Spline interpolation between hyperspaces of convex or fuzzy sets

Thomas Vetterlein
Department of computer sciences 1
University of Dortmund
44221 Dortmund
Thomas.Vetterlein@uni-dortmund.de

January 2005 - revised May 2006

Abstract

We consider the interpolation problem for functions whose range and whose domain both consist of convex or fuzzy subsets of a real Euclidean space. This problem arises in fuzzy controlling, namely when the functional dependence between two fuzzy vectors is known only for finitely many cases. To have a criterion for an appropriate choice of an interpolation function, we generalise the well-known idea from spline interpolation: the function should be "as smooth as possible".

1 Introduction

We discuss in this paper ways of interpolation between hyperspaces of convex or fuzzy sets. The underlying base sets are bounded regions of Euclidean spaces, which may be of any (finite) dimension.

Our work is based on ideas developed in the Collaborative Research Centre SFB531-A1 in Dortmund; we may refer e.g. to H. Thiele's paper [18]. The intuitive background is the following. Let α and β be variables which take values either from the closed bounded convex subsets of some $\mathbb{R}^m$ resp. $\mathbb{R}^n$, or from the standard fuzzy sets (i.e. the normal, support-bounded, upper semicontinuous, and fuzzy-convex fuzzy sets) over the $\mathbb{R}^m$ resp. $\mathbb{R}^n$. As usual, α and β should be thought of as representing imprecise values, where the impreciseness may well be genuine or otherwise caused by insufficient information. We further assume that α uniquely determines β and that we know about this dependence only from finitely many cases. The problem is then how to reconstruct the dependence of β from α in other cases – if possible, in analogy to the interpolation of crisp data.

The sketched problem occurs for instance in fuzzy controlling, in connection with the construction of devices processing multidimensional fuzzy data. These devices are typically based on what is called a fuzzy if-then rule base, that is, a collection of pairs $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k) \in \mathcal{F}(\Omega) \times \mathcal{F}(\Psi)$, where $\mathcal{F}(\Omega)$ and $\mathcal{F}(\Psi)$ are universes of fuzzy sets over $\Omega \subseteq \mathbb{R}^m$ and $\Psi \subseteq \mathbb{R}^n$, respectively. The meaning of such a rule base is that if the input value is α_i, then the output value must be set to β_i; $i = 1, ..., k$. Since an actual input value certainly need not be among the finitely many ones occurring in the rule base, we have to extend the rule base to a total function from $\mathcal{F}(\Omega)$ to $\mathcal{F}(\Psi)$.

In fuzzy controlling, this problem has found practical solutions. According to a typical procedure, a fuzzy relation between Ω and Ψ is computed from the rule base, and the compositional rule of inference (CRI) is applied. This pragmatic procedure, which originates from [20], is in most cases satisfactory because fuzzy data occurs only internally; the sharp input data is first fuzzified, next the fuzzy data is processed under the CRI, and the resulting fuzzy data is finally defuzzified. Under these circumstances, the middle part, i.e. the function mapping $\mathcal{F}(\Omega)$ to $\mathcal{F}(\Psi)$, should not be viewed isolated; what counts is the whole three-step device mapping sharp input data to sharp output data.

In the present paper, we do not intend to question the value of this procedure, which has been used successfully in innumerably many applications. We do not even claim that our proposals can lead to a higher quality in fuzzy controlling; we simply would like to communicate our ideas and inspirit the general discussion.

We are led by the observation that the function which maps fuzzy values to fuzzy values on the base of a fuzzy relation and which may actually be considered as the key part of a controlling device, does not operate according to some predefined and easily acceptable criterion. We raise the question if we cannot apply some natural principle, determining a function which maps the input fuzzy values to output fuzzy values in a way that the most basic expectations are met; cf. [18].

This aim brings interpolation theory into play. Now, the theory of interpolation between Euclidean spaces has been very lively for many decades, and we wonder if we cannot generalise well-accepted methods of interpolation between “crisp” data to those between universes of convex sets or fuzzy data.

First to mention, in recent years the theory of radial basis functions (RBFs) has turned out to be a very general and flexible and the same time very practical tool for multivariate interpolation. Unfortunately, an adaptation of the theory of RBFs to our setting is far from being straightforward. The transition from reproducing kernel Hilbert spaces, which consist of real-valued functions, to spaces consisting of maps between hyperspaces is not at all trivial. But we refer to [10, 13] for work on RBFs. Here, we will rather do one step back in the younger history of interpolation theory. What we consider in this paper is the formalism of spline interpolation in the abstract form which was developed mainly by Atteia from the 60s on [3, 4]. The idea underlying spline interpolation is well-known: The interpolating function minimises the norm of its image under a certain differential operator, that is, a value which measures its smoothness.

2 Outline of our approach

Assume that we are given an if-then rule base $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k) \in \mathcal{C}(\Omega) \times \mathcal{C}(\Psi)$, where $\mathcal{C}(\Omega)$ and $\mathcal{C}(\Psi)$ contain the compact convex subsets of $\Omega \subseteq \mathbb{R}^m$ and $\Psi \subseteq \mathbb{R}^n$ respectively. Our question is how to single out a function with reasonable properties $f\colon\ \mathcal{C}(\Omega) \to \mathcal{C}(\Psi)$ such that $f(\alpha_1) = \beta_1, \ldots, f(\alpha_k) = \beta_k$. To this end, we are going to adapt the formalism of spline interpolation.

