LIBERTAS MATHEMATICA, VOL. XVII (1997)

Methods of Solving Optimization Problems and Linear Equations in the Space of Fuzzy Vectors

Dan Butnariu*

ABSTRACT

The aim of this note is to present iterative methods of solving linear equations and convex optimization problems in the space of fuzzy vectors. These methods produce sequences which approximate the corresponding solutions with respect to the L^2-norm induced on the space of fuzzy vectors via Bobylev's transform.

1. Introduction

This note is aimed at presenting algorithms for finding solutions to optimization problems and to a class of linear equations having fuzzy vectors as unknowns. In order to precise the problem we are dealing with recall that a *fuzzy vector* is a function $u : \mathbb{R}^n \rightarrow [0, 1]$ such that, for each $\alpha \in (0, 1]$, the set

$$[u]^\alpha := \{x \in \mathbb{R}^n ;\ u(x) \geq \alpha\}$$

is nonempty, convex and closed and the set

$$[u]^0 := cl\left(\bigcup_{\alpha \in (0,1]} [u]^\alpha\right)$$

is bounded. The collection of all fuzzy vectors is denoted by $\mathcal{E}^n$, as usual. This set is

*This work was done during the author's visit to the Department of Mathematics of the University of Texas at Arlington, Arlington, Texas 71019, USA in September 1996.

provided with the linear operations induced by Zadeh's extension principle from the operations in $\mathbb{R}^n$, that is, if $u, v \in \mathcal{E}^n$ and if $c \in \mathbb{R}$, then $u + v$ and, respectively, cu is the unique fuzzy vector such that, for any $\alpha \in (0, 1]$,

$$[u + v]^\alpha = [u]^\alpha + [v]^\alpha$$

and, respectively,

$$[cu]^\alpha = c[u]^\alpha.$$

A subset $\mathcal{C}$ of $\mathcal{E}^n$ is said to be *convex* if, for each $\alpha \in (0,1)$ and for any two fuzzy vectors $u, v \in \mathcal{C}$, we have $\alpha u + (1 - \alpha)v \in \mathcal{C}$. If $\mathcal{C}$ is convex, then a function $F : \mathcal{C} \to \mathbb{R}$ is called *convex* if

$$F(\alpha u + (1 - \alpha)v) \leq \alpha F(u) + (1 - \alpha)F(v) \tag{1}$$

whenever $\alpha \in (0,1)$ and $u, v \in \mathcal{C}$. A function T from $\mathcal{C}$ to $\mathcal{E}^n$ is called *semiaffine* if, for every $\alpha \in (0,1)$ and for any $u, v \in \mathcal{C}$,

$$T(\alpha u + (1 - \alpha)v) = \alpha T(u) + (1 - \alpha)T(v). \tag{2}$$

A question of interest in practice is whether or under what conditions, given a convex function F on a convex set of fuzzy vectors $\mathcal{C}$, we can compute a fuzzy vector $u \in \mathcal{C}$ such that

$$F(u) = min\{F(v); v \in \mathcal{C}\} \tag{3}$$

provided that such a vector u exists. In what follows this will be called the *convex optimization problem (with fuzzy vectors as unknowns)*. Another question of practical interest is that of finding a way of computing solutions to what we call *linear equations* (*with fuzzy vectors as unknowns*), that is solutions in $\mathcal{C}$ to equations in the form

$$T(u) = g, \tag{4}$$

where $T : \mathcal{C} \to \mathcal{E}^n$ is semiaffine and $g \in Range(T)$. The purpose of this work is to emphasize algorithms for solving the problems (3) and (4). Problems similar to those mentioned above are widely discussed in the fuzzy set theoretical literature of the last two decades and we refer the reader to the results and bibliography in [1, Chapter 14] for more information on this topic.

2. The framework

In what follow we endow the set of fuzzy vectors $\mathcal{E}^n$ with the linear structure indicated in the previous section and with the pseudometric defined as follows. To each fuzzy vector u we associate its *support function* $s_u : [0, 1] \times \mathbb{R}^n \to \mathbb{R}$ given by

$$s_u(\alpha, x) = sup\{ < a, x > ; \; a \in [u]^\alpha \}.$$

This function is well-defined (see [1]). The *Bobylev representation of the fuzzy vector* u is the function $\widehat{u} : B^n \to \mathbb{R}$ given by

$$\widehat{u}(x) = s_u(\|x\|, x),$$

for any x belonging to the closed unit ball B^n in $\mathbb{R}^n$. According to [1, Proposition 6.4.1], the function $\widehat{u}$ is upper semicontinuous on B^n. Therefore, $\widehat{u}$ is also measurable as a real function on the set B^n provided with the σ-algebra $\mathcal{A}$ of its Lebesgue measurable subsets. From [1, Proposition 6.1.12] we deduce that, for any $\alpha \in [0,1]$ and $x \in B^n$,

