K$ for which the integral over the norm of the second derivative is smallest.

4 Interpolation of fuzzy data

Assume now that we have to solve a problem analogous to the one of the last section, with the difference that both the data from the domain and from the range are fuzzy vectors. Our fuzzy data is, according to Section 1, represented by real-valued functions on $[0, 1] \times S^{n-1}$. To reduce the interpolation problem to a finite-dimensional one, we shall, according to the pragmatic approach announced, approximate these functions by their values in finitely many points.

Accordingly, we work with the following structure, modeled upon Definition 1.2 and Proposition 1.3.

Definition 4.1 *Let $\Omega \subseteq \mathbb{R}^p$ be a domain, let I_f be a finite subset of $[0, 1]$ and S_f a finite subset of S^{p-1}. We call a function $s\colon\ I_f \times S_f \to \mathbb{R}$ an approximative* *support function w.r.t.* Ω, I_f, S_f *if (i) for all* $e \in S_f$, *there is an* $x \in \mathbb{R}^p$ *such that* $(x, e) = s(e)$ *and* $(x, f) \leq s(f)$ *for all* $f \in S_f \setminus \{e\}$*; (ii) for all* $e \in S_f$, $s(\cdot, e)$ *is decreasing; and (iii)* $s \leq s_\Omega|_{I_f \times S_f}$.

Moreover, let $L(I_f \times S_f)$ *be the space of real-valued functions on* $I_f \times S_f$, *endowed with the supremum norm and the pointwise order.*

In an analogous way, also the notion of an interpolating function and of a fuzzy if-then rule base is adapted from Definition 1.4.

Let us fix domains $\Xi \subseteq \mathbb{R}^m$ and $\Upsilon \subseteq \mathbb{R}^n$, and finite subsets $S_d \subseteq S^{m-1}$, $S_r \subseteq S^{n-1}$, $I_f \subseteq [0, 1]$.

Definition 4.2 *Let* $\mathcal{D}$ *be the set of approximative support functions w.r.t* Ξ, I_f, S_d, *and* $\mathcal{R}$ *those w.r.t.* Υ, I_f, S_r. *Set*

$$\begin{aligned} F \stackrel{\text{def}}{=} \{f\colon \mathcal{D} \to \mathcal{R} :\ & f \text{ is continuous and of class } \mathcal{C}^1, \\ & Df \text{ is bounded and Lipschitz continuous,} \\ & D^2 f \text{ is } L^1,\ \ f \text{ preserves the order}\}. \end{aligned}$$

For a subset $\mathcal{X}$ *of* $\mathcal{D}$, *a function* $f\colon \mathcal{X} \to \mathcal{R}$ *is called fuzzy approximative if there is an* $\bar{f} \in F$ *extending* f.

In case $\mathcal{X}$ *if finite, we call* f *an approximative fuzzy if-then rule base.*

We may now apply Theorem 3.2.

Theorem 4.3 *Let* $(u_1, v_1), \ldots, (u_k, v_k) \in \mathcal{D} \times \mathcal{R}$ *be an approximative fuzzy if-then rule base such that there is no non-trivial affine function* $\mathcal{D} \to L(I_f \times S_r)$ *which maps* $u_1, \ldots, u_k$ *to* 0. *Set*

$$m(f) = \int_{\mathcal{D}} ||D^2 f(x)||dx.$$

Then there is a unique $f\colon \mathcal{D} \to \mathcal{R}$ *such that* $f(u_1) = v_1, \ldots, f(u_k) = v_k$ *which minimises* $m(f)$.

Proof. The set K of all functions $f \in F$ such that $f(u_1) = v_1, \ldots, f(u_k) = v_k$ is convex and closed. Furthermore, by definition of an approximative fuzzy if-then rule base, K is not empty. So the claim follows from Theorem 3.2.

References

[Att] M. Atteia, Fonctions ‘spline’ definies sur un ensemble convexe, *Numer. Math.* **12** (1968), 192 - 210.

[DiKl] P. Diamond, P. Kloeden, “Metric spaces of fuzzy sets: theory and applications”, World Scientific, Singapore 1994.

[Got] S. Gottwald, Fuzzy relation equations and approximate solutions. Some recent results, Proceedings of 3rd EUSFLAT conference (Zittau 2003), 6 - 11.

[Haj] P. Hájek, "Metamathematics of fuzzy logic", Kluwer Acad. Publ., Dordrecht 1998.

[Hol] R. B. Holmes, "Geometric functional analysis and its applications", Springer-Verlag, New York and Berlin 1975.

[Jen] S. Jenei, Interpolation and extrapolation of fuzzy quantities revisited – an axiomatic approach, *Soft Comput.* **5** (2001), 179 - 193.

[JeKlKo] S. Jenei, E. P. Klement, R. Konzel, Interpolation and extrapolation of fuzzy quantities - the multiple-dimensional case, *Soft Comp.* **6** (2002), 258 - 270.

[KoHi] O. Kóczy, K. Hirota, Approximate reasoning by linear rule interpolation and general approximation, *Int. J. Approx. Reasoning* **9** (1993), 197 - 225.

[Nov] V. Novák, A horizon shifting method of linguistic hedges for approximate reasoning, Proceeding I of the 5th IEEE conference (New Orleans 1996), 423 - 427.