In a first step, the hyperspace $\mathcal{C}(\mathbb{R}^m)$ of compact convex subsets of $\mathbb{R}^m$ needs to be embedded into a Banach space. This is most easily done by identifying every set $\alpha \in \mathcal{C}(\mathbb{R}^m)$ with its support function

$$s_\alpha\colon\ S^{m-1} \to \mathbb{R}, \quad e \mapsto \max\{(x, e)\colon\ x \in \alpha\};$$

cf. e.g. [16]. This function has the following geometric meaning. For any unit vector e, there is a hyperplane H_e with normal vector e supporting the convex set α; this means that H_e touches, on the side opposite to the direction of e, the surface of α. Furthermore, the hyperplane H_e is uniquely determined by e and by the (signed) distance from the origin from H_e; this distance is $s_\alpha(e)$. In other words, the hyperplanes determined by e and $s_\alpha(e)$, where e varies over the unit vectors, enclose the convex set α and uniquely characterise α in this way.

So the hyperspace of compact convex sets is identified with certain real functions on S^{m-1}. Thus, we may embedded it, for instance, into the Banach space $C(S^{m-1})$ of continuous real functions on S^{m-1}, endowed with the supremum norm. The support functions of convex subsets form a positive cone in $C(S^{m-1})$; moreover, the addition and the multiplication with positive reals of support functions correspond to the same operations performed pointwise with the corresponding convex subsets. For further details on convex sets, we recommend [16].

The problem we encounter next is easily stated: The space $C(S^{m-1})$ is infinite-dimensional, and when considering functions between two such spaces, fundamental difficulties arise. For instance, there is no way to define a measure on $C(S^{m-1})$ or $\mathcal{C}(\mathbb{R}^m)$ with reasonable properties; measure on appropriate subsets of the hyperspace of compact convex sets is an own problem, to be treated in another paper; cf. [5].

In practice, however, we have anyhow only finitely many parameters to specify a convex set. So it does not seem to be too problematic to switch over to a finite dimensional space approximating $C(S^{m-1})$. This is most simply done by restricting the real functions on S^{m-1} to their values in finitely many points; and this is what we will do here. So we do what might have been expected: We restrict to (convex) polytopes with prescribed normal vectors of their bounding hyperplanes. Note that the set of all polytopes lies dense in the set of compact convex sets e.g. w.r.t. the Hausdorff metric.

So having fixed, let's say, r normal vectors, we will, in a first step, see how the set of support functions restricted to these vectors is to be characterised. Each such restricted support function is given by an r-tuple of real numbers; so what we have to describe is a subset of $\mathbb{R}^r$ containing those r-tuples which describe convex polytopes. The interpolation itself is then done between two such parameter subsets.

The method which we choose is spline interpolation, in the form due to Atteia [4]. Very roughly speaking, it works as follows. The space of functions used for the interpolation between two Euclidean spaces is a Sobolev space, which contains functions whose (weak) derivatives of q-th order are square integrable. Here, q must be chosen sufficiently large, depending on the number of interpolation knots. For each such function, the square integral over the norms of the q-th derivatives is taken as a measure of smoothness, and this value is minimised.

We shall furthermore generalise the procedure for fuzzy sets. This step is, from the theoretical point of view, easy. The crucial point is: We work with fuzzy sets in their level-wise representation. Recall that a fuzzy set is usually considered as a function u from an $\mathbb{R}^n$ to the real unit interval $[0, 1]$. However, when restricting to the normal, support-bounded, upper semicontinuous, and fuzzy-convex fuzzy sets, we may equivalently view fuzzy sets as functions from $[0, 1]$ to compact convex sets, by associating with each $\alpha \in [0, 1]$ the α-level set $[u]^\alpha$. See e.g. [7].

We will see that all statements about hyperspaces of convex sets may be generalised to spaces of fuzzy sets in a straightforward way. The argumentation just becomes technically slightly more involved.

3 Other approaches

Before beginning the mathematical part, let us discuss the relationship of our work to others'. We proceed from far-related to closely related approaches.

It must first be pointed out that our work is not comparable to fuzzy inference formalised in the framework of some formal system of many-valued logics. A rule base may well be understood as a set of logical implications and may be formalised, for instance, within some first-order version of Hájek's Basic Logic [8]. Let us stress the difference to our concerns: In the logical framework, we investigate what is expressed by a rule base as it stands, not being interested in 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.

It might also be suggestive to say that 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. We must add, however, that our "horizontal" viewpoint is not principally incompatible with concepts of logics; V. Novák and I. Perfilieva define in [15, 14] linguistic hedges which do refer to the structure of the base set.

Second, we would like to point out that there are numerous works concerned with the problem of interpolation of fuzzy data, where it is assumed that the domain consists of crisp values. For an overview, we refer to [7, Chapter 12.3]; for more recent work, we may mention e.g. [1].

Third, as mentioned above, a function mapping fuzzy data to fuzzy data is typically required to be induced by a fuzzy relation. It is clear that this condition seriously 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 all suprema are preserved.

When working with functions induced by fuzzy relations, two problems naturally arise. There might be no function interpolating the entries of the rule base, and to find a proper alternative is not easy to determine. Moreover, if there is such a function, there are in general many, and again it is not easy to find arguments justifying a canonical choice.

Here, we do not work with fuzzy relations, but we remark the following. We may certainly make use of the approach presented in this article *together* with the requirement that the interpolating function is based on a fuzzy relation. The latter may simply be chosen as smooth as possible and modelling the rule base as well as possible. This possibility will be discussed elsewhere.

Finally, a further series of works comes quite close to our ideas. We have in mind the well-known work of L. T. Kóczy and K. Hirota [12], which has been further elaborated in numerous articles; we may refer e.g. to [6, 19]. A further approach is due to S. Jenei [11]. In view of the present article, the methods along these lines are creditable by the fact that they are guided by applicational needs and that they are technically comparably easily realizable. Our own work is notable for the fact that we are guided by the idea that the interpolating function should fulfil a certain criterion; although a "smooth" mathematical theory is the result, it is difficult to realise it in practice.

The quality of the approach presented here is probably best appreciated by considering the case that the rule base contains one-dimensional crisp data. The smoothness requirement then leads to results comparable to crisp spline interpolation. As opposed to that, many methods defined in the past just reduce in this case to the linear interpolation between neighbouring points.