$$|s_u(\alpha, x)| \leq sup\{\|y\|; \; y \in [u]^0\},$$

where the set $[u]^0$ is bounded. Consequently, the function $\widehat{u}$ is bounded too. Being measurable and bounded the function $\widehat{u}$ is p-integrable on B^n with respect to the Lebesgue measure λ for any $p \in [1, \infty)$. Obviously, this implies that $\widehat{u}$ is p-integrable with respect to the probability measure

$$\mu = (\lambda(B^n))^{-1}\lambda.$$

Hence, the function $\Delta_p : \mathcal{E}^n \times \mathcal{E}^n \to \mathbb{R}_+$ given by

$$\Delta_p(u, v) = \left(\int_{B^n} |\widehat{u} - \widehat{v}|^p d\mu \right)^{1/p}$$

is well defined and it is a pseudometric on the set of fuzzy vectors, i.e., it has the following properties:

(i) $\Delta_p(u, v) = \Delta_p(v, u)$ and

(ii) $\Delta_p(u, v) \leq \Delta_p(u, w) + \Delta_p(w, v)$,

for any $u, v, w \in \mathcal{E}^n$. Note that $\Delta_p(u, v) = 0$ iff $u = v$, μ-a.e.

Recall (see [1, Section 6.4]) that the function $\jmath$: $u \mapsto \widehat{u}$ is one-to-one and that, for any $u, v \in \mathcal{E}^n$ and for any two numbers $\alpha, \beta \geq 0$ we have

$$\jmath(\alpha u + \beta v) = \alpha \jmath(u) + \beta \jmath(v).$$

This means that $\jmath$ is a isometric embedding of the space $\mathcal{E}^n$ provided with the pseudometric Δ_p into a convex cone of the Banach space $\mathcal{L}^p : = \mathcal{L}^p(B^n, \mathcal{A}, \mu)$ provided with its usual algebraic and metric structures. In particular, this applies when $p = 2$, a situation we will use in the sequels. We denote $\widehat{\mathcal{E}}^n = \jmath(\mathcal{E}^n)$. Combining Proposition 6.4.1 in [1] with the fact that any convergent sequence in $\mathcal{L}^p$ has a subsequence which converges uniformly μ-a.e. on B^n we obtain that $\widehat{\mathcal{E}}^n$ is closed in $\mathcal{L}^p$.

If $\mathcal{C}$ is a subset of $\mathcal{E}^n$, then $\widehat{\mathcal{C}}$ denotes its image via $\jmath$ in the space $\mathcal{L}^2$. Observe that if $\mathcal{C}$ is convex, then so is $\widehat{\mathcal{C}}$. Also, if $\mathcal{C}$ is a convex set of fuzzy vectors and the function $F : \mathcal{C} \to \mathbb{R}$ is convex, then so is the function $\widehat{F} : \widehat{\mathcal{C}} \to \mathbb{R}$ defined by

$$\widehat{F}(\widehat{v}) = F(\jmath^{-1}(\widehat{v})) \tag{5}$$

where $\jmath^{-1} : \widehat{\mathcal{E}}^n \to \mathcal{E}^n$ is the inverse of $\jmath$. Note that the proof of Proposition 6.4.2 in [1] indicates a way of finding $u : = \jmath^{-1}(\widehat{u})$ when $\widehat{u} \in \widehat{\mathcal{E}}^n$ is given.

Let $\mathcal{C}$ be a nonempty convex set of fuzzy vectors and let $F : \mathcal{C} \to \mathbb{R}$ be a convex function. It is easy to see that, if $u \in \widehat{\mathcal{C}}$ is a solution of the optimization problem

$$\widehat{F}(\widehat{u}) = min\{\widehat{F}(\widehat{v});\ \widehat{v} \in \widehat{\mathcal{C}}\} \tag{6}$$

then $u = \jmath^{-1}(\widehat{u})$ is a solution of (3) and conversely. These show that for solving the convex optimization problem (3) it is sufficient to find a solution $\widehat{u} \in \widehat{\mathcal{C}}$ of the optimization problem (6) and, then, $u = \jmath^{-1}(\widehat{u})$ is a solution of (3). Similarly, for solving the linear equation (4) it is sufficient to find a solution $\widehat{u} \in \widehat{\mathcal{C}}$ of the equation

$$\widehat{T}(\widehat{v}) = \widehat{g} \tag{7}$$

with

$$\widehat{T}(\widehat{v}) : = \jmath(T(\jmath^{-1}(\widehat{v}))) \tag{8}$$

and, then, $u = \jmath^{-1}(\widehat{u})$ is a solution of (4). Our main results, presented below, show ways of solving the problems (6) and (7) in the Hilbert space $\mathcal{L}^2$ and, by doing that, they allow solving the problems (3) and (4).

