A Differential Evolution Algorithm for computing the Fuzzy Variance

Luciano Stefanini
University of Urbino 'Carlo Bo', Italy
luciano.stefanini@uniurb.it

Abstract

We propose a differential evolution (DE) algorithm for the calculation of the interval and fuzzy variance. In particular, we see that the DE methods can be efficient for the fuzzy variance of a relatively high number of fuzzy data; computational results with up to 100 data show that the number of function evaluations to obtain the estimated global solutions grows less then quadratically with the number of data.

Keywords: Fuzzy Variance, Differential Evolution.

1 Introduction

It is well known that the calculation of the variance in the case of interval or fuzzy data is NP-hard and finding heuristic procedures with good computational performance is an important field of research.

The problem at hand is a global constrained minimization and maximization (over box constraints of the form $a_i \leq x_i \leq b_i$) of a convex function; the minimization of a convex function is not difficult and most procedures will produce the optimum point (it is unique by the local=global minimum theorem for convex functions); but the maximization of a convex function is quite a different question and finding the global max value is an open problem).

On the other hand, the box-constrained max of a convex function is combinatorial as the solution is always at a vertex of the box (the proof is easy and well known) and in n dimensions the number of vertices is 2^n as any of the vertices is a candidate solution.

The exact solution in the worst case requires the evaluation of the function at all the vertices of the box and this cannot be done even for n of the order of 30-50 (2^{50}=1.1259E+15). For recent results on the box-constrained maximization of a convex function see [11].

For these reasons, it is necessary to adopt more complex procedures as the problem is incomparably harder then min of a convex function.

Obtaining a good (possibly near-optimal) solution for the fuzzy (or interval)-valued variance is well solved for low-dimensional problems (see [1],[2] and [3]), but in many applications of interval and data analysis and related fields, reasonable dimensions is 100-200 or more, so that the search for heuristic (possibly fast) procedures is of great importance and interest.

We first report (section 2) some properties of the variance problem, in particular the combination of invariance to translations and homegenity of degree two give some indications to the structure of the problem and some ideas to develop heuristic procedures.

In section 3 we describe the DE (differential evolution) algorithm for the global optimization of a function with box constraints. Similarly to all evolutionary algorithms for optimization, DE uses a finite population (not a single point) of solutions; at each generation, a population of potential solution points (chromosomes) is recombined to produce a new generation of candidate points which contains a better solution than in previous generations, i.e. at each generation the solution is tentatively improved.

We then apply the DE procedure to the variance problem and, in the final section, we report some computational results with up to 100 data. The fact that the experienced computational effort grows less then quadratically with the dimension of the problem (i. e. the number of

L. Magdalena, M. Ojeda-Aciego, J.L. Verdegay (eds): Proceedings of IPMU'08, pp. 1559–1566
Torremolinos (Málaga), June 22–27, 2008

fuzzy data) seems of interest also for high dimensional problems (section 4).

2 General Setting

We consider the problem of calculating the interval and the fuzzy estimation of the variance of n compact intervals $\tilde{x}_i = [x_i^-, x_i^+]$ of the set $\mathbb{R}$ of real numbers or n fuzzy numbers (in the $\alpha - cut$ representation): for $i = 1,...,n$

$\widetilde{X}_i = \{[\widetilde{X}_i]_\alpha = [x_{i,\alpha}^-, x_{i,\alpha}^+], \alpha \in [0,1]\}$. As the fuzzy case can be reduced to the calculation of the interval for each cut (i.e. for each $\alpha \in [0,1]$) we first describe the case of intervals.

The interval variance is defined to be the interval $\tilde{v} = [v^-, v^+]$ containing the possible values of the standard statistical variance $v = k \sum_{i=1}^{n} (x_i - m)^2$ for all possible values of $x_i \in [x_i^-, x_i^+]$ and the corresponding mean value $m = \frac{1}{n} \sum_{i=1}^{n} x_i$; the usual values of the coefficient k are $k = \frac{1}{n}$ or $k = \frac{1}{n-1}$. As the variance can be considered to be a function of the n observations $x_1,...,x_n$ the problem becomes that of extending it to the interval arguments $\tilde{x}_i$; consider simply the function

$$w = f(x_1,...,x_n) = \sum_{i=1}^{n} \left(x_i - \frac{1}{n} \sum_{j=1}^{n} x_j \right)^2 \tag{1}$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} \left((n-1)x_i - \sum_{j=1, j \neq i}^{n} x_j \right)^2 ; \tag{2}$$

as $v = kw$, the interval variance $\tilde{v} = k\tilde{w}$ is obtained by scalar multiplication of the interval extension $\tilde{w}$ of w by the constant factor k.