4 Parametrising convex sets

For the interpolation between hyperspaces of compact convex sets, it is hardly possible to work with these hyperspaces as a whole: we have to specify a strict, but still reasonably large subspace. Here, we will restrict to those (convex) polytopes the normal vectors of the facets of which are among a fixed finite set of unit vectors. This is a standard approach; cf. e.g. [17].

Let us fix in this section a dimension $n \geq 1$. We will follow [16] in terminology and notation. In particular, we will denote by $H^-_{e,a} = \{x \in \mathbb{R}^n \colon\ (x,e) \leq a\}$ the closed halfspace specified by the unit vector e and the real number a; similarly, $H^+_{e,a} = \{x \in \mathbb{R}^n \colon\ (x,e) \geq a\}$; both these halfspaces are bounded by the hyperplane $H_{e,a} = \{x \in \mathbb{R}^n \colon\ (x,e) = a\}$.

Definition 4.1 A finite set $\mathcal{E} = \{e_1, \ldots, e_r\}$ of unit vectors of the $\mathbb{R}^n$ is called a *direction set* if the positive hull of the $e_1, \ldots, e_r$ is the whole R^n. By an *$\mathcal{E}$-polytope* we then mean the bounded intersection of halfspaces $H^-_{e_1,a_1}, \ldots, H^-_{e_r,a_r}$, where $a_1, \ldots, a_r \in \mathbb{R}$.

Moreover, let $\alpha = \bigcap_{e \in \mathcal{E}} H^-_{e,s(e)}$ be an *$\mathcal{E}$-polytope*, and let $R \in \mathbb{R}^+$. Then α is called *R-bounded* if $|s(e)| \leq R$ for every $e \in \mathcal{E}$.

The representation of an $\mathcal{E}$-polytope by an intersection of halfspaces $H^-_{e,s(e)}$, $e \in \mathcal{E}$, is not unique. We get uniqueness by requiring that every parameter $s(e)$ is the smallest possible one.

Definition 4.2 Let $\mathcal{E}$ be a direction set of the R^n. We then call a mapping $s \colon\ \mathcal{E} \to \mathbb{R}$ an *$\mathcal{E}$-support function* if (i) $\bigcap_{e \in \mathcal{E}} H^-_{e,s(e)}$ is non-empty and (ii) $\bigcap_{e \in \mathcal{E}, e \neq f} H^-_{e,s(e)} \cap H^-_{f,s(f)-\varepsilon} \subsetneq \bigcap_{e \in \mathcal{E}} H^-_{e,s(e)}$ for any $f \in \mathcal{E}$ and $\varepsilon > 0$.

In other words, $\mathcal{E}$ being a direction set, an $\mathcal{E}$-support function $s\colon\ \mathcal{E} \to \mathbb{R}$ is simply the support function of some polytope α restricted to $\mathcal{E}$. Namely, the associated $\mathcal{E}$-polytope is $\alpha = \bigcap_{e\in\mathcal{E}} H^-_{e,s(e)}$. Note that we established a one-to-one correspondence between the $\mathcal{E}$-support functions and the $\mathcal{E}$-polytopes.

Given some fixed direction set $\mathcal{E}$, we next have to characterise the set of all $\mathcal{E}$-support functions algebraically. Recall that support functions of convex sets, which are defined on the whole unit sphere S^{n-1}, are exactly the restrictions of positive homogeneous and subadditive functions to S^{n-1}; see e.g. [16, Thm 1.7.1]. However, we could not find this fact useful to derive the following proposition.

Proposition 4.3 *Let $\mathcal{E}$ be a direction set of the R^n. Then $s\colon\ \mathcal{E} \to \mathbb{R}$ is an $\mathcal{E}$-support function if and only if the following condition holds:*

(P) *Let $e_1, \ldots, e_n \in \mathcal{E}$ be a basis of $\mathbb{R}^n$ and let p be the point in the intersection of the hyperplanes $H_{e_1,s(e_1)}, \ldots, H_{e_n,s(e_n)}$; let $e \in \mathcal{E}$. If then e is a positive combination of $e_1, \ldots, e_n$, we have $p \in H^+_{e,s(e)}$; if $-e$ is a positive combination of $e_1, \ldots, e_n$, we have $p \in H^-_{e,s(e)}$.*

Proof. Assume that s is an $\mathcal{E}$-support function. Then $\alpha = \bigcap_{e\in\mathcal{E}} H^-_{e,s(e)}$ is a polytope, and every $H_{e,s(e)}$ is a support hyperplane of α.

Let $e_1, \ldots, e_n$ be a basis of $\mathbb{R}^n$, and let p be the point in the intersection of the hyperplanes $H_{e_1,s(e_1)}, \ldots, H_{e_n,s(e_n)}$. Then $\gamma = H^-_{e_1,s(e_1)} \cap \ldots \cap H^-_{e_n,s(e_n)}$ is the polyhedral set generated by the point p and by the rays $\cap_{j\neq i} H_{e_i,s(e_i)} \cap H^-_{e_j,s(e_j)}$, $j = 1, ..., n$, starting at p. If now e is a positive combination of $e_1, \ldots, e_n$, there is a support hyperplane $H_{e,a}$ of γ at p; then $\alpha \subseteq \gamma \subseteq H^-_{e,a}$. On the other hand, $H_{e,s(e)}$ supports α, and it follows $H^-_{e,s(e)} \subseteq H^-_{e,a}$. So $s(e) \leq a$ and $p \in H^+_{e,s(e)}$.

If $-e$ is a positive combination of $e_1, \ldots, e_n$, we may similarly conclude $p \in H^-_{e,s(e)}$. So (P) is proved.

