FUZZY INFORMATION ON DISCRETE AND CONTINUOUS DOMAINS: APPROXIMATION RESULTS

PAULO J.S.G. FERREIRA^{a,*} and MANUEL J.C.S. REIS^{b,†}

^a*Departamento de Electrónica e Telecomunicações/IEETA, Universidade de Aveiro 3810-193, Aveiro, Portugal;* ^b*Departamento de Engenharias da Universidade de Trás-os-Montes e Alto Douro, CETAV, Portugal*

(Received 25 July 2003; In final form 21 January 2004)

Measures of information based on fuzzy sets have been defined both for finite and for continuous universes. In the continuous case, the measure of information $I(f)$ depends on the concept of non-increasing rearrangement of the function f . It has been observed that $I(f)$ can be obtained as a limit of discrete distributions $\pi^{(N)}$ approximating f . We consider the approximation problem in more detail, and study the convergence of $I(\pi^{(N)})$ to $I(f)$ in terms of the smoothness properties of f itself (modulus of continuity and Lipschitz exponent).

Keywords: Approximation, fuzzy information, measures of information, non-increasing rearrangement, sorting, uncertainty-based information

1. INTRODUCTION

Quoting from Klir and Wierman (1999: p. 138), which gives an overview of uncertainty-based information, “a host of problems regarding relative measures of non-specificity for bodies of evidence defined on infinite universal sets remains unresolved at this time”.

Consider, for example, measures of information based on fuzzy sets. These have been defined for finite and for continuous universes of discourse (Ramer, 1990). In the finite case, any information function is of the form

$$J(\pi) := \sum [\tau(p_i) - \tau(p_{i+1})] \log i,$$

where τ is a non-decreasing mapping of $[0,1]$ onto itself, and the possibility values p_i are assumed to be sorted by descending order. In other words, p_i denotes the i -th largest value of $\pi(x)$. The U -uncertainty corresponds to the case $\tau(x) = x$,

$$U(\pi) := \sum (p_i - p_{i+1}) \log i = \sum p_i [\log i - \log(i-1)].$$

In the continuous case, π can be identified with a function f , ranging over a continuous set of values, and these values must somehow be sorted. The non-increasing rearrangement

*E-mail: pjf@det.ua.pt

†Corresponding author. E-mail: mcabral@utad.pt

(Hardy *et al.*, 1952) of f (from now on, simply called “rearrangement”, and denoted by $\bar{f}$) solves the problem of “sorting”. Its definition is recalled in the next section.

Taking the matter of sorting functions with a continuous range of values as solved, one is left with the problem of finding an adequate measure of information $I(f)$. The U -uncertainty can be rewritten

$$U(\pi) = \sum p_i \nabla \log i,$$

where $\nabla \log i = \log i - \log(i-1)$. Assuming for definiteness that $f : [0, 1] \rightarrow [0, 1]$, this suggests the consideration of

$$\int_0^1 \bar{f}(x) \frac{dx}{x}.$$

But possibility distributions satisfy $\sup f = 1$. Consequently, $\bar{f}(0) = 1$, and the integral diverges. To circumvent this difficulty, Ramer (1990) considered an information distance between the given density and the uniform density, rather than an absolute measure. This leads to

$$I(f) := \int_0^1 \frac{1 - \bar{f}(x)}{x} dx. \quad (1)$$

It has been observed that $I(f)$ can be obtained as a limit of discrete distributions $\pi^{(N)}$ (with N distinct values) approximating f . However, as far as we know, this approximation problem has not been considered in detail. It is a matter of interest, since in general it is impossible to determine the rearrangement of a given function in closed form.

In this paper, we study the convergence of $I(\pi^{(N)})$ to $I(f)$ in terms of the smoothness properties of f itself (modulus of continuity and Lipschitz exponent α). We will show that

$$|\log N - U(\pi^{(N)}) - I(f)| \leq \frac{\log N}{N^\alpha}$$

under certain conditions on f . The next section introduces some necessary background, related to distribution functions, rearrangements (or sortings) and Lipschitz classes.

2. DISTRIBUTION FUNCTIONS, REARRANGEMENTS AND LIPSCHITZ CLASSES

The modulus of continuity ω_f of a continuous function f will be subsequently used. It is defined by:

$$\omega_f(\delta) := \sup_{|x-y| \leq \delta} |f(x) - f(y)|,$$

where the supremum is taken with respect to all pairs x and y in the domain of f and such that $|x - y| \leq \delta$. The modulus of continuity of a continuous function converges to zero as $\delta \rightarrow 0$. The class of functions defined by the condition $\omega_f(\delta) \leq \alpha \delta^\gamma$, where α and γ are positive constants, will also be useful. It is identical to the class of functions that satisfy a Lipschitz condition of order γ and constant α , for some positive γ and α , i.e. the set of all f such that the inequality

$$|f(x) - f(y)| \leq \alpha |x - y|^\gamma \quad (2)$$

holds for all pairs x and y in the domain of f .

The distribution function M_f associated with a function f will play an important role throughout the rest of the paper. It is defined by:

$$M_f(y) := \text{meas} \{x : f(x) > y\},$$

where “meas S ” denotes the Lebesgue measure of the set S . It is tacitly assumed that M_f is finite almost everywhere.

Any two functions with the same distribution (for example, f and a translated version of f) are called “equidistributed”. The distribution function M_f is non-increasing and right-continuous, because the underlying measure is monotone and continuous