We have

$$\tilde{w} = \{f(x_1,...,x_n) \mid x_i \in [x_i^-, x_i^+], \forall i\} \tag{3}$$

and, being f a continuous function,

$$\widetilde{w} = [w^-, w^+] \text{ where}$$

$$w^- = \min\{f(x_1,...,x_n) \mid x_i \in [x_i^-, x_i^+]\} \tag{4}$$

$$w^+ = \max\{f(x_1,...,x_n) \mid x_i \in [x_i^-, x_i^+]\} \tag{5}$$

The problem then becomes that of globally minimize and maximize the quadratic function $f(x_1,...,x_n)$ with the box constraints $x_i \in [x_i^-, x_i^+]$, $i = 1,...,n$.

It is well known that the global minimum (4) is easy to compute, but the maximization (5) is NP-hard.

So the main interest is in the computation of the upper level w^+ by the maximization (5).

In the following, we denote by $\hat{x} = (\hat{x}_1,...,\hat{x}_n)$ the point corresponding to the global maximum of $f(x_1,...,x_n)$ over the box (multidimensional interval) $\mathbb{X} = \bigtimes_{i=1}^{n} [x_i^-, x_i^+]$ obtained as cartesian product of the single intervals.

We first observe some properties of the quadratic function $f(x_1,...,x_n)$ in (1).

1. The first partial derivatives of f are

$$\frac{\partial f}{\partial x_i} = \frac{2(n-1)}{n^2} \sum_{i=1}^{n} \left((n-1)x_i - \sum_{j=1, j \neq i}^{n} x_j \right)$$

and we may have globally positive or globally negative partial derivatives if for some index i one of the inequalities hold:

$$\text{If } (n-1)x_i^- - \sum_{j=1, j \neq i}^{n} x_j^+ \geq 0 \text{ then } \frac{\partial f}{\partial x_i} \geq 0, \forall x \in \mathbb{X} \tag{6}$$

$$\text{If } (n-1)x_i^+ - \sum_{j=1, j \neq i}^{n} x_j^- \leq 0 \text{ then } \frac{\partial f}{\partial x_i} \leq 0, \forall x \in \mathbb{X}. \tag{7}$$

In case (6) the maximization point $\hat{x} = (\hat{x}_1,...,\hat{x}_n)$ has $\hat{x}_i = x_i^+$ and in case (7) it has $\hat{x}_i = x_i^-$; these two rules can possibly reduce the number of variables in the maximization to be solved; a similar result, where x_i^+ and x_i^- are inverted, holds for the minimization problem.

2. Function f in (1) is invariant to diagonal transformations, i.e. for every $x = (x_1,...,x_n)$ and for any real number α, we have $f(x + \alpha e) = f(x)$ where $e = (1,1,...,1)$; this implies that there exist many points of the box corresponding to the same value of $f(x_1,...,x_n)$.

Three interesting consequences of this property are:

(i) without changing the resulting interval variance, we can translate all the data intervals such that $x_i^- \geq l > 0$ by the transformations

$$x_i^- \to x_i^- + \bar{t},\ \ x_i^+ \to x_i^+ + \bar{t} \ \text{ with}$$
$$\bar{t} = l + \max_{x_i^- < 0} | x_i^- | \ ;$$

(ii) if $h = \min_{i=1,\ldots,n} \{x_i^+ - x_i^-\}$ is the minimal length of the intervals, then

$f(x_1^-,\ldots,x_n^-) = f(x_1^- + t,\ldots,x_n^- + t), \forall t \in [0,h]$;

(iii) if $x^* = (x_1^*,\ldots,x_n^*)$ is a feasible point and

$h^- = \max_{i=1,\ldots,n} \{x_i^- - x_i^*\} \leq 0$,

$h^+ = \min_{i=1,\ldots,n} \{x_i^+ - x_i^*\} \geq 0$.

Then

$f(x_1^*,\ldots,x_n^*) = f(x_1^* + t,\ldots,x_n^* + t)$,

$\forall t \in [h^-,h^+]$; this implies that the optimal solution can be always found at the boundary of the box.

3. Function f in (1) is homogeneous of degree two, i.e. for every $x = (x_1,\ldots,x_n)$ and for any real number t, we have $f(tx) = t^2 f(x)$.

Invariance to diagonal transformations and homogeneity imply that $f(tx + \alpha e) = t^2 f(x)$ for all feasible x and all t, α such that $tx + \alpha e$ is feasible.

4. Denote the level sets of function (1) by

$L(c) = \{x \mid x = (x_1,\ldots,x_n),\ f(x_1,\ldots,x_n) = c\}$;

we have $L(c) = \varnothing$ if $c < 0$ and

$L(0) = \{(t,\ldots,t) \mid t \in \mathbb{R}\}$.