There is a significant advantage of solving (6) and, respectively (7), and using the final results for determining the solutions of (3) and, respectively, (4). Namely, all numerical computations are done in a real Hilbert space and not in the space of fuzzy vectors where computing is a complicated process because addition and multiplication should be done level by level. On the other hand, the convergence of our algorithms is guaranteed in the Δ_2 pseudometric whose induced topology is weaker than the uniform Hausdorff metric topology on the space of fuzzy vectors. It is an interesting open question whether, or in what conditions, the algorithms presented below still converge in the uniform Hausdorff metric or in any other metric on $\mathcal{E}^n$ among those discussed in [1] and in [2] and in the references therein.

3. The algorithms

3.1 *Solving linear equations in $\mathcal{E}^n$*. In this section the space $\mathcal{E}^n$ is provided with the pseudometric Δ_2, C is a convex closed subset of $\mathcal{E}^n$, $T : C \to \mathcal{E}^n$ denotes a semiaffine function and $\widehat{T}$ is the function defined by (8). We associate with T and with each $x \in B^n$ the function $\widehat{T}_x : \widehat{\mathcal{E}}^n \to \mathbb{R}$ defined by

$$\widehat{T}_x(\widehat{u}) = \widehat{T}(\widehat{u})(x).$$

Suppose that $g \in Range(T)$. Then, since T is semiaffine, it follows that, for all $x \in B^n$, the set

$$Q_x : = \{\widehat{v} \in \widehat{C};\ \widehat{T}_x(\widehat{v}) = \widehat{g}(x)\}$$

is nonempty and convex. If $\widehat{T}_x$ is continuous, then Q_x is closed. Therefore, in this case, for each $\widehat{v} \in \widehat{C}$, there exists an unique orthogonal (metric) projection $P_x(\widehat{v})$ onto the set Q_x. Note that the point-to-set mapping $x \to Q_x : B^n \to \mathcal{L}^2$ has an integrable selector, namely, the *constant* function $x \to \widehat{u}$ where u is such that $T(u) = g$. This implies that, for any $\widehat{v} \in \widehat{C}$, the function $x \to P_x(\widehat{v}) : B^n \to \mathcal{L}^2$ is Bochner μ-integrable (see [3, Lemma 2.1]). Thus, Theorem 2.2 in [3] as well as Theorem 4.4 in [4] apply. They lead to the next result:

Theorem 1. *Suppose that the following conditions hold:*
(i) For each $x \in B^n$ the function $\widehat{T}_x$ is continuous;
(ii) $g \in Range(T)$.

Then, for any initial point $u^0 \in C$, the sequence $\{\widehat{u}^k\}$ in $\mathcal{L}^2$ recursively generated according to the rule

$$\widehat{u}^{k+1} = \int_{B^n} P_x(\widehat{u}^k) d\mu(x)$$

is contained in $\widehat{C}$ and converges weakly (in $\mathcal{L}^2$) to a point $\widehat{u}$ in $\widehat{C}$ such that $u : = \jmath^{-1}(\widehat{u})$ is a solution of the linear equation (4). Moreover, if C is compact, then the sequence $\{\widehat{u}^k\}$ converges strongly to $\widehat{u}$ in $\mathcal{L}^2$ and the sequence $\{u^k : = \jmath^{-1}(\widehat{u}^k)\}$ converges in $\mathcal{E}^n$ to a solution u of the linear equation (4).

It should be noted that practical applicability of the algorithm for approximating solutions of the linear equation (4) described in Theorem 1 above depends on the possibility of effectively computing the projections P_x. In some circumstances, this can be easily done. For instance, if T is the extension by Zadeh's principle of a linear function

$S:\mathbb{R}^n \to \mathbb{R}^n$, then the functions $\widehat{T}_x$ are necessarily continuous (the argument in this respect is based on [5, Theorem 5]) and determining the projections P_x can be achieved by using a closed formula presented in [3, Equation (15)].

