UNCERTAINTY AND FUZZY SETS: CLASSIFYING THE SITUATION

Volodymyr Donchenko

Abstract: The so called "Plural Uncertainty Model" is considered, in which statistical, maxmin, interval and Fuzzy model of uncertainty are embedded. For the last case external and internal contradictions of the theory are investigated and the modified definition of the Fuzzy Sets is proposed to overcome the troubles of the classical variant of Fuzzy Subsets by L. Zadeh. The general variants of logit- and probit- regression are the model of the modified Fuzzy Sets. It is possible to say about observations within the modification of the theory. The conception of the "situation" is proposed within modified Fuzzy Theory and the classifying problem is considered. The algorithm of the classification for the situation is proposed being the analogue of the statistical MLM(maximum likelihood method). The example related possible observing the distribution from the collection of distribution is considered

Keywords: Uncertainty, Fuzzy subset, membership function, classification, clusterization.

ACM Classification keywords: I.5.1.Pattern Recognition: Models Fuzzy sets; G.3. Probability and Statistics: Stochastic processes; H.1.m. Models and Principles: miscellaneous

Introduction

Classical conception of Fuzzy Subsets (ClasFsS) proposed Lotfi Zadeh [Zadeh, 1965] (see methodological view in [Kaufmann, 1982]) seemed to propose the practice a new method to manipulate with uncertainty. This method from the very beginning considered to be alternative to the ones already had been in use that time: statistical, maxmin, interval. The proposition to take into account and formalize the idea of the intermediate, transitional domains between "crisp" alternatives was the essence of the ClasFsS.

As it was mentioned earlier from the very beginning the ClasFsS considered by it founders on one hand as being alternative to those had been being in use by the moment of birth of the theory("isolationism"), and on the other hand as the theory alternative to the classical, "crisp", set theory. Particularly, ClasFsS considered having nothing mutual with the statistical methods. Both of these pretensions seem to be symptoms of the theory coming into being.

Indeed, as relating set theory ClasFsS is in the "naive" faze and nothing like "crisp" set theory axioms there exit. As to the last, in the [Donchenko, 2005] (see also [Donchenko, 1998-3], [Donchenko, 1998-4], also [Donchenko, 2004]) the attempt is represented to propose some form of the abstraction axiom. Also, the full absence of somewhat that one may be nominated to be logical calculus characterizes the situation in ClasFsS now, though there are attempts to say about "Fuzzy logic". But the one called "Fuzzy logic" is simply analogue of "crisp" propositions calculus [Kaufmann, 1982] see also [Kaufmann and Gupta, 1991]). It is reasonably to note opportunely, that functions of this "fuzzy calculus" are not full in the space of all function on "fuzzy propositions" as it is in Boolean algebra.

Also, pretensions of ClasFsS to be alternative to statistical method which namely is the way to investigate uncertainty through the frequencies of results, are not fruitful for the ClasFsS because all the power of statistical methodology including interpretation turned out to be cut off. Statistical interpretation is natural for a membership function [Donchenko, 2005]. Besides, the considerations of that work demonstrating the gap in the object of uncertainty in the definition of ClasFsS pointed out the way to make the definition of ClasFsS correct in this aspest. This modification of the ClasFsS embodied in conception of Modified Fuzzy Sets (MoFS) in the paper has been already referred to [Donchenko, 2005] and earlier publications [Donchenko, 1998-3], [Donchenko, 1998-4], also [Donchenko, 2004]). Namely, in MoFS classical membership function $\mu(e), e \in E$ becomes to be a function of two arguments: from $e \in E$ and from "crisp" predicate $P \in \wp$ – or correspondent "crisp" set. In such variant of determination "property $P \in \wp$ is described fuzzy way by the modified membership function $\mu^{(P)}(e), e \in E, P \in \wp$" with preserving general restriction $\mu^{(P)}(e) \in [0,1], e \in E, P \in \wp$, P being a parameter.