Let $\mathbb{S} = \{(u_1,\ldots,u_n) \mid u_i \in \mathbb{R},\ \sum_{i=1}^n u_i = 1\}$ be the unit simplectic hyperplane (the unit simplex is the subset $\mathbb{S}^+$ of $\mathbb{S}$ in the positive ortant). We have $f(u_1,\ldots,u_n) = \sum_{i=1}^n (u_i - \frac{1}{n})^2$ for all $u \in \mathbb{S}$

so that, on $\mathbb{S}$, the function in (1) is separable.

5. We have $L(0) = \{(\frac{1}{n},\ldots,\frac{1}{n})\}$ and, for $c > 0$,

$$L(c) \cap \mathbb{S} = \{u \mid \sum_{i=1}^n (u_i - \frac{1}{n})^2 = c,\ \ \sum_{i=1}^n u_i = 1\};$$

this means that the projections of each $L(c)$ into $\mathbb{S}$ is a $(n-1)$-dimensional sphere centered at $(\frac{1}{n},\ldots,\frac{1}{n})$ with radius $\sqrt{c}$.

6. Let $\mathbb{S}_{\mathbb{X}}^+ = \{u \mid u \in \mathbb{S}\ ,\ \ u_i \geq 0\ \ \forall i,\ \exists t \in \mathbb{R}^+ \text{ s.t. } tu \in \mathbb{X}\}$ be the points u of the positive unit simplex such that, for at least one positive value of t it is $x_i^- \leq tu_i \leq x_i^+\ \ \forall i$. Clearly, assuming (eventually after a transformation) that $x_i^- > 0\ \ \forall i$, we have also $u_i > 0\ \ \forall i\ \ \forall u \in \mathbb{S}_{\mathbb{X}}^+$. The condition for $u \in \mathbb{S}$ to be $u \in \mathbb{S}_{\mathbb{X}}^+$ is that there is a $t \in \mathbb{R}^+$ with

$$\frac{x_i^-}{u_i} \leq t \leq \frac{x_i^+}{u_i}\ \ \forall i = 1,\ldots,n$$

or, equivalently that

$$\max\{\frac{x_i^-}{u_i} \mid i = 1,\ldots,n\} \leq \min\{\frac{x_i^+}{u_i} \mid i = 1,\ldots,n\}.$$

It is easy to see that $\mathbb{S}_{\mathbb{X}}^+$ is a convex subset of the unit simplex.

7. For a given $u \in \mathbb{S}_{\mathbb{X}}^+$, define

$$t_u^+ = \min\{\frac{x_i^+}{u_i} \mid i = 1,\ldots,n\} \geq 1$$
then
$$t_u^- = \max\{\frac{x_i^-}{u_i} \mid i = 1,\ldots,n\} \leq 1;$$

$$w^- = \min\{(t_u^-)^2 \sum_{i=1}^n (u_i - \frac{1}{n})^2 \mid u \in \mathbb{S}_{\mathbb{X}}^+\}$$
$$w^+ = \max\{(t_u^+)^2 \sum_{i=1}^n (u_i - \frac{1}{n})^2 \in \mathbb{S}_{\mathbb{X}}^+\}.$$

8. After the properties above, the determination of w^- and w^+ is equivalent to solving the optimization problems

$$w^- = \underset{t,u}{Min}\ \ t^2 \sum_{i=1}^n (u_i - \tfrac{1}{n})^2$$
$$w^+ = \underset{t,u}{Max}\ \ t^2 \sum_{i=1}^n (u_i - \tfrac{1}{n})^2$$
s.t.

$$\sum_{i=1}^n u_i = 1$$
$$x_i^- \leq tu_i \leq x_i^+,\ i = 1,\ldots,n$$
$$t \geq 0, u_i > 0,\ i = 1,\ldots,n.$$

For a given feasible u, the determination of the optimal value of t is easy:

$$t^-(u) = \max\{\frac{x_i^+}{u_i} \mid i = 1,\ldots,n\} \text{ for } w^-$$
$$t^+(u) = \min\{\frac{x_i^-}{u_i} \mid i = 1,\ldots,n\} \text{ for } w^+.$$

9. If $x^* = (x_1^*,...,x_n^*)$ is a feasible point, then $\exists t^-, t^+ \geq 0$ and $\exists \alpha^-, \alpha^+ \geq 0$ such that $t^- x^* + \alpha^-$ and $t^+ x^* - \alpha^+$ are feasible and

$$f(t^- x^* + \alpha^-) \leq f(tx^* + \alpha) \leq f(t^+ x^* - \alpha^+)$$

$\forall t \in [t^-, t^+]$ and $\forall \alpha \in [-\alpha^+, \alpha^-]$; t^-, α^- are obtained by solving the two dimensional LP minimization

$$\min_{t,\alpha \geq 0} t \quad \text{s. t. } x_i^- \leq t x_i^* + \alpha \leq x_i^+ \ \forall i \tag{8}$$

and t^+, α^+ are obtained by solving the two dimensional LP maximization

$$\max_{t,\alpha \geq 0} t \quad \text{s. t. } x_i^- \leq t x_i^* - \alpha \leq x_i^+ \ \forall i. \tag{9}$$