3.2 Solving convex optimization problems. In this section $\mathcal{C}$ is assumed to be a nonempty, compact, convex subset of $\mathcal{E}^n$ and $F:\mathcal{C}\to\mathbb{R}$ is supposed to be convex and continuous. Therefore, the convex optimization problem (3) has a solution in $\mathcal{C}$ and, then, the problem (6) has a solution in $\widehat{\mathcal{C}}$. For each $\widehat{v}\in\widehat{\mathcal{C}}$, denote

$$\mathcal{U}_{\widehat{v}} = \{\widehat{u}\in\widehat{\mathcal{C}};\ \widehat{F}(\widehat{u}) - \widehat{F}(\widehat{v}) \le 0\}.$$

Since $\widehat{F}$ is continuous, it follows that the point to set mapping $\widehat{v}\to\mathcal{U}_{\widehat{v}}:\mathcal{C}\to\mathcal{L}^2$ is measurable with respect to the σ-algebra $\mathcal{B}$ of all Borel subsets of $\widehat{\mathcal{C}}$ (cf. [6, Theorem 8.2.9]). If $\widehat{u}$ is a solution of (6), then, for any complete probability measure ν on $\mathcal{B}$ the constant function $\widehat{v}\to\widehat{u}:\widehat{\mathcal{C}}\to\mathcal{L}^2$ is a square integrable selector of the point-to-set mapping $\mathcal{U}$. Hence, according to Lemma 2.1 in [3], for any complete probability measure ν on $\mathcal{B}$ and for any $\widehat{w}\in\widehat{\mathcal{C}}$, the function $\widehat{v}\to\Pi_{\widehat{v}}(\widehat{w}):\widehat{\mathcal{C}}\to\widehat{\mathcal{C}}$ which assigns to each $\widehat{v}\in\widehat{\mathcal{C}}$ its orthogonal projection $\Pi_{\widehat{v}}(\widehat{w})$ onto the set $\mathcal{U}_{\widehat{w}}$ is ν-square integrable. These show that in our circumstances, Theorem 4.4 in [4] is applicable and it gives the following result:

Theorem 2. *For any complete probability measure ν on the σ-algebra $\mathcal{B}$ of all Borel subsets of $\widehat{\mathcal{C}}$ the sequence $\{\widehat{w}^k\}$ defined in $\mathcal{L}^2$ by the recursive formula*

$$\widehat{w}^{k+1} = \int_{\widehat{\mathcal{C}}}\Pi_{\widehat{w}}(\widehat{w}^k)d\nu(\widehat{w}),$$

exists no matter how the initial point $\widehat{w}^0\in\widehat{\mathcal{C}}$ is chosen. This sequence converges strongly in $\mathcal{L}^2$ to a point $\widehat{u}\in\widehat{\mathcal{C}}$ which is a solution of the problem (6) and such that $\mathrm{u}:=\jmath^{-1}(\widehat{u})$ is a solution of the convex optimization problem (3).

There are two essential difficulties when it comes to apply this result to solving practical convex optimization problems. The first is that of finding a complete probability measure ν on $\mathcal{B}$ such that determining the integrals involved in generating the sequence $\{\widehat{w}^k\}$ will be computationally feasible. The second difficulty concerns finding the orthogonal projections $\Pi_{\widehat{w}}(\widehat{w}^k)$ at each computational step of the procedure. This last difficulty can be easily overcome in some situations as, for instance, in the case when F is semiaffine (in particular, when F is the extension by Zadeh's principle of a linear functional on $\mathbb{R}^n$). In such a case the sets $\mathcal{U}_{\hat{v}}$ are subsets of semispaces in $\mathcal{L}^2$ and closed formula are available for determining the required projections (see [3]). The first difficulty may be more demanding. However, a careful analysis of the algorithm presented in Theorem 2 suggests that it may be possible to find a sequence $\{\widehat{s}^m\}$ in $\widehat{\mathcal{C}}$ such that the averages

$$\frac{1}{m}\sum_{i=1}^{m}\Pi_{s^i}(\widehat{w}^k)$$

converge to the same limit as the sequence $\{\widehat{w}^k\}$ itself. If and how this can be done is a question whose answer we do not know.

References

[1] P. Diamond and P. Kloeden, *Metric Spaces of Fuzzy Sets*, World Scientific, Singapore, 1994.

[2] M. Rojas-Medar and H. Roman-Flores, *On the equivalence of convergences of fuzzy sets*, Fuzzy Sets and Systems, 80, 1996, 217-224.

[3] D. Butnariu, *The expected projection method: Its behavior and applications*, J. Applied Analysis, 1, 1995, 93-108.

[4] D. Butnariu and S. Flam, *Strong convergence of the expected projection method in Hilbert spaces*, Numer. Funct. Anal. and Optim., 16, 1995, 601-636.

[5] L.C. de Barros, R.C. Bassanezi and P.A. Tonelli, *On the continuity of Zadeh's extension principle*, Proc. of the IFSA'97 Congress, Prague, 1997.

[6] J.-P. Aubin and H. Frankowska, *Set Valued Analysis*, Birkhauser, Basel, 1990.