

GAUSSIAN MEMBERSHIP FUNCTIONS ARE MOST ADEQUATE IN REPRESENTING UNCERTAINTY IN MEASUREMENTS

V. Kreinovich¹, C. Quintana¹, L. Reznik²

¹Computer Science Department, University of Texas at El Paso, El Paso, TX 79968 USA

²Sikeirosa 15-1-46, St. Petersburg 194354, Russia

Abstract. In rare situations like fundamental physics we perform experiments without knowing what their results will be. In the majority of real-life measurement situations, we more or less know beforehand what kind of results we will get. Of course, this is not the precise knowledge of the type “the result will be between $a - \delta$ and $a + \delta$ ”, because in this case, we would not need any measurements at all. This is usually a knowledge that is best represented in uncertain terms, like “perhaps (or “most likely”, etc.) the measured value x is between $a - \delta$ and $a + \delta$ ”.

Traditional statistical methods neglect this additional knowledge and process only the measurement results. So it is desirable to be able to process this uncertain knowledge as well. A natural way to process it is by using fuzzy logic. But there is a problem: we can use different membership functions to represent the same uncertain statements, and different functions lead to different results. What membership function to choose?

In the present paper, we show that under some reasonable assumptions, Gaussian functions $\mu(x) = \exp(-\beta x^2)$ are the most adequate choice of the membership functions for representing uncertainty in measurements. This representation was efficiently used in testing jet engines for airplanes and spaceships.

1. INTRODUCTION

Usually in measurement situations there is some prior knowledge. In rare situations like fundamental physics we perform experiments without knowing what their results will be. In the majority of real-life measurement situations, we more or less know beforehand what kind of results we will get. Of course, this is not the precise knowledge of the type “the result will be between $a - \delta$ and $a + \delta$ ”, because in this case we would not need any measurements at all. This is usually a knowledge that is best represented in uncertain terms, like “perhaps (or “most likely”, etc.) the measured value x is between $a - \delta$ and $a + \delta$ ”.

Traditionally the uncertain prior knowledge is not used in measurement processing. Traditional statistical methods neglect this additional knowledge and process only the measurement results. So it is desirable to be able to process this uncertain knowledge as well.

The usage of fuzzy logic and related problems. A natural way to process uncertainty is by using fuzzy logic [Z65]. This way we represent every statement of the type “most likely, $|x - a| \leq \delta$ ” by a membership function $\mu(x)$ that for each x gives us a degree to which we are certain that this particular x satisfies the given condition. But there is a problem: we can use different membership functions to represent the same uncertain statements. What membership function to choose?

What we are planning to do? In the present paper, we show that under some reasonable assumptions, Gaussian functions $\mu(x) = \exp(-\beta x^2)$ are the most adequate choice of the membership functions for representing uncertainty in measurements. This representation was efficiently used in testing jet engines for airplanes and spaceships.

2. MOTIVATION OF THE FOLLOWING DEFINITIONS

We must have in mind that different experts can have different opinions. Therefore the final resulting knowledge about the value of a physical quantity does not consist of a single

statement, but can be formed by adding several statements of several expert, e.g., “most likely, $|x - a_1| \leq \delta_1$ ”, “most likely, $|x - a_2| \leq \delta_2$ ”, ... The resulting statement is “most likely, $|x - a_1| \leq \delta_1$, and most likely, $|x - a_2| \leq \delta_2, \dots$ ” In order to represent this resulting knowledge we must choose some operation $*$ for $\&$. Then the resulting membership function will be equal to $\mu(x) = \mu_1(x) * \mu_2(x) * \dots$, where $\mu_i(x)$ corresponds to the opinion of i -th expert.

What $\&$ -operation to choose? Experimental results given in [HC76], [O77], and [Z78], show that among all possible operations $a, b \rightarrow \min(a, b)$ and $a, b \rightarrow ab$ are the best fit for human reasoning.

The \min operation does not seem to be adequate for our purposes, because if we use \min , then, e.g., the degree, to which a function $x(t)$ satisfies the condition “for all t , most likely $|x(t)| \leq M$ ”, is equal to the minimal of the degrees of the statements “most likely, $|x(t)| \leq M$ ” for all t . This minimum is attained when the value of $|x(t)|$ is the biggest possible. Therefore, the function $x_1(t)$ that is everywhere equal to $2M$, gets the same degree of consistency with the above-given rule, as the function that is almost everywhere equal to 0, and is attaining the value $2M$ only on a small interval. Intuitively, however, for the first function, for which the inequality is not true in a single point, our degree of belief that $x_1(t)$ satisfies this condition is practically 0, while for the second function, for which this inequality is almost everywhere true, our degree of belief must be close to 1.