This suggests a possible heuristic procedure:

Step 1. Select randomly a set of P feasible points $x^{(p)}$, $p = 1,...,P$ with $x_i^- \leq x_i^{(p)} \leq x_i^+$ $\forall i$;

Step 2. For each $p = 1,...,P$, solve the two LP problems (8) and (9) using $x^* = x^{(p)}$ and let $(t_-^{(p)}, \alpha_-^{(p)})$ and $(t_+^{(p)}, \alpha_+^{(p)})$ be the corresponding solutions; evaluate $w_-^{(p)} = f(t_-^{(p)} x^{(p)} + \alpha_-^{(p)})$ and $w_+^{(p)} = f(t_+^{(p)} x^{(p)} - \alpha_+^{(p)})$;

Step 3. Approximated values of w^- and w^+ are, respectively, $\min\{w_-^{(p)} \mid p = 1,2,...,P\}$ and $\max\{w_+^{(p)} \mid p = 1,2,...,P\}$.

3 A DE Algorithm for Fuzzy Variance

Consider a feasible point $x \in \mathbb{X}$ such that $x_i^- \leq x_i \leq x_i^+$, $i = 1,...,n$ and transform $x = ru$ with $r = \sum_{i=1}^{n} x_i > 0$ and $u_i = \frac{x_i}{r}$ (clearly, $u_i > 0 \ \forall i$). Then, better points x' and x'' for the Min and Max problems are, respectively,

$$x' = t_u^- u \text{ with } t_u^- = \max\{\frac{x_i^+}{u_i} \mid i = 1,...,n\}$$

$$x'' = t_u^+ u \text{ with } t_u^+ = \min\{\frac{x_i^-}{u_i} \mid i = 1,...,n\}$$

as, in fact (the considered values of t are such that $tu \in \mathbb{X}$)

$$w^-(x') = (t_u^-)^2 \sum_{i=1}^{n} (u_i - \frac{1}{n})^2 \leq t^2 \sum_{i=1}^{n} (u_i - \frac{1}{n})^2$$

and

$$w^+(x'') = (t_u^+)^2 \sum_{i=1}^{n} (u_i - \frac{1}{n})^2 \geq t^2 \sum_{i=1}^{n} (u_i - \frac{1}{n})^2$$

This produces an advantage in performing the optimizations, as to each trial point $x \in \mathbb{X}$ we can immediately substitute a (boundary) feasible point x' or x'' (along the direction $u = \frac{x}{\|x\|_1}$, $\| x \|_1 = \sum_{i=1}^{n} | x_i |$ being the usual $1 - norm$) where the objective function is better then at x.

This fact will in general speed up any optimization procedure and we will see that even high dimensional problems (up to 100 variables) can be handled by a suitable implementation of a Differential Evolution algorithm.

To take more advantage in the calculation of the fuzzy variance, we represent all the fuzzy numbers the LU (Lower-Upper) parametrization introduced in [6]; in such a way, the number of optimization problems is reduced without loosing the desired precision in the determination of the membership function of the fuzzy variance (see also [7] for the application of similar ideas to the general fuzzy extension of functions).

The simpler LU parametrization of the $\alpha - cuts$ $[u_\alpha^-, u_\alpha^+]$ of a fuzzy number u is obtained on the trivial decomposition of interval $[0,1]$ with $N = 1$ (without internal points) and $\alpha_0 = 0, \alpha_1 = 1$:

$$\begin{aligned} u_\alpha^- &= u_0^- + (u_1^- - u_0^-) p(\alpha; \delta u_0^-, \delta u_1^-) \\ u_\alpha^+ &= u_0^+ + (u_1^+ - u_0^+) p(\alpha; \delta u_0^+, \delta u_1^+). \end{aligned} \tag{10}$$

The shape function $p(\alpha; \delta u_0, \delta u_1)$ is taken from a family (see [6]) of monotonic increasing functions over $[0,1]$ and such that $p(0) = 0$, $p(1) = 1$, $p'(0) = \delta u_0$ and $p'(1) = \delta u_1$. In this simple case, u can be represented by a vector of 8 components

$$u = (u_0^-, \delta u_0^-, u_0^+, \delta u_0^+; u_1^-, \delta u_1^-, u_1^+, \delta u_1^+) \tag{11}$$

with u_0^-, δu_0^-, u_1^-, δu_1^- for the lower branch u_α^- and u_0^+, δu_0^+, u_1^+, δu_1^+ for the upper branch u_α^+; the values are such that $u_0^- \leq u_1^- \leq u_1^+ \leq u_0^+$ to have the support $[u_0^-, u_0^+]$ and the core $[u_1^-, u_1^+]$ of u.