Definition. By the Modified Fuzzy Set (MoFS) we will mean the pair (E, $\mu^{(P)}(e)$, $P \in \wp$), where E – abstract set (supporter) and $P \in \wp$ with the $\wp$ - the set of "crisp" predicate on abstract "crisp" universal set U_P (or correspondent "crisp" subset of U_P).

MoFS Model Example: generalized variants of logit- and probit-regression

Logit- and probit- regressions and its Generalized variants are the best example of MoFS. As it is known, in these variants of the regressions are the dependence of the Bernoulli distribution parameter on the vector $e \in E = R^m$ is considered. This dependence has the next $\beta \in R^m$ parameterization:

$$P_e\{Y = 1\} = G(e^T \beta),$$

where Y - Bernoulli - random variable, G(z) appropriate known distribution function or the tail of the distribution.

Parameter $\beta \in R^m$ is to be estimated via observations (sample) $(e_1, y_1),...,(e_n, y_n)$:

$e_i \in R^m, i = \overline{1,n}$ - non-random,

$y_i \in \{0,1\}, i = \overline{1,n}$ - the values (realizations) of independent Bernoulli - random variable $Y_i, i = \overline{1,n}$ with the parameters, correspondingly:

$$P_{e_i}\{Y_i = 1\} = G(e_i^T \beta),\ i = \overline{1,n}.$$

In this model case:

- $E = R^m$;
- P as a predicate defined by the relation: $P = \{Y = 1\}$;
- $\mu^{(P)}(e) = G(e^T \beta)$.

Evidently, $\mu^{(P)}(e)$ is parameterized by $\beta \in R^m$.

General character of the MoFS model example: statistical interpretation of the MoFS

In the paper [Donchenko, 2005] (see also [Donchenko, 1998-3], [Donchenko, 1998-4]) two theorems have been proved made possible statistical interpretation of the ClasFsS and MoFS . These theorems, formulated for discrete and non discrete supporters E are the next.

Theorem 1. For any finite collection of the ClasFsS $(E, \mu_{\underline{A_i}}(e))$, $i = \overline{1,n}$ with the one and the same supporter E one may find discrete probability space (Ω, B_Ω, P), collection of the evens $A_i \in B_\Omega, i = \overline{1,n}$ and complete collection of the events He : $H_e \in B_\Omega$, e∈E, – within this probability space, that all of the membership functions $\mu_{\underline{A_i}}$, $i = \overline{1,n}$ may be represented as the systems of conditional probabilities in the next way:

$$\mu_{\underline{A_i}}(e) = P\{A \mid H_e\}, \text{ for any } e \in E,\ i = \overline{1,n}.$$

Theorem 2. Given the:

- (E, $\Im$, m)- the space with a measure;
- $(E, \mu_{\underline{A_i}}(e))$, $i = \overline{1,n}$ μ_i(e), i > 0 a collection on Fuzzy subsets with the equal supporters E;
- all of the membership functions $\mu_{\underline{A_i}}(e)$, $i = \overline{1,n}$ are $\Im$, £, measurable (£– Borel σ-algebra on R^1),

then:

- exist probability space (Ω, B_Ω, P),
- exist ξ discrete random S_p – valued random variable on (Ω, B_Ω, P) with S_p is n-element set with elemens say S_i, $i = \overline{1,n}$;
- exist η random E – valued random variable on (Ω, B_Ω, P)

such, that for any $i = \overline{1,n}$

$\mu_{\underline{A_i}}(e)$ = P{ ξ = S_i | η = e },

where P{ξ = S_i | η } – conditional distribution of r.v. ξ respectively r.v.. η.

Remark 1. Both of the theorems demonstrate, that for ClasFsS exist latent “crisp” predicates-events (or correspondent “crisp”sets): $A_i \in B_\Omega$, $i = \overline{1,n}$ for the first case and $\{\xi_i = S_i\} \in B_\Omega$, $i = \overline{1,n}$, - which are characterized in a Fuzzy way. For the MoFS variants, as it has been demonstrated by the logit-, probit- model example, these predicates are presented in the modified definition patently.