So, using \min in our problem is inconsistent with our intuition, and therefore we must use the product for $\&$.

Comment. Other arguments for choosing different $\&$ operations are given in our previous publications [KR86] and [KQLFLKBR92].

We want to describe membership functions for the following statements. We are interested in describing statements of the type “most likely, $|x - a| \leq \delta$ ”, where x is unknown, and a, δ are known values. So we must describe, to what extent any given value x satisfies this condition.

All these membership functions can be obtained from one of them. Evidently, x satisfies the inequality $|x - a| \leq \delta$ if and only if the value $y = (x - a)/\delta$ satisfies the inequality $|y| \leq 1$. Therefore, it is natural to assume that the statement “most likely, $|x - a| \leq \delta$ ” has the same degree of belief as the statement “most likely, $|y| \leq 1$ ”, where $y = (x - a)/\delta$. So, if we will be able to describe a membership function $\mu(y)$ that corresponds to the statement “most likely, $|y| \leq 1$ ”, then we will be able to describe our degree of belief $\mu_1(x)$ that x satisfies the condition “most likely, $|x - a| \leq \delta$ ” as $\mu((x - a)/\delta)$. So the main problem is to find an appropriate function $\mu(x)$.

What if we ask several experts. A statement “most likely, $|x - a| \leq \delta$ ” means that an expert estimates x as a , and his own estimate of his precision is $\approx \delta$. Since such estimates are often very crude, it is reasonable to ask the opinion of several experts. After we have asked k experts, we get k statements of the same type: “most likely, $|x - a_i| \leq \delta_i$ ”, where $i = 1, 2, \dots, k$, and a_i and δ_i are the estimates of the i -th expert. The corresponding membership functions are $\mu((x - a_i)/\delta_i)$.

Since all of them are experts, we believe in what all of them say, and therefore our resulting knowledge is: “most likely, $|x - a_1| \leq \delta_1$, and most likely, $|x - a_2| \leq \delta_2$, and ...” Since we agreed to represent “and” as a product, the resulting membership function is equal to $\nu(x) = \mu((x - a_1)/\delta_1)\mu((x - a_2)/\delta_2)\dots\mu((x - a_k)/\delta_k)$.

In case we have a precise knowledge, and each of the experts describes an interval, in which the unknown value x must be, the resulting knowledge is that x belongs to the intersection of all

these intervals. This intersection is itself an interval, and therefore the only effect of asking several experts is that we decrease uncertainty. We do not change the form of the knowledge: it is still an interval, and in principle one smart expert could have named it from the very beginning.

In a similar way, it seems reasonable to assume that in the general fuzzy case, by combining the opinions of several experts, we do not seriously add any additional knowledge; we may diminish slightly an uncertainty domain for the unknown x , but that's all.

How to describe this argument mathematically: we must apply normalization. In mathematical terms, we would like to postulate that the resulting membership function $\nu(x)$ coincides with one of the functions $\mu((x - a)\delta)$, and in principle it could represent the opinion of just one smart expert.

We cannot, however, postulate precisely that. The reason is as follows. The bigger y , the smaller is our belief that “most likely, $|y| \leq 1$ ”. So, the function $\mu(y)$ must be monotonously decreasing for $y > 0$. Its maximum m is attained, when $y = 0$. So, when we combine the two statements “most likely, $|x| \leq 1$ ”, and “most likely, $|x - 0.3| \leq 1$ ”, the resulting membership function $\nu(x) = \mu(x)\mu(x - 0.3)$ is always smaller than m^2 , because both factors are $\leq m$, and for $x \neq 0$ the first factor is $< m$, and for $x = 0$ the second. So even if $m = 1$, the function $\nu(x)$ never attains m , and thus it cannot be equal to $\mu((x - a)/\delta)$.

The solution to this problem is well known in fuzzy logic: we can *normalize* $\nu(x)$, i.e., turn from $\nu(x)$ to $\nu'(x) = N\nu(x)$, where the normalization constant N is equal to $N = 1/(\max_y \nu(y))$.

Comment. A motivation for using namely this type of normalization is given in [KQLFLKBR92].

Now we are ready to formulate our demand.

3. MATHEMATICAL FORMULATION OF THE PROBLEM AND THE MAIN RESULT

Definition 1. By a *membership function* we will understand a continuous function $\mu(x)$ from the set R of all real numbers into the interval $[0,1]$.

Definition 2. We say that two membership functions $\mu(x)$ and $\nu(x)$ are *equivalent* if $\mu(x) = C\nu(x)$ for some constant $C > 0$.