Assume now (P). Let $e_1, \ldots, e_k \in \mathcal{E}$ be such that (i) $e_1, \ldots, e_k$ span $\mathbb{R}^n$ and (ii) the polyhedron $\alpha' = H^-_{e_1,s(e_1)} \cap \ldots \cap H^-_{e_k,s(e_k)}$ is non-empty. Note that for $k = n$, this condition always holds. Now, let $e \in \mathcal{E}$ be distinct from $e_1, \ldots, e_k$. We shall show that then both $H^+_{e,s(e)} \cap \alpha' \neq \emptyset$ and

$H^-_{e,s(e)} \cap \alpha' \neq \emptyset$. It will then follow by induction, starting with $k = n$ and ending with the number of elements of $\mathcal{E}$, that s is an $\mathcal{E}$-support function.

If α' contains a ray with direction f such that $(e, f) > 0$, then clearly $H^+_{e,s(e)} \cap \alpha' \neq \emptyset$. Let us assume the opposite case. Then the support set of α' belonging to e contains an extreme point p of α'. By [16, Lemma 2.3.13], we know that for any $\varepsilon > 0$ there are $b_1, \ldots, b_k \in \mathbb{R}$ such that $|b_1 - s(e_1)|, \ldots, |b_k - s(e_k)| \leq \varepsilon$, and $\alpha'' = H^-_{e_1,b_1} \cap \ldots \cap H^-_{e_k,b_k}$ is a polyhedron whose extreme points are contained in exactly n facets, and p is still an extreme point of α''. By [16, Thm 2.4.9], e is in the positive hull of the normal vectors of the facets to which p belongs, that is, e is a positive combination of n vectors among $e_1, \ldots, e_k$, let's say of the first n vectors $e_1, \ldots, e_n$. So p is the point in $H_{e_1,b_1} \cap \ldots \cap H_{e_n,b_n}$ and because ε was arbitrary, p is also contained in $H_{e_1,s(e_1)} \cap \ldots \cap H_{e_n,s(e_n)}$. It follows by assumption that $p \in H^+_{e,s(e)}$, and so $H^+_{e,s(e)} \cap \alpha' \neq \emptyset$. Similarly, we conclude that $H^-_{e,s(e)} \cap \alpha' \neq \emptyset$. □

For some fixed r-element direction set $\mathcal{E}$, we may consider the collection of $\mathcal{E}$-support functions as a subset of the $\mathbb{R}^r$ in the straightforward way.

Definition 4.4 Let $\mathcal{E}$ be an ordered direction set, that is, a direction set endowed with a fixed order of its elements. Let $e_1, \ldots, e_r$ be the elements of $\mathcal{E}$ in this order; we then define the *polytope parameter space* associated to $\mathcal{E}$ as the set $\mathcal{P}_{\mathcal{E}}$ of all $p = (a_1, \ldots, a_r) \in \mathbb{R}^r$ such that

$$s_p \colon \ \mathcal{E} \to \mathbb{R}, \ \ e_1 \mapsto a_1, \ldots, e_r \mapsto a_r$$

is an $\mathcal{E}$-support function.

Moreover, for some $R > 0$, we define the polytope parameter set associated to $\mathcal{E}$ and R as the set $\mathcal{P}_{\mathcal{E},R}$ of all $(a_1, \ldots, a_r) \in \mathcal{P}_{\mathcal{E}}$ such that $|a_1|, ..., |a_r| \leq R$.

Since $\mathcal{E}$-support functions are in a one-to-one correspondence with $\mathcal{E}$-polytopes, we have thus identified the collection of all $\mathcal{E}$-polytopes with a certain subset of the $\mathbb{R}^r$ – namely with $\mathcal{P}_{\mathcal{E}}$, the polytope parameter space associated to $\mathcal{E}$. Moreover, we have identified the R-bounded $\mathcal{E}$-polytypes with $\mathcal{P}_{\mathcal{E},R}$.

It is now of central importance to see which shape $\mathcal{P}_{\mathcal{E}}$ has.

Theorem 4.5 *Let $\mathcal{E}$ be an ordered direction set with r elements and let $\mathcal{P}_{\mathcal{E}}$ the polytope parameter space associated to $\mathcal{E}$. Then $\mathcal{P}_{\mathcal{E}}$ is the intersection of halfspaces of $\mathbb{R}^r$ which contain the origin, and $\mathcal{P}_{\mathcal{E}}$ has non-empty interior.*

In particular, for any $R > 0$, $\mathcal{P}_{\mathcal{E},R}$ is a polytope in $\mathbb{R}^r$ with non-empty interior.

Proof. Let $s\colon\ \mathcal{E} \to \mathbb{R}$; we will use condition (P) of Proposition 4.3 to formulate exact conditions for s to be an $\mathcal{E}$-support function. In what follows, we denote by $v^1, \dots, v^n$ the coefficients of some $v \in \mathbb{R}^n$ w.r.t. the canonical basis of $\mathbb{R}^n$.

Let $e, e_1, \dots, e_n \in \mathcal{E}$ such that $e_1, \dots, e_n$ is a basis of $\mathbb{R}^n$ and e is distinct from $e_1, \dots, e_n$. Let $a = s(e)$, $a_1 = s(e_1), \dots, a_n = s(e_n)$. The point p in the intersection of the hyperplanes $H_{e_1,a_1}, \dots, H_{e_n,a_n}$ fulfils $(p, e_1) = a_1, \dots, (p, e_n) = a_n$; so

$$p = \begin{pmatrix} e_1^1 & \dots & e_1^n \\ \vdots & & \vdots \\ e_n^1 & \dots & e_n^n \end{pmatrix}^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

So we see that the condition $(p, e) \le a$ or $(p, e) \ge a$ may be written in the form

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \\ a \end{pmatrix} \begin{pmatrix} q_1 \\ \vdots \\ q_n \\ q \end{pmatrix} \le 0$$

for properly chosen $q_1, \dots, q_n \in \mathbb{R}, q \in \{-1, 1\}$. In view of (P), it follows that s is an $\mathcal{E}$-support function iff the (ordered) range of s lies in the intersection of certain halfspaces of the $\mathbb{R}^r$ through the origin.