MoFS and MoFS model example in the context of plural models of uncertainties

When saying about ClasFsS or MoFS role in uncertainty description it is interesting to create “general platform” in which basic theory for uncertainty manipulating can take their own places. It is likely the so call “plural” model of uncertainties to be such platform.

“Plural” uncertainties model start from the conception of “observation” and “observation situation”.

When saying “observation situation” we mean “conditions” (denoted by κ) plus “observation” by itself. In its turn “conditions” in the “observation situation” consist of “varying part” (denoted by x) and on default part (denoted by f). As to “observation” then compulsory part of them is the response y on the conditions κ. But there is no precise meaning of the “observation”. To put it more precisely “observation” may be interpreted in the next three variants:

$$\text{“observation”} = \begin{cases} y \\ (x, y) \\ (\kappa, y) \end{cases} .$$

It is necessary to say, that standard meaning of a sequence of “observations”(real or virtual) is $(x_1, y_1),...,(x_N, y_N)$ while it is necessarily to be $(\kappa_1, y_1),...,(\kappa_N, y_N)$.

__Definition__. “Plural” model of uncertainties is the model, based on response y plurality in the sequence of “observations”: real or virtual.

So, when saying about uncertainties one have to answer himself what content of “observation” is in use.

Indeed, plurality in y take place in deterministic case when “conditions” are of the form $\kappa_1 = (x_1, f),...,\kappa_N = (x_N, f)$, but “observations” treat to be $y_1,..., y_N$ .This case may be interpreted as the “latent” parameter case, as it qualified in physics.

Plurality in y under just and the same $\kappa \equiv \kappa_i, i = \overline{1,N}$ is the object of application of statistical method. As it is well known all result observing or may be observed in observations are described and the frequencies (may be its limits: probabilities) of the results or collections of the results (events) are considered.

Classical regressions illustrates the application of the statistical methods when there is plurality in y in observations with common κ , when response y is real and variability in x take place. Application of Least Square Method (LS) or its modifications are common in this case.

Generalized variants of ligit- and probit-regressions illustrates the application of the statistical methods when there is plurality in y in observations with common κ , when response y i is binary: $y \in \{0,1\}$ - and variability in x take place. Application of MLM(Maximum Likelihood Method) characterize this case of plurality in observations of binary response y.

And, at last, in minmax approach varying part of the conditions is considered to consist of two parts:

$$x_i = (x_i^{(1)}, x_i^{(2)}) \ , i = \overline{1,N},$$

while each of observations treats to be $(x_i^{(1)}, y_i), i = \overline{1,N}$, or more precisely $((x_i^{(1)}, f), y_i), i = \overline{1,N}$, with $x_i^{(2)} \in X_2$, where X_2 is known.

MoFS approach or modification of the ClasFsS within the plural model of uncertainties is interpreted just as the generalized logit- and probit- regression. Indeed, accordingly to basic statistical interpretation theorem from [Donchenko, 2005] membership function $\mu^{(P)}(e), e \in E$ has the uncertainty object P. So, MoFS observations may be treated as

$$(e_i, y_i), i = 1, N : e_i \in E, y_i = \begin{cases} 1, \text{when P is observed} \\ 0, \text{when P is not observed} \end{cases}.$$

Collections of the MoFS: situation

MoFS definition of the a fuzzy set as a pair $(E, \mu^{(P)}(\cdot)), \mu : E \to [0,1]$ with the membership function $\mu^{(P)}(\cdot)$ being the function of two arguments with one of them fixed ($P \in \wp$) let the problem of classification or clusterization to be considered: ascription each of elements $e \in E$ to the one of the K classes, described by predicates $P_k \in \wp, k = \overline{1,K}$ from a MoFS collection $(E, \mu_k^{(P_k)}(\cdot)), P_k \in \wp, k = \overline{1,K}$ with $\mu_k^{(P_k)}(\cdot) : \sum_{k=1}^{K} \mu_k^{(P_k)}(e) \stackrel{\le}{\ge} 1,$

such collection may be complete:

$$\forall e \in E \sum_{k=1}^{K} \mu_k^{(P_k)}(e) = 1, \tag{1}$$

as well as incomplete: with strong inequality.