Definition 3. We say that a membership function $\mu(x)$ is *adequate for describing uncertainty of measurements* if it satisfies the following conditions:

- it is symmetric ($\mu(-x) = \mu(x)$),
- $\mu(x)$ is strictly decreasing on $(0, \infty)$ and tends to 0 as $x \rightarrow \infty$
- for every finite sequence of pairs $(a_1, \delta_1), (a_2, \delta_2), \dots, (a_k, \delta_k)$ there exist a and δ such that the product $\mu((x - a_1)/\delta_1)\mu((x - a_2)/\delta_2)\dots\mu((x - a_k)/\delta_k)$ is equivalent to $\mu((x - a)/\delta)$.

THEOREM. Any membership function, that is adequate for describing uncertainty of measurements, is equivalent to $\exp(-\beta x)$ for some $\beta > 0$.

(The proof is given in Section 5).

Comment. So we conclude that Gaussian functions are the only adequate membership functions. These functions are really widely used [K75], [BCDMMM85], [YIS85], [KM87, Ch. 5], etc. Alternative explanation of why Gaussian functions are used is given in [KR86] and in Section 8 of [KQLFLKBR92].

4. HOW THIS RESULT CAN BE USED AND HOW IT WAS USED

How it can be used. If for some physical quantity x several experts give their estimates a_1, a_2, \dots, a_k , and they estimate the precision of their estimates as correspondingly $\delta_1, \delta_2, \dots, \delta_k$, then the resulting membership function is equal to $\mu(x) = \exp(-\beta(x - a)^2/\delta^2)$, where $\delta = (\delta_1^{-2} + \delta_2^{-2} + \dots + \delta_k^{-2})^{-1/2}$ and $a = (a_1\delta_1^{-2} + \dots + a_k\delta_k^{-2})/(\delta_1^{-2} + \dots + \delta_k^{-2})$.

Comments.

1. These formulas can be easily obtained by explicitly computing $\mu(x)$ as a result of normalization of the product $\mu_1(x)\mu_2(x)\dots\mu_k(x)$, where $\mu_i(x) = \exp(-\beta(x - a_i)^2/\delta_i^2)$.
2. These formulas are surprisingly identical with the statistical formulas that correspond to the case when we have k statistical estimates a_i with precisions δ_i and apply the least squares method $\sum_i (a - a_i)^2/\delta_i^2 \rightarrow \max_a$ to get the resulting estimate for a . This is not such a big surprise, because least squares method is based on the assumption of a Gaussian distribution. The positive side is that not only the resulting formulas are extremely simple to implement, but maybe there is no need to implement them at all, because we can copy the existing statistical software.

How this result was actually used. Expert estimates are extremely important in testing the jet engines. The reason is that an important part of this testing is trying to figure out what is going on in the high-temperature regions, and the temperatures are so high there that we cannot place any sensors. So the only available information about these regions is the experts estimates.

One of the authors (L.R.) used this fuzzy representation of uncertainty in designing software for the automatized jet engines testing system IVK-12 [KR86]. This system was actually used to test jet engine for aircraft and spaceships.

Possible other applications. One area where we believe this approach can be useful is when we determine the position of a Space Shuttle. The existing systems use several different types of sensors, with different precisions, and often with only experts estimates of that precision. In order to make appropriate control decisions we must combine these estimates into a single value. Fuzzy approach allows us to do that.

5. PROOF OF THE THEOREM

Comment. This proof contains some mathematical ideas from our previous publications [KR86] and [KQLFLKBR92].

1. Assume that $\mu(x)$ is an adequate function in the sense of the above definition. It is easy to check that if $\mu(x)$ is an adequate choice, then the result $\mu(x)/(\max \mu(y))$ of its normalization is also an adequate choice. Since $\mu(x)$ is monotone, this maximum is attained for $x = 0$, and therefore the result of this normalization satisfied the condition $\mu(0) = 1$.

So, without losing any generality, we will further assume that $\mu(0) = 1$.

2. From the definition of an adequate function it follows, in particular, that $\mu(x)\mu(x) = C\mu((x - a)/\delta)$ for some a, C and δ . The left hand side attains its maximum ($= 1$) at $x = 0$, the right-hand side attains its maximum (that is equal to C) for $x = a$. Since these two sides are one and the same function, we conclude that $a = 0$ and $C = 1$, i.e., that $\mu^2(x) = \mu(k_2x)$ for some constant $k_2 (= 1/\delta)$. For $l(x) = \log \mu(x)$ we conclude that $2l(x) = l(k_2x)$.