That the intersection of all these halfspaces has full dimension follows from the fact that there is an $s\colon\ \mathcal{E} \to \mathbb{R}$ and an $\varepsilon > 0$ such that all $s'\colon\ \mathcal{E} \to \mathbb{R}$ fulfilling $|s'(e) - s(e)| \le \varepsilon$ for any $e \in \mathcal{E}$ are $\mathcal{E}$-support functions; see e.g. [16, Lemma 2.4.12].

Finally, let $R > 0$. Then $\mathcal{P}_{\mathcal{E},R}$ is the intersection of $\mathcal{P}_{\mathcal{E}}$ with the cube $\{p \in \mathbb{R}^r\colon\ ||p||_\infty \le R\}$; this is clearly a polytope. $\square$

We conclude the section with one more remark. The parametrisation of polytopes which we use is natural mainly because in our parameter space $\mathcal{P}_\mathcal{E}$, the addition and multiplication with positive reals coincides with the same operations performed pointwise with the corresponding polytopes. For, we have to do with the restrictions of support functions; and the linear operations performed pointwise with support functions correspond to the respective operations performed pointwise with convex sets. For this topic, we refer once again to [16].

Moreover, polytopes are partially ordered by inclusion; and also this partial order is compatible with the structure of the parameter space. We state this fact explicitely.

Definition 4.6 Let $\mathcal{P}_\mathcal{E}$ be the polytope parameter space associated to some ordered direction set $\mathcal{E}$. Then we endow $\mathcal{P}_\mathcal{E}$ as well as $\mathcal{P}_{\mathcal{E},R}$ for any $R > 0$, with the pointwise natural order.

Proposition 4.7 *The order of a polytope parameter space and the inclusion relation between the corresponding polytopes coincide.*

Note that a bounded parameter space $\mathcal{P}_{\mathcal{E},R}$, $R > 0$, has a largest element.

5 Parametrising fuzzy sets

Having restricted the hyperspace of compact convex sets to a hyperspace of certain polytopes, we shall continue along the same lines to define a reasonably wide, but still manageable space of fuzzy sets. Again, we fix some dimension $n \geq 1$.

In the present context, it would clearly be impractical to view fuzzy sets as functions from the base set to the real unit interval. We will adapt the "level-set viewpoint" instead: By a fuzzy set, we mean a bounded and decreasing (w.r.t. $\subseteq$) and left-continuous (w.r.t. Hausdorff metric) function $a\colon (0,1] \to \mathcal{C}(\mathbb{R}^n)$. But we do not need this general notion; in accordance with the last section, we will restrict to finitely many levels only.

Definition 5.1 Let $\mathcal{E}$ be a direction set, and let Δ be a finite subset of the real unit interval $[0, 1]$. Let Δ be endowed with the natural order, and let the set of $\mathcal{E}$-polytopes be partially ordered by the inclusion relation. By a $\mathcal{E}, \Delta$-fuzzy set, we then mean a decreasing function a from Δ to the set of $\mathcal{E}$-polytopes.

Moreover, for some $R > 0$, a $\mathcal{E}, \Delta$-fuzzy set is called R-bounded if all $\mathcal{E}$-polytope in its range are R-bounded.

These are the objects under study. Note that $\mathcal{E}$, the direction set, has the meaning explained in the Section 4; it contains the directions in which we may bound convex sets. Moreover, the finite set Δ of reals between 0 and 1 is new; each element in Δ corresponds to one of the levels of the fuzzy sets which we want to parametrise. Both for $\mathcal{E}$ and for Δ, we have to make a fixed choice.

Like in the case of the polytopes, we will identify the $\mathcal{E}, \Delta$-fuzzy sets with elements of an appropriate parameter space.

Definition 5.2 Let $\mathcal{E}$ be an ordered direction set with r elements, and let $\Delta = \{\lambda_1, \dots, \lambda_d\} \subseteq [0, 1]$. We then define the *fuzzy set parameter space* associated to $\mathcal{E}$ and Δ as the set $\mathcal{F}_{\mathcal{E},\Delta}$ of all d-tuples $(p_1, \dots, p_d) \in \mathcal{P_E}^d$ such that the function mapping λ_i, $i = 1, ..., d$, to the $\mathcal{E}$-polytope represented by p_i is an $\mathcal{E}, \Delta$-fuzzy set.

Moreover, for some $R > 0$, we define the fuzzy set parameter space associated to $\mathcal{E}$, Δ, and R as the set $\mathcal{F}_{\mathcal{E},\Delta,R}$ of all d-tuples $(p_1, \dots, p_d) \in \mathcal{F}_{\mathcal{E},\Delta}$ such that p_1 represents an R-bounded $\mathcal{E}$-polytope.

Consisting of d-tuples of r-tuples, we will treat $\mathcal{F}_{\mathcal{E},\Delta}$ as a subset of $\mathbb{R}^{dr}$.

Theorem 5.3 *Let $\mathcal{E}$ be an ordered direction set with r elements, and let $\Delta = \{\lambda_1, \dots, \lambda_d\} \subseteq [0, 1]$. Let $\mathcal{F}_{\mathcal{E},\Delta}$ be the fuzzy set parameter space associated to $\mathcal{E}$ and Δ. Then $\mathcal{F}_{\mathcal{E},\Delta}$ is the intersection of halfspaces of $\mathbb{R}^{dr}$ which contain the origin, and $\mathcal{F}_{\mathcal{E},\Delta}$ has non-empty interior.*

In particular, for any $R > 0$, $\mathcal{F}_{\mathcal{E},\Delta,R}$ is a polytope in $\mathbb{R}^{dr}$ with non-empty interior.