Predicates collection $P_k \in \wp, k = \overline{1,K}$ from MoFS collection $(E, \mu_k^{(P_k)}(\cdot)), P_k \in \wp, k = \overline{1,K}$ may be interpreted as a collection of the alternatives which may take place for each of the elements of $e \in E$ with some probabilities, described by $H\,\mu_k^{(P_k)}(e)), e \in E, k = \overline{1,K}$.

In the ClasFsS the values of membership functions may be considered classically: as a confidence (certainty) functions. Actually, in this case the collection (list) of the objects of uncertainty (alternatives) are to be described, we will say, that MoFS collection $(E, \mu_k^{(P_k)}(\cdot)), P_k \in \wp, k = \overline{1,K}$, describe the situation for the elements $e \in E$ or the situation made concrete by $e \in E$. Ascription the $e \in E$ under consideration to the one of K classes, described by predicates $P_k \in \wp, k = \overline{1,K}$, we will call the classifying the situation.

Classifying the situation conception

When interpreting MoFS collection as a situation, each of the $e \in E$ make the situation concrete while membership functions $\mu_k^{(P_k)}(e), k = \overline{1,K}$ for each of the fix $e \in E$ evince the “degree of appearance” for the $P_k \in \wp, k = \overline{1,K}$. Such conception of the situation make it possible to estimate the situation for that or this $e \in E$ by the maximum (or minimum) of the confidence in each of $P_k \in \wp, k = \overline{1,K}$ for fixed $e \in E$ and correspondingly to make ascription the element $e \in E$ to one of the classes, determined by $P_k \in \wp, k = \overline{1,K}$. Such approach to classifying the situation namely realizes the idea embodied in MLM (Maximum Likelihood Method). It is reasonable to remark that in MLM this idea is realized in a posteriori form: when having the observations.

Definition. Function $\hat{P}(e), e \in E, \hat{P}: E \to \{P_1,...,P_K\}$, determined on the common E of the MoFS collection $(E, \mu_k^{(P_k)}(\cdot)), P_k \in \wp, k = \overline{1,K}$ (situation) by the relation:

$$\hat{P}(e) = P_{k^*}, k^* = \arg\max_{k=\overline{1,K}} \mu_k^{(P_k)}(e), e \in E \tag{2}$$

said to be situation estimation for $e \in E$.

Remark 2. Generally speaking situation estimation may be plural: when maximum in (1) reached simultaneously for some k $k \in \{1,...,K\}$. In this case $\hat{P}(e), e \in E$ turned out to be plural: $\hat{P}(e) \subseteq \{P_1,...,P_K\}, e \in E$ and determined by the modification of the relation (2):

$$\hat{P}(e) = \{P_k : k \in K^* = \operatorname{Arg}\max_{k=\overline{1,K}} \mu_k^{(P_k)}(e), e \in E\}. \tag{3}$$

Remark 3.Just as in (1) or (2) situation is estimated to be “best”, it may be estimated to be worst. In this case “min” is substituted instead “max” in (1) or (2).

Remark 4. Term “situation estimation” by no means do not restrict “classification” character of the $\hat{P}(e), e \in E$, when $P_k, k = \overline{1,K}$ define classes and $\mu_k^{(P_k)}(\cdot)), k = \overline{1,K}$ define the membership probabilities for each $e \in E$.

In this case $\hat{P}(e), e \in E$ said to be classifying function. Note that classes are not necessarily mutually exclusive. Under exclusive alternatives the condition (1) is natural.

Remark 5. Situation estimation according (1) or (2) with “max” or “min” demonstrates that when operating with “fuzzy logic” operations it is reasonable to take into account the arguments on which the result of operation is reached but no the result of the operation by itself.