Let $\tilde{w} = \tilde{f}(\tilde{x}_1, \tilde{x}_2, ..., \tilde{x}_n)$ denote the fuzzy variance of n fuzzy variables $[\tilde{x}_k]_\alpha = [x_{k,\alpha}^-, x_{k,\alpha}^+]$;

to obtain the parametric LU representation of $\widetilde{w}$ we need to calculate the usual minima w_0^-, w_1^- and maxima w_0^+, w_1^+ but also the slopes $\delta w_0^-, \delta w_1^-$ and $\delta w_0^+, \delta w_1^+$, obtained by the following rule: for the two membership levels α_i, $i = 0,1$ where $\alpha_0 = 0, \alpha_1 = 1$, denote by $\widehat{x}_i^- = (\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-)$ and $\widehat{x}_i^+ = (\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+)$ the points where the $\min$ and the $\max$ take place; then

$w_i^- = f(\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-)$, $\quad w_i^+ = f(\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+)$;

the slopes δw_i^-, δw_i^+ are computed (as f is differentiable and the derivatives are easy to calculate in closed form) by

$$\begin{aligned}\delta w_i^- = &\sum_{\substack{k=1\\ \widehat{x}_{k,i}^- = x_{k,i}^-}}^{n} \frac{\partial f(\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-)}{\partial x_k} \delta x_{k,i}^- \\ &+ \sum_{\substack{k=1\\ \widehat{x}_{k,i}^- = x_{k,i}^+}}^{n} \frac{\partial f(\widehat{x}_{1,i}^-, ..., \widehat{x}_{n,i}^-)}{\partial x_k} \delta x_{k,i}^+\end{aligned} \tag{12}$$

$$\begin{aligned}\delta w_i^+ = &\sum_{\substack{k=1\\ \widehat{x}_{k,i}^+ = x_{k,i}^-}}^{n} \frac{\partial f(\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+)}{\partial x_k} \delta x_{k,i}^- \\ &+ \sum_{\substack{k=1\\ \widehat{x}_{k,i}^+ = x_{k,i}^+}}^{n} \frac{\partial f(\widehat{x}_{1,i}^+, ..., \widehat{x}_{n,i}^+)}{\partial x_k} \delta x_{k,i}^+.\end{aligned} \tag{13}$$

The idea of DE to find *Min* or *Max* is to start with an initial "population" $x^{(1)} = (x_1, ..., x_n)^{(1)}, ..., x^{(p)} = (x_1, ..., x_n)^{(p)}$ of p feasible points for each generation (i.e. for each iteration) to obtain a new set of points by recombining randomly the individuals of the current population and by selecting the best generated elements to continue in the next generation. To define a starting population to initialize the DE procedure, as it is rare to have information about promising subsets to privilege, it is natural to seed random points belonging to the feasible box (satisfying the boundary constraints). The initial population is then chosen randomly to "cover" uniformly the entire parameter space. Denote by $x^{(k,g)}$ the $k-$th vector of the population at iteration (generation) g and by $x_j^{(k,g)}$ its $j-$th component. At the phases of the procedure, a new trial population is generated and each new individual is compared with its counterpart in the current population. The points with the best value of the objective function $f(x_1, ..., x_n)$ will "survive" to the next generation as they generally better than their actual counterparts.

At each iteration, the method generates a set of candidate points $y^{(k,g)}$ to substitute the elements $x^{(k,g)}$ of the current population, if $y^{(k,g)}$ is better. To generate $y^{(k,g)}$ two operations are applied: recombination and crossover.

A typical recombination operates on a single component $j \in \{1, ..., n\}$ by generating a new perturbed vector of the form $v_j^{(k,g)} = x_j^{(r,g)} + \gamma[x_j^{(s,g)} - x_j^{(t,g)}]$, where $r, s, t \in \{1, 2, ..., p\}$ are chosen randomly and $\gamma \in]0, 2]$ is a constant (eventually chosen randomly for the current iteration) that controls the amplification of the variation.

The potential diversity of the population is controlled by a crossover operator, that construct the candidate $y^{(k,g)}$ by crossing randomly the components of the perturbed vector $v_j^{(k,g)}$ and the old vector $x_j^{(k,g)}$ ($j_1, j_2, ..., j_h$ are random),

$$y_j^{(k,g)} = \begin{cases} v_j^{(k,g)} & \text{if } j \in \{j_1, j_2, ..., j_h\} \\ x_j^{(k,g)} & \text{if } j \notin \{j_1, j_2, ..., j_h\} \end{cases}$$

and the components of each individual of the current population are modified to $y_j^{(k,g)}$ by a given probability q. Typical values are $\gamma \in [0.2, 0.95]$, $q \in [0.7, 1.0]$ and $p \geq 5n$ (the higher p, the lower γ).