Proof. $\mathcal{F}_{\mathcal{E},\Delta}$ is a subset of the cardinal product of polytope parameter spaces, which are polytopes. Furthermore, the conditions determining this subset are inequalities between certain pairs of components; so $\mathcal{F}_{\mathcal{E},\Delta}$ is the meet of two polytopes, which gives again a polytope. This shows the first part; the second one follows similarly. $\square$

Like in the case of convex set, our parameter set reflects appropriately addition, multiplication with positive reals and the partial order of the corresponding fuzzy sets. The last point gives rise to the following definition.

Definition 5.4 Let $\mathcal{F}_{\mathcal{E},\Delta}$ be the fuzzy set parameter space associated to some ordered direction set $\mathcal{E}$ and $\Delta \subseteq [0,1]$. Then we endow $\mathcal{F}_{\mathcal{E},\Delta}$ as well as $\mathcal{F}_{\mathcal{E},\Delta,R}$ with the pointwise natural order.

Proposition 5.5 *The order of the fuzzy set parameter space and the inclusion relation between the corresponding fuzzy sets coincide.*

6 Monotone functions between sets of polytopes

Before turning to the interpolation problem itself, we shall take care of one specific property which an interpolating function must have. In the present context, convex subsets or fuzzy sets are meant to represent propositions in a way that the larger the subset or fuzzy set is, the weaker is the represented proposition. Furthermore, a pair (α, β) of a rule base is meant to model an implication of the form "if the input value is α, then the output value is β", and the same interpretation is assumed about any extension of the rule base. Hence it is natural to require that a rule base and any extension of a rule base to a total function should be monotone: it should preserve the (subset or fuzzy) inclusion relation.

Definition 6.1 Let $\mathcal{P}_{\mathcal{D}}$ and $\mathcal{P}_{\mathcal{E}}$ be the polytope parameter spaces associated to some ordered direction sets $\mathcal{D}$ and $\mathcal{E}$, respectively. By a *rule* w.r.t. $\mathcal{P}_{\mathcal{D}}$ and $\mathcal{P}_{\mathcal{E}}$, we mean a pair $(\alpha, \beta) \in \mathcal{P}_{\mathcal{D}} \times \mathcal{P}_{\mathcal{E}}$; a finite set of rules is called a *rule base*. A rule base $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k)$ is called *monotone* if $\alpha_i \leq \alpha_j$ implies $\beta_i \leq \beta_j$ for any i, j.

A rule base respecting the inclusion relation can not only be extended to a monotone total function; by mollification, even any smoothness requirement can be fulfilled. In the sequel, we denote the ε-neighbourhood of some point or some subset of $\mathbb{R}^r$ by $U_\varepsilon(\cdot)$.

Lemma 6.2 *Let $\mathcal{P}_{\mathcal{D},R}$, $\mathcal{P}_{\mathcal{E},S}$ be the polytope parameter spaces associated to the ordered direction sets $\mathcal{D}$, $\mathcal{E}$ and positive reals R, S, respectively. Let $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k) \in \mathcal{P}_{\mathcal{D},R} \times \mathcal{P}_{\mathcal{E},S}$ a monotone rule base. Then there is an analytic monotone function $f: \mathcal{P}_{\mathcal{D},R} \to \mathcal{P}_{\mathcal{E},S}$ such that $f(\alpha_1) = \beta_1, \ldots, f(\alpha_k) = \beta_k$.*

Proof. Let $\varepsilon > 0$ be small enough that for any i, j and $\alpha_i' \in U_{2\varepsilon}(\alpha_i)$, we have $\alpha_i' \leq \alpha_j$ only in case $\alpha_i \leq \alpha_j$. Extend the pointwise partial order of $\mathcal{P}_{\mathcal{D},R}$ to $U_\varepsilon(\mathcal{P}_{\mathcal{D},R})$, and let

$$\begin{aligned} \hat{f} \ &: \ U_\varepsilon(\mathcal{P}_{\mathcal{D},R}) \to \mathcal{P}_{\mathcal{E},S}, \\ &\alpha \mapsto \min\,\{\beta_i\colon\ 1 \leq i \leq k \text{ and } \alpha' \leq \alpha_i \text{ for some } \alpha' \in U_\varepsilon(\alpha)\}, \end{aligned}$$

where $\min\,\emptyset$ is understood to be the maximal element of $\mathcal{P}_{\mathcal{E},S}$. Then $\hat{f}$ is monotone, $\hat{f}(\alpha_i) = \beta_i$, and $\hat{f}$ is constant in the ε-neighbourhood of α_i; $i = 1, ..., k$.

Let now φ: $\mathbb{R}^n \to \mathbb{R}^+$ be an analytic function such that $\int_{\mathbb{R}^n} \varphi(x)dx = 1$ and the support of φ is in the ε-neighbourhood of the origin. Let f: $\mathcal{P}_{\mathcal{D},R} \to \mathcal{P}_{\mathcal{E},S}$, $\alpha \mapsto \int_{U_\varepsilon(0)} \hat{f}(\alpha - x)\varphi(x)dx$. Then f is still monotone, $f(\alpha_i) = \beta_i$ for any i, and f is analytic. □

We note that by Lemma 6.2, a function f extending a monotonous rule base may be chosen monotone, but this does not imply in general that f preserves lattice operations. By requiring only that the rule base is monotone, f need not even be extendible to a lattice homomorphism between the finite sublattices generated by the rule entries. To ensure the existence of an f which even preserves lattice operations, the stricter condition would be needed that if an infimum of left-side entries is below a supremum of other left-side entries, the same should hold for the corresponding right-side entries.

7 Spline interpolation

We finally turn to the interpolation problem itself. We shall first of all outline our procedure. Let us assume that we are given an if-then rule base $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k)$ consisting of pairs of convex compact subsets of the $\mathbb{R}^m$ and $\Psi \subseteq \mathbb{R}^n$, respectively. Then we proceed as follows:

- We fix a direction set $\mathcal{D}$ in the $\mathbb{R}^m$, and a direction set $\mathcal{E}$ in the $\mathbb{R}^n$. We identify each $\alpha_1, \ldots, \alpha_k$ with an element of the polytope parameter space associated to $\mathcal{D}$, i.e. with some element of $\mathcal{P}_\mathcal{D}$; and we identify each $\beta_1, \ldots, \beta_k$ with an element of $\mathcal{P}_\mathcal{E}$.