Situation Model example: clustering the distributions, probe sets (a priori data)

The classification problem for several distributions demonstrates the “classifying situation” approach proposed earlier. More precisely, let observed $e \in R^m$ may represent one of the K distributions $P^{(k)}(B)$,B –Borel set in $R^m, k = \overline{1,K}$ or, equally, may be the value of the one of the random (multivariate) variable (r.v.) $\varepsilon_k, k = \overline{1,K}$: $P^{(k)}(B) = P\{\varepsilon_k \in B\}, k = \overline{1,K}$,B – Borel set in R^m. One can consider the values of the distributions on “probe” sets $e + \pi$, $e \in E$, for appropriate fixed Borel π. As a π ta ball $S_\rho(0)$ for the fixed radius $\rho > 0$ can be considered. Obviously $\mu_{k,\rho}^{(P_k)}(e), e \in R^m, k = \overline{1,K}$, determined by the relations

$$\mu_{k,\rho}^{(P_k)}(e) = P^{(k)}(S_\rho(e)) = P^{(k)}(e + S_\rho(0)) = P\{\varepsilon_k \in e + S_\rho(0)\}, k = \overline{1,K}, e \in R^m, \tag{4}$$

are membership functions for MoFS with the “crisp”-predicates P_k : “to have distribution $P^{(k)}$ ”. $k = \overline{1,K}$.

Also π one can be chosen to be symmetric, close convex set $V_\rho(0)$ with fixed radius $\rho > 0$: $e + \pi = e + V_\rho(0)$, $e \in E$. Last set denoted to be $V_\rho(e)$:

$$V_\rho(e) = e + V_\rho(0)\,. \qquad (5)$$

So, the collection $(R^m, \mu_{k,\rho}^{(P_k)}(e)), k = 1, K$ is the MoFS collection with the "crisp" predicates P_k :" property to have distribution $P^{(k)}$", $k = \overline{1, K}$,- which describes the situation.

The balls $S_\rho(e) = e + S_\rho(0)$, $\rho > 0$ centered in $e \in R^m$ in a natural way probe the distributions. The results are represented by the MoFS membership functions $\mu_{k,\rho}^{(P_k)}(e)), k = 1, K$.

Remark 6. It is advisable to say that in the example under consideration memberships functions of the MoFS collection, which describe the situation are obviously statistically transparent. Ana in a addition standart statistical representation accordingly to theorem 2 above from for example [Donchenko, 2005] turned out to be of the next form:

$$\mu_{k,\rho}^{(P_k)}(e)) = P\{\varepsilon_k \in S_\rho(\xi) \,|\, \xi = e\}, k = \overline{1, K} \qquad (6)$$

With R^m -valued vector fandom variable (r.v.) ξ determined on the probability space common with the collection $\varepsilon_k, k = \overline{1, K}$, independent of them and having the distribution to be nonsingular.

Indeed, we have:

$$P\{\varepsilon_k \in S_\rho(\xi) \,|\, \xi\} = M\{\chi_{S_\rho(\xi)}(\varepsilon_k) \,|\, \xi\} = M\{\chi_{S_\rho(0)}(\varepsilon_k - \xi) \,|\, \xi\} =$$

$$= M\{\chi_{S_\rho(0)}(\varepsilon_k - e)\Big|_{e=\xi} = P\{\varepsilon_k - e \in S_\rho(0)\}\Big|_{e=\xi} = P\{\varepsilon_k \in S_\rho(e)\}\Big|_{e=\xi}\,.$$

So, indeed, relation (6) gives the example of the standard (universal) statistical representation for collection of the ClasFsS or MoFS membership function accordingly to theorem 2 from the [Donchenko, 2005] has been already mentioned above.

Turning back to the MoFS membership functions (6) from the collection, defined the situation, we note that $\hat{P}(e), e \in R^m$ from (2) or (3) denoted as $\hat{P}_\rho(e), e \in R^m$ below are defined by one of the next two relations:

$$\hat{P}_\rho(e) = P_{k*} : k^* = \arg\max_{k=\overline{1,K}} \mu_{k,\rho}^{(P_k)}(e), e \in R^m\}\,, \qquad (7)$$

$$\hat{P}_\rho(e) = \{P_k : k \in K^* = \mathrm{Arg}\max_{k=\overline{1,K}} \mu_{k,\rho}^{(P_k)}(e), e \in R^m\}\,. \qquad (8)$$

So, classifying the situation accordingly to (7) or (8) obviously turned out to be the "maximum likelihood" estimation by P_k - probabilities for ρ - neighborhood of $e \in R^m$.