Likewise, if we consider 3, 4, etc terms, we conclude that $3l(x) = l(k_2x)$, $4l(x) = l(k_4x)$, etc.

3. The function $\mu(x)$ for $x > 0$ is monotonously decreasing from 1 to 0. Therefore, $l(x)$ is monotonously decreasing from 0 to $-\infty$. Since μ is continuous, the function $l(x)$ is also continuous, and, therefore, there exists an inverse function $i(x) = l^{-1}(x)$, i.e., such a function that $i(l(x)) = x$ for every x .

For this inverse function, the equality $nl(x) = l(k_n x)$ turns into $i(nl(x)) = i(l(k_n x)) = k_n x = k_n i(l(x))$. So, if we denote $l(x)$ by X , we conclude that for every n , there exists a k_n such that $i(nX) = k_n i(X)$.

If we substitute $Y = nX$, we conclude that $i(Y) = k_n i(Y/n)$, and therefore, $i(Y/n) = (1/k_n)i(Y)$.

From these two equalities, we conclude that $i((m/n)X) = (1/k_n)i(nX) = (k_m/k_n)i(X)$. So, for every rational number r , there exists a real number $k(r)$ such that $i(rX) = k(r)i(X)$.

Therefore, the ratio $i(rX)/i(X)$ is constant for all rational r .

4. Since $i(X)$ is a continuous function, and any real number can be represented as a limit of a sequence of rational numbers, we conclude that this ratio is constant for real values of r as well. Therefore, for every real number r there exists a $k(r)$ such that $i(rX) = k(r)i(X)$.

All monotone solutions of this functional equation are known: they are $i(X) = AX^p$ for some A and p [A66]. Therefore, the inverse function $l(x)$ ($x > 0$) also takes the similar form $l(x) = Bx^m$ for some k and m . Taking into consideration that $\mu(x)$ and hence $l(x)$ are even functions, we conclude that $l(x) = B|x|^m$ for all x .

5. Now, from the demand that a function $\mu(x)$ is adequate, we conclude that for every $a > 0$ we have $\mu(x-a)\mu(x+a) = C\mu((x-a_1)/\delta)$ for some a_1 and δ . The left-hand side of this equation is an even function, so the right-hand side must also be even, and therefore $a_1 = 0$. So, $\mu(x-a)\mu(x+a) = C\mu(x/\delta)$. For $x = 0$ we get $\mu(a)\mu(a) = C$. Turning to logarithms, we conclude that for every a , there exists a $k(a)$ such that $l(x-a) + l(x+a) = l(k(a)x) + 2l(a)$. If we substitute here $l(x) = B|x|^m$, and divide both sides by B , we conclude that $|x-a|^m + |x+a|^m = k(a)^m|x|^m + 2a^m$.

6. When $x > 0$, and a is sufficiently small, then $x+a$, x , and $x-a$ are all positive, and, therefore, $(x-a)^m + (x+a)^m = k(a)^m x^m + 2a^m$. If we move $2a^m$ to the left-hand side, and divide both sides by x^m , we conclude that $(1 - (a/x))^m + (1 + (a/x))^m - 2(a/x)^m = k(a)^m$. The left-hand side of the resulting equality depends only on $z = a/x$, the right-hand side only on a . Therefore, if we choose any positive real number λ , and take $a' = \lambda a$ and $x' = \lambda x$ instead of a and x , then we can conclude that the left-hand side will be still the same, and therefore, the right-hand side must be the same, i.e., $k(a)^m = k(\lambda a)^m$. Since λ was an arbitrary number, we conclude that $k(a)$ does not depend on a at all, i.e., $k(a)^m$ is a constant. Let us denote this constant by k .

So the equation takes the form $(1-z)^m + (1+z)^m = k + 2z^m$. When $z \rightarrow 0$, then the left-hand side tends to 2 and right-hand side to k , so from their equality we conclude that $k = 2$.

The left-hand side is an analytical function of z for z close to 0. Therefore the right-hand side must also be a regular analytical function in the neighborhood of 0 (i.e., it must have a Taylor expansion for $z = 0$). Hence, m must be an integer.

The values $m < 2$ are impossible, because for $m = 0$ our equality turns into a false equality $2 = 3$, and for $m = 1$ it turns into an equality $1 - z + 1 + z = 2 + z$, which is true only for $z = 0$. So $m \geq 2$.