- We approach the interpolation problem as follows. We search for a function from the polytope parameter space $\mathcal{P}_\mathcal{D}$ to the polytope parameter space $\mathcal{P}_\mathcal{E}$, such that each α_i is mapped to β_i, $i = 1, ..., k$, under the above identification.

- As the criterion for how to choose the interpolating function, we adapt the concept of spline interpolation. To this end, we use a Sobolev space of functions $f\colon \mathcal{P}_\mathcal{D} \to \mathcal{P}_\mathcal{E}$; we define a function from this space to the positive reals; and we minimise the image under this function.

The case of fuzzy sets is analogous.

We recall next the formalism of spline interpolation in its general form. It is not possible to give here any detailed explanations; we have to refer to the literature. The formalism is due to Atteia [3, 4], and the following version is due to [9, §21B].

We summarise in a few words the meaning of the following theorem. The space F is supposed to be the space of functions among which interpolants are chosen. The functions having prescribed values at some given knots are supposed to be in K. Furthermore, d is a differential operator the images of which are contained in D; and $||d(f)||_D$ should represent the "degree of smoothness" of an $f \in K$.

Theorem 7.1 *Let F and D be Hilbert spaces; let K be a non-empty closed convex subset of F; and let $d\colon\ F \to D$ be a surjective linear* *operator. Assume that* Ker d *is finite-dimensional and* Ker $d \cap C_K = \{0\}$*, where* $C_K = \{f \in F\colon\ K + f \subseteq K\}$ *is the recession cone of* K*. Then there is a unique* $f \in K$ *minimising* $\{||d(f)||_D\colon\ f \in K\}$*, where* $||\cdot||_D$ *is the norm in* D*.*

The abstract Theorem 7.1 is needed to prove the following lemma. Here, the (closure of the) set A and the set B stand for the two polytope parameter spaces mentioned before. Furthermore, we consider a Sobolev space of functions from A to B; for the definition and basic properties of Sobolev spaces, we refer to [2].

The reason for proving that polytope parameter spaces are polytopes, becomes now clear. Namely, we assume that A fulfils the strong local Lipschitz condition, a notion explained in [2, 4.9], and fulfilled in particular by polytopes. It ensures by [2, 4.12] that the Sobolev space $W^{q,2}(A)$, and hence $W^{q,2}(A, \mathbb{R}^s)$ for any s, consists of continuous functions extendible to $\overline{A}$, assumed that q is large enough.

Finally, a finite set $A_0 \subseteq \mathbb{R}^r$ is called q-unisolvent if there is no non-trivial polynomial of degree q on $\mathbb{R}^r$ vanishing on A_0.

Lemma 7.2 *Let* $r, s \geq 1$*; let* A *be a bounded open subset of* $\mathbb{R}^r$ *fulfilling the strong local Lipschitz condition; and let* B *be a closed convex subset of* $\mathbb{R}^s$*. Let* $\overline{A}$ *and* B *be endowed with the pointwise natural order.*

Furthermore, let $q \geq \frac{r+3}{2}$ *and* $F = W^{q,2}(A, \mathbb{R}^s) = \{f \in L^2(A, \mathbb{R}^s)\colon D^t f \in L^2(A, S((\mathbb{R}^r)^t, \mathbb{R}^s)$ *for all* $t \leq q\}$*, where* $S((\mathbb{R}^r)^t, \mathbb{R}^s)$ *is the space of symmetric* t*-linear forms from* $\mathbb{R}^r$ *to* $\mathbb{R}^s$*. Let* $D = L^2(A, S((\mathbb{R}^r)^q, \mathbb{R}^s))$*; and let* $d\colon\ F \to D,\ \ f \mapsto D^q f$*.*

Finally, let A_0 *a finite subset* $\overline{A}$ *which is* q*-unisolvent; and let* $f_0 : A_0 \to B$ *be a monotone function. Let* $K(f_0) = \{f \in F :\ f(\overline{A}) \subseteq B, f$ *is monotone,* $f|_{A_0} = f_0\}$ *be non-empty.*

Then there is a unique $f \in K(f_0)$ *minimising* $\{||D^q f||_D\colon\ f \in K(f_0)\}$*.*

Proof. $K(f_0)$ is closed, convex, and by assumption non-empty. Because $d(F)$ is closed in D, we may consider d as a surjective function from F to $d(F)$. Ker $d = \mathrm{Pol}_q(A)$ is the space of polynomials of degree q, hence finite-dimensional. Furthermore, if $f \in C_{K(f_0)}$, then $f|_{A_0} = 0$, and if f is a polynomial of degree less than q, we have by assumption $f = 0$. So Theorem 7.1 applies. $\square$

It remains as a purely technical task to apply this lemma for the interpolation between spaces of convex or fuzzy sets.

Theorem 7.3 *Let $\mathcal{D}$ and $\mathcal{E}$ be ordered direction sets, containing r and s vectors, respectively; and let $R, S > 0$. Let $A = \mathcal{P}_{\mathcal{D},R}$ and $B = \mathcal{P}_{\mathcal{E},S}$ be the polytope parameter spaces associated to $\mathcal{D}$ and R and to $\mathcal{E}$ and S, respectively. Let $q \geq \frac{r+3}{2}$, and let $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k) \in A \times B$ such that $\{\alpha_1, \ldots, \alpha_k\}$ is q-unisolvent. Define $F = W^{q,2}(A, \mathbb{R}^s)$ and $D = L^2(A, S((\mathbb{R}^r)^q, \mathbb{R}^s))$. Then, among the monotone functions $f : A \to B$ in F such that $f(\alpha_1) = \beta_1, \ldots, f(\alpha_k) = \beta_k$, there is a unique one minimising $||D^q f||_D$.*

Proof. By Theorem 4.5, A is a polytope with non-empty interior, and thus $\overset{\circ}{A}$ is an open set fulfilling the strong local Lipschitz property. By Lemma 6.2, there is at least one $f \in F$ which maps A monotonously to B such that $f(\alpha_1) = \beta_1, \ldots, f(\alpha_k) = \beta_k$. So the assertion follows by Lemma 7.2. $\square$

Furthermore, we may formulate a similar result for fuzzy sets.