Remark 7. It is interesting to note that alternatives in (8) are not exclusive.

Situation Model example, continue: limit by the infinitive decreasing the measures of probe sets

Classifying the situation accordingly (7) or (8) shows transparent statistical (probabilistic) content. It is classification by maximum probability of the distributions collection under consideration on the "probe" set, neighboring $e \in E$. The items relating the classifying problem under "decreasing" the the "probe" set. It is naturally to normalize the membership functions in some way. It is turned out, that the classifying situation algorithm reduces the problem to the classification by the maximum of density functions for the distribution under consideration.

Indeed, let the distributions $P_k, k = 1, K$ have the continues densities $h_k(z), z \in R^m, k = \overline{1,K}$, and the "size" ρ of the "probe" sets $S_\rho(e) = e + S_\rho(0)$ or $V_\rho(e) = e + V_\rho(0)$ is infinitely decreasing: $\rho \to 0$. As it has been mentioned about membership functions are to be normalized in some way. There are two variants for normalization: one for m=1 and another for the general case: for m>1. It is $\rho^{-1}, \rho > 0$ in the first case (m=1) and inverse of Lebeague measure ϑ in R^m on the probe set in another (m>1). It is naturally to demand of "non singularity" for the density functions $h_k(z), k = 1, K$ of distributions $P_k, k = 1, K$."Non singularity" is understudied as non zero value of the distributions on the ""probe" sets for all $\rho > 0$.

Then the next pairs of statements take place.

Theorem 3. Let the distribution functions $h_k(z), z \in R^m, k = \overline{1,K}$ of distributions $P_k, k = 1, K$ are continuous and non singular.
Then

$$\lim_{\rho\to 0} \rho^{-1}\mu_{k,\rho}^{(P_k)}(e)) = h_k(e) \| \mathrm{grad}_z h_k(e) \|, e \in R^1, k = \overline{1,K}, \quad (9)$$

$$\lim_{\rho\to 0} \{\vartheta(S_\rho(0))\}^{-1}\mu_{k,\rho}^{(P_k)}(e)) = h_k(e), e \in R^m, m > 1,\ k = \overline{1,K}. \quad (10)$$

Similar result take place for MoFS membership functions built by the "probe" sets $V_\rho(e) = e + \rho V$ with symmetric, convex, close set with radius equal to one, i.e. with the collection of the MoFS membership functions of the next type:

$$\mu_{k,\rho}^{P_k}(e) = P^{(k)}(V_\rho(e)) = P^{(k)}(e + \rho V) = P\{\varepsilon_k \in e + \rho V\}, e \in R^m, k = \overline{1,K}. \quad (11)$$

Theorem 4. Let conditions of theorem 3 take place and $P_k(V_\rho(e)) > 0, k = 1, K$. Then there is exist $\varphi : 0 < \varphi \le 1$ such, that:

$$\lim_{\rho\to 0} \rho^{-1}\mu_{k,\rho}^{(P_k)}(e)) = \varphi h_k(e) \| \mathrm{grad}_z h_k(e) \|, e \in R^1, k = \overline{1,K}, \quad (12)$$

$$\lim_{\rho\to 0} \{\vartheta(\rho V_1(0)))\}^{-1}\mu_{k,\rho}^{(P_k)}(e)) = h_k(e), e \in R^m, m > 1,\ k = \overline{1,K}. \quad (13)$$

Classifying the situation: "a priori maximum likelihood method"

Relations (9)-(10), (12)-(13) shows straight connection between situation classifying in the model example and classification by the density functions, namely, between classifying by the probabilities of distributions $P_k, k = 1, K$ on the "probe" sets and classification by density functions of correspondent distributions.