The candidate $y^{(k,g)}$ is then compared to the existing $x^{(k,g)}$ by evaluating the objective function at $y^{(k,g)}$: if better then substitution occurs in the new generation $g+1$, otherwise $x^{(k,g)}$ is retained.

Many variants of the recombination schemes have been proposed and some seem to be more effective than others (see [9] and [10], [4] for constraints handling).

To take into account the particular nature of our problem, we modify the basic procedure: start with the $(\alpha = 1) - cut$ back to the $(\alpha = 0) - cut$ so that the optimal solutions at a given level can be inserted into the "starting" populations of lower levels; use two distinct populations and perform the recombinations such that, during generations, one of the populations specializes to find the minimum and the other to find the maximum. For the details, the interested reader can consult reference [8].

4 Computational results

We have implemented the DE procedure using MATLAB and we have run a series of examples with randomly generated fuzzy data. We also have implemented two additional heuristic procedures based on the properties illustrated in section 2. DE will indicate the differential evolution algorithm illustrated in section 3; HeurLP will indicate the heuristic procedure illustrated at point 9. of section 2; Heur2 will indicate a heuristic procedure similar to HeurLP but, instead of solving the LP problem for each random initial feasible point, we apply the following procedure:

Step 1. start with a feasible $x^{(p)}$ and obtain $\hat{\beta} < 0$ such that $x^{(p)} - \hat{\beta}$ is feasible and substitute $x^{(p)}$ with $x^{(p)} - \hat{\beta}$;

Step 2. determine the maximum value of $t^+ > 1$ such that $t^+ x^{(p)}$ is feasible;

Step 3. iterate steps 1. and 2. while $\hat{\beta} < 0$ and $t^+ > 1$ are found (or for a given fixed number of iterations); at the final point $x^{(p)}$ the value of the variance is better (greater) then for the initial randomly generated point; repeat the procedure for a prefixed number of randomly generated initial feasible points and record the biggest found variance as an estimation for the maximum variance.

The fuzzy numbers for calculating the fuzzy variance are generated randomly to be triangular $\tilde{x}_i = \langle a_i, b_i, c_i \rangle$ with $\alpha - cuts$

$$[\tilde{x}_i]_\alpha = [a_i + (b_i - a_i)\alpha, c_i + (b_i - c_i)\alpha]$$

and $a_i < b_i < c_i$ obtained as follows (Matlab instructions, D is the dimension n of the problem):

$xmin = 5.0; xmax = 10.0;$

$dmin = 1.0; dmax = 4.0;\ rand('state', 0);$

$b = xmin + rand(1, D) * (xmax - xmin);$

$a = b - dmin - rand(1, D) * (dmax - dmin);$

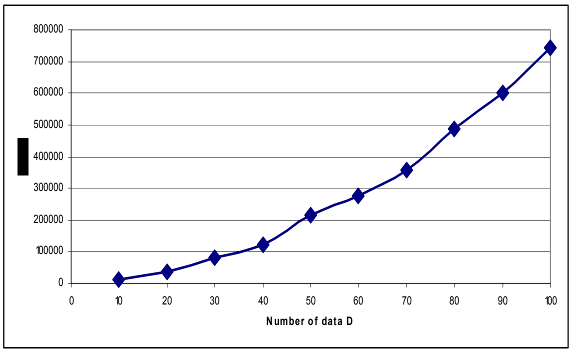
$c = b + dmin + rand(1, D) * (dmax - dmin);$

To test the procedures, we have generated 10 different problems with D from 10 to 100. For all algorithms we give the found fuzzy variance. For algorithm DE we report the number of function evaluations; for algorithm HeurLP we report the number of function evaluations and the number of 2-dimensional LP problems needed; for algorithm Heur2 we give the number of function evaluations and the total number of internal steps 1. and 2. In all the problems, the LU-fuzzy parametrization with $N = 4$ (five points) is used so that for each fuzzy variance four global maximization problem have been solved.

In the illustrated examples, the solution found by the DE algorithm was never worse then other procedures; HeurLP is better than Heur2 but its computational time is bigger.