Theorem 7.4 *Let $\mathcal{D}$ and $\mathcal{E}$ be ordered direction sets, containing r and s vectors, respectively; let $R, S > 0$; and let $\Delta \subseteq [0,1]$ have cardinality d. Let $A = \mathcal{F}_{\mathcal{D},\Delta,R}$ and $B = \mathcal{F}_{\mathcal{E},\Delta,S}$ be the fuzzy set parameter spaces associated to $\mathcal{D}$, Δ, R and to $\mathcal{E}$, Δ, S, respectively. Let $q \geq \frac{dr+3}{2}$, and let $(\alpha_1, \beta_1), \ldots, (\alpha_k, \beta_k) \in A \times B$ such that $\{\alpha_1, \ldots, \alpha_k\}$ is q-unisolvent. Define $F = W^{q,2}(A, \mathbb{R}^{ds})$ and $D = L^2(A, S((\mathbb{R}^{dr})^q, \mathbb{R}^{ds}))$. Then, among the monotone functions $f : A \to B$ in F such that $f(\alpha_1) = \beta_1, \ldots, f(\alpha_k) = \beta_k$, there is a unique one minimising $||D^q f||_D$.*

Proof. This theorem is seen analogously to the preceeding one. $\square$

We conclude that, by Theorems 7.3 and 7.4, we know that a function extending a given if-then rule base and having the characteristic property of being the smoothest one, exists.

Unfortunately, the theorems do not provide any hint how to calculate this function. We are provided with neither more nor less than the knowledge that in principle, we can determine a function with properties comparable to the case of crisp spline interpolation.

8 Conclusion

We have proposed a way how to realise fuzzy inference by methods of interpolation: We showed that a fuzzy if-then rule base may be extended to a total function between fuzzy set universes, fulfilling a condition analogous to the case of spline interpolation between crisp values: it minimises a real value measuring smoothness.

The remarkable point about this approach is that we make use of a single principle, not allowing ad-hoc assumptions. As a disadvantage, we mention that the number of entries in the rule base must be rather large to ensure uniqueness of the interpolating function. We must moreover admit that the question how to realise the presented concept in practice is difficult. If technical realisability is a too hard task, it is our hope that, at least, our point of view on fuzzy inference will contribute to the development of further methods designed along similar lines.

Acknowledgements

This research was supported by the German Research Foundation (DFG) as part of the Collaborative Research Centre "Computational Intelligence" (SFB 531, subproject A1).

I would furthermore like to express my gratitute to the referees of this paper for their thorough, constructive critics.

References

[1] A. M. Anile, B. Falcidieno, G. Gallo, M. Spagnuolo, S. Spinello, Modeling uncertain data with fuzzy B-splines, *Fuzzy Sets Syst.* **113** (2000), 397 - 410.

[2] R. A. Adams, J. J. F. Fournier, “Sobolev spaces”, Academic Press, Oxford, 2-nd edition 2003.

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

[4] M. Atteia, “Hilbertian kernels and spline functions”, North-Holland, Amsterdam 1992.

[5] C. Bandt, G. Baraki, Metrically invariant measures on locally homogeneous spaces and hyperspaces, *Pac. J. Math.* **121** (1986), 13 - 28.

[6] P. Baranyi, L. T. Kóczy, T. D. Gedeon, A Generalized Concept for Fuzzy Rule Interpolation, *IEEE Trans. on Fuzzy Systems* **12** (2004), 820 - 837.

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

[8] P. Hájek, “Metamathematics of fuzzy logic”, Kluwer Acad. Publ., Dordrecht 1998.

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

[10] J-S. R. Jang, C.-T. Sun, Functional equivalence of radial basis function networks and fuzzy inference systems, *IEEE Trans. on Neural Networks* **4** (1993), 156 - 159.

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

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

[13] L. T. Kóczy, D. Tikk, T. D. Gedeon, On functional equivalence of certain fuzzy controllers and RBF type approximation schemes, *Int. J. of Fuzzy Systems* **2**, (2000), 164 - 175.

[14] V. Novák, I. Perfilieva, J. Močkoř, "Mathematical principles of fuzzy logic", Kluwer Academic Publishers, Dordrecht 1999.

[15] V. Novák, I. Perfilieva, Evaluating linguistic expressions and functional fuzzy theories in fuzzy logic, in: L. A. Zadeh et al. (eds.), "Computing with words in information / intelligent systems 1. Foundations", Physica-Verlag, Heidelberg 1999; pp. 383 - 406.

[16] R. Schneider, "Convex bodies: the Brunn-Minkowski theory", Cambridge University Press, Cambridge 1993.

[17] J. Stoer, C. Witzgall," Convexity and Optimization in finite dimensions I", Springer-Verlag, Berlin 1970.

[18] H. Thiele, Investigating the logical structure of fuzzy if-then rule bases using concepts of mathematical logic and of functional analysis, Technical Report CI-33/98, Dortmund university 1998.

[19] D. Tikk, P. Baranyi, Comprehensive analysis of a new fuzzy rule interpolation method, *IEEE Trans. on Fuzzy Systems* **8** (2000), 281 - 296.

[20] L. A. Zadeh, Outline of a new approach to the analysis of complex systems and decision processes, *IEEE Trans. Syst. Man Cybern.* **3** (1973), 28 - 44.