So, let for the MoFS collection $(R^m, \mu_{k,\rho}^{(P_k)}(\cdot)), k = \overline{1,K}$ there exist limits $\lim_{\rho\to 0} \rho^{-1}\mu_{k,\rho}^{(P_k)}(e))$, denote it by $d_k(e)$, for $E = R^1$ or limit $\lim_{\rho\to 0} \{\vartheta(\rho V)\}^{-1}\mu_{k,\rho}^{(P_k)}(e))$ also denoted by $d_k(e)$, for some symmetric, convex close set V with radius equal 1.

Definition. Function $\hat{P}(e), e \in E, \hat{P} : E \to \{P_1,..., P_K\}$, defined by MoFS collection $(R^m, \mu_{k,\rho}^{(P_k)}(\cdot)), k = \overline{1,K}$, by the one of the relations:

$$\hat{P}_{\infty}(e) = P_{k^*}, k^* = \arg\max_{k=\overline{1,K}} d_k(e), e \in R^m, \quad (14)$$

$$\hat{P}_{\infty}(e) = P_{k^*}, k^* = \arg\max_{k=\overline{1,K}} d_k(e) \| grad_z h_k(e) \|, e \in R^m, \quad (15)$$

said to be classifying the situation by the limit of normalized membership functions.

Model example shows the interpretation of the classifying algorithm represented by (14) or (15).

Theorems 3-4 show the relation between maximum probabilities classification and classifying the situation for the MoFS collection on the model example. It is reasonable to recollect here similar classification procedure; main-shift classification (see, for example, [Comaniciu, 2002]). In the mean-shift algorithm classification s implemented by the density functions generated for each of the classes by the observations of learning sample, i.e. that one may be called "a posteriori maximum likelihood method".

Conclusion

In the paper modified variant of fuzziness from [Donchenko, 2005] (see also [Donchenko, 1998-3], [Donchenko, 1998-4] and [Donchenko, 2004]), is considered to develop the conception proposed by the ideas of "situation" and "classifying the situation". The algorithm for "classifying situation" is proposed having natural probability content. This probability content is illustrated by the model example based on the idea of "probe" sets. Limit behavior of the objects in the model example establish the relation between "classifying the situation" and "a priori likelihood method". "A priori likelihood method" is proposed to be extended on "classifying situation" construction. Similar classification algorithm: mean-shift method – is recollected having the same idea, but a posteriori character.

References

[Comaniciu, 2002] .D. Comaniciu. A robust approach towards Feature space analysis.//IEEE Transactions on Pattern Analysis and Machine Intelligence. – V.24,№5, May 2002.– p. 603-617.

[Donchenko, 1998-3] Donchenko V.S. Conditional distributions and Fuzzy sets. // Bulletin of Kiev University. Series: Physics and Mathematics, №3, 1998. (In Ukrainian)

[Donchenko, 1998-4] Donchenko V.S. Probability and Fuzzy sets. // Bulletin of Kiev University. Series: Physics and Mathematics, №4,1998. (In Ukrainian)

[Donchenko, 2004] Donchenko V.S. Statistical models of observations and Fuzzy sets. // Bulletin of Kiev University. Series: Physics and Mathematics, №1, 2004. (In Ukrainian)

[Donchenko, 2005]. Donchenko V. Fuzzy sets: Abstraction axiom, Statistical Interpretation. Observations of Fuzzy Sets. //International Journal "Information Theories and Applications".– V.13, №3.– 2005.–p.233-239.

[Kaufmann, 1982]. Introduction to Fuzzy subset theory [Рус. Перевод Кофман А. Введение в теорию нечетких множеств.- Г.: Радио и связь. 1982.- 322 с.

[Kaufmann and Gupta, 1991] Arnold Kaufmann and Madan Gupta. Introduction to Fuzzy Arithmetic. - Thomson Computer Press, 1991.

[Zadeh, 1965] . Zadeh, Lotfi, Fuzzy Sets.// Information and Control, 8(3). June 1965. pp. 338-353.

Authors' Information

Donchenko, Volodymyr – Professor, Kyiv National Taras Shevchenko University, Cybernetics Faculty, Systems analysis and Decision Making Department; e-mail: vsdon@unicyb.kiev.ua, voldon@unicyb.kiev.ua.