D	F_{DE}	F_{HeurLP}	N_{LP}	F_{Heur2}	N_{Iter}
10	13750	905	1449	505	1200
20	36500	995	1612	995	1200
30	81750	1085	1638	1485	1400
40	121000	1175	1744	1975	1400
50	216250	1265	1792	2465	1600
60	274500	1355	1763	2955	1600
70	358750	1445	1772	3445	1800
80	488000	1545	1945	3935	1800
90	600750	1625	1955	4425	2000
100	742500	1715	1962	4915	2000

It appears that good near-optimal solutions require a large number of function evaluations or of solutions to auxiliary LP problems; but increasing the dimension D of the number of fuzzy data, the number of function evaluations required by the DE procedure will increase less then quadratically, as illustrated by the following graph:

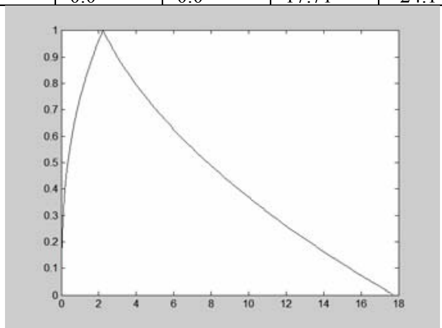


In the following tables, the LU-fuzzy parametrization of the found solutions is reported only for the first (core) and the last (support) $\alpha - cuts$;

the plots are obtained by the full $N = 4$ parametrization.

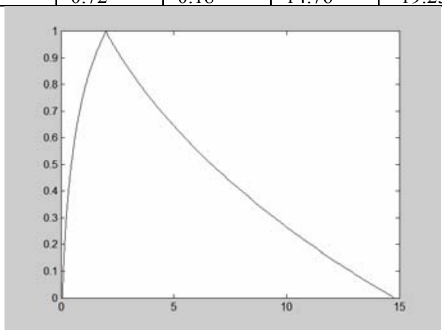
Problem 1 (D=10)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	2.28	6.08	2.28	-6.82
0.0	0.0	0.0	17.71	-24.14



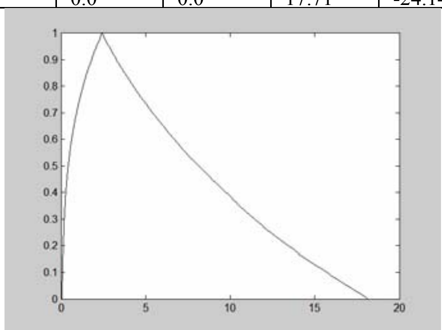
Problem 2 (D=20)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	1.99	5.65	1.99	-6.23
0.0	0.72	0.18	14.76	-19.25



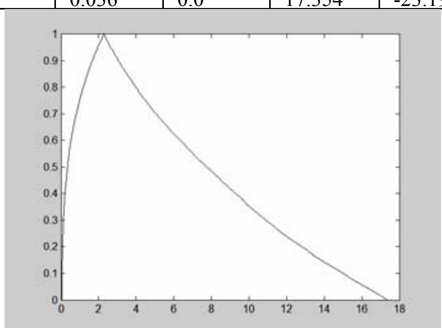
Problem 3 (D=30)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	2.23	6.03	2.23	-6.82
0.0	0.0	0.0	17.71	-24.14



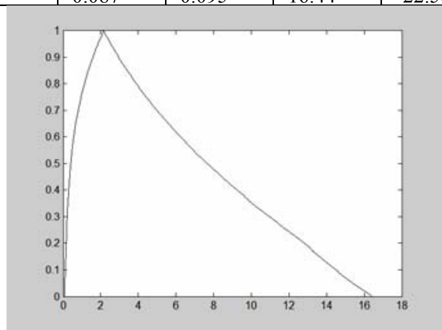
Problem 4 (D=40)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	2.284	6.78	2.284	-7.08
0.0	0.036	0.0	17.354	-23.19



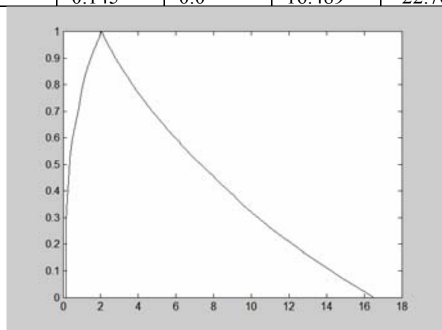
Problem 5 (D=50)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	2.17	6.501	2.17	-7.061
0.0	0.087	0.095	16.44	-22.58



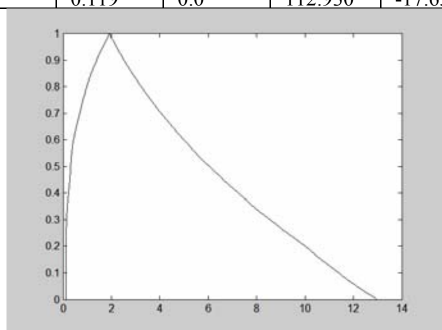
Problem 6 (D=60)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	2.058	5.955	2.058	-6.648
0.0	0.145	0.0	16.489	-22.762