Since both sides are analytical in z , the second derivatives of both sides at $z = 0$ must be equal to each other. The second derivative of the left-hand side at $z = 0$ is equal to $m(m-1)$. The second derivative of the right-hand side is equal to $2m(m-1)z^{m-2}$. If $m > 2$, then this derivative equals 0 at $z = 0$ and therefore cannot be equal to $m(m-1)$. So $m \geq 2$, and m cannot be greater than 2. Therefore, $m = 2$.

So, $l(x) = Bx^2$, and hence $\mu(x) = \exp(-\beta x^2)$ for some $\beta > 0$. Q.E.D.

6. CONCLUSIONS

How to represent in mathematical terms uncertain numeric statements about the value x of a physical quantity, e.g., statements of the type “most likely x is between $a - \delta$ and $a + \delta$ ”? Reasonable arguments lead us to the conclusion that the most adequate membership functions for such statements are Gaussian functions $\mu(x) = \exp(-\beta(x-a)^2/\delta^2)$.

If we use these membership functions, then we can apply simple algorithms to combine the opinions of several experts. Namely, if k experts give estimates a_1, \dots, a_k , and they estimate the precision of their estimates as correspondingly $\delta_1, \delta_2, \dots, \delta_k$, then the resulting membership function is equal to $\mu(x) = \exp(-\beta(x-a)^2/\delta^2)$, where $\delta = (\delta_1^{-2} + \delta_2^{-2} + \dots + \delta_k^{-2})^{-1/2}$ and $a = (a_1\delta_1^{-2} + \dots + a_k\delta_k^{-2})/(\delta_1^{-2} + \dots + \delta_k^{-2})$.

These formulas coincide with the ones that result from applying the statistical least squares method, so we do not even have to write a new software.

This approach was applied to testing jet engine for aircraft and spaceships, and it may be useful in many other applications, e.g., in combining the results of several coordinate and distance sensors in spaceship navigation.

ACKNOWLEDGEMENTS

This work was started with the support of the Soviet Space Shuttle Program and was continued under NSF Grant No. CDA-9015006, NASA Research Grant No. 9-482, and the Institute for Manufacturing and Materials Management grant. The authors are greatly thankful to Harold Brown (GE Aircraft Engines) and Bob Lea (NASA Johnson Space Center) for inspiring discussions.

REFERENCES

- [A66] J. Aczel. *Lectures on functional equations and their applications*. Academic Press, N.Y. and London, 1966.
- [BCDMMM85] G. Bartolini, G. Casalino, F. Davoli, M. Mastretta, R. Minciardi, and E. Morten. *Development of performance adaptive fuzzy controllers with applications to continuous casting plants*, in: M. Sugeno (editor). *Industrial applications of fuzzy control*, North Holland, Amsterdam, 1985, pp. 73–86.
- [HC76] H. M. Hersch and A. Caramazza. *A fuzzy-set approach to modifiers and vagueness in natural languages*. *J. Exp. Psychol.: General*, 1976, Vol. 105, pp. 254–276.
- [K75] A. Kauffman. *Introduction to the theory of fuzzy subsets. Vol. 1. Fundamental theoretical elements*, Academic Press, N.Y., 1975.
- [KQLFLKBR92] V. Kreinovich, C. Quintana, R. Lea, O. Fuentes, A. Lokshin, S. Kumar, I. Boricheva, and L. Reznik. *What non-linearity to choose? Mathematical foundations of fuzzy control*. Proceedings of the 1992 International Fuzzy Systems and Intelligent Control Conference, Louisville, KY, 1992, pp. 349–412.
- [KM87] R. Kruse and K. D. Meyer. *Statistics with vague data*. D. Reidel, Dordrecht, 1987.
- [O77] G. C. Oden. *Integration of fuzzy logical information*, *Journal of Experimental Psychology: Human Perception Perform.*, 1977, Vol. 3, No. 4, pp. 565–575.

[KR86] V. Kreinovich and L. K. Reznik. *Methods and models of formalizing prior information (on the example of processing measurements results)*. In: *Analysis and formalization of computer experiments*, Proceedings of Mendeleev Metrology Institute, Leningrad, 1986, pp. 37-41 (in Russian).

[YIS85] O. Yagishita, O. Itoh, and M. Sugeno. *Application of fuzzy reasoning to the water purification process*, in: M. Sugeno (editor). *Industrial applications of fuzzy control*, North Holland, Amsterdam, 1985, pp. 19–40.

[Z65] L. Zadeh. *Fuzzy sets*. *Information and control*, 1965, Vol. 8, pp. 338–353.

[Z78] H. J. Zimmermann. *Results of empirical studies in fuzzy set theory*. In: *Applied General System Research* (G. J. Klir, ed.) Plenum, New York, 1978, pp. 303–312.