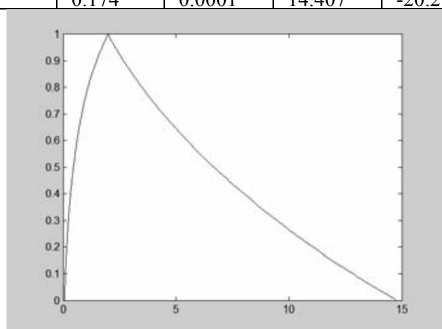
Problem 7 (D=70)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	1.933	6.006	1.933	-5.318
0.0	0.119	0.0	112.930	-17.634



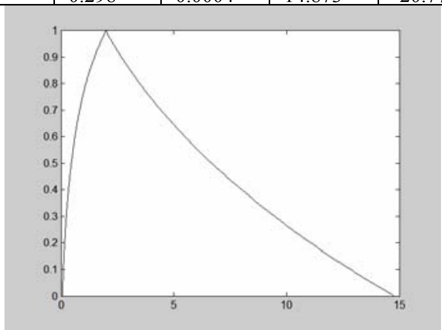
Problem 8 (D=80)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	1.779	5.631	1.779	-5.647
0.0	0.174	0.0001	14.407	-20.21



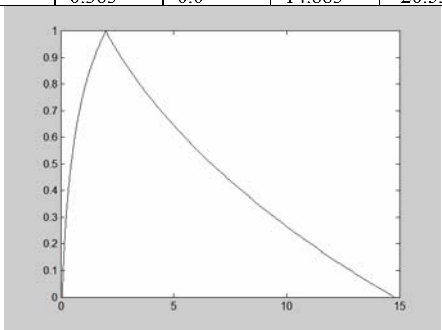
Problem 9 (D=90)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	1.859	5.549	1.859	-6.82
0.0	0.298	0.0004	14.873	-20.779



Problem 10 (D=100)

α	w_α^-	δw_α^-	w_α^+	δw_α^+
1.0	1.985	6.053	1.985	-5.971
0.0	0.363	0.0	14.883	-20.53



5 Conclusions

We can conclude that the DE methods, whose nice behaviour in many hard optimization problems is still appreciated ([5],[9],[10]), seems to be a promising tool (and easy to implement as in [7],[8]) for applications in the interval and fuzzy contexts.

In particular, it appears that they can benefit of some properties of the fuzzy variance problem (e.g. quadratic homogenity and diagonal translation invariance) to facilitate the finding of good solutions.

Further work is planned to find improved DE strategies (e.g. multi populations and constraint handling [4]).

References

[1] D. Dubois, H. Fargier, J. Fortin, "A generalized vertex method for computing with fuzzy intervals", in Proceedings of the IEEE International Conference of Fuzzy Systems, Budapest, 2004, 541-546.

[2] S. Ferson, L. Ginzburg, V. Kreinovich, L. Longpré , M. Aviles, "Computing variance for interval data in NP-hard.", ACM SIGACT News, 33 (2002) 108-118.

[3] V. Kreinovich, H.T. Nguyen, B. Wu, "On line algorithms for computing mean and variance of interval data and their use in intelligent systems", Information Sciences, 177 (2007) 3228-3238.

[4] K. Deb, "An efficient constraint handling method for genetic algorithms", *Computer methods in applied mechanics and engineering*, 186, 2000, 311-338.

[5] K. Price, "An introduction to differential evolution", in D. Corne, M. Dorigo, F. Glover (Ed.), *New Ideas in Optimization*, McGraw Hill, 1999, 79-108.

[6] L. Stefanini, L. Sorini, M. L. Guerra, "Parametric representation of fuzzy numbers and application to fuzzy calculus", *Fuzzy Sets and Systems*, 157, 2006, 2423-2455.

[7] L. Stefanini, "Differential Evolution Methods for the Fuzzy Extension of Functions", Working Paper Series EMS n. 103, 2006; available online at the *RePEc* web page, *http://ideas.repec.org/f/pst233.html*.

[8] L. Stefanini, "A Differential Evolution Algorithm for the Fuzzy Variance", Working Paper Series EMS n. 112, University of Urbino, 2008; in preparation; it will be available online at the *RePEc* web page, *http://ideas.repec.org/f/pst233.html*.

[9] R. Storn, K. Price, "Differential Evolution: a simple and efficient heuristic for global optimization over continuous spaces", ICSI technical report TR-95-012, Berkeley University, 1995. Also, *Journal of Global Optimization*, 11, 1997, 341-359.

[10] R. Storn, "System design by constraint adaptation and differential evolution", *IEEE Transactions on Evolutionary Computation*, 3, 1999, 22-34.

[11] Vandenbussche, D., Nemhauser, G.: A branch-and-cut algorithm for nonconvex quadratic programs with box constraints. Math. Program. 102(3), 559–575 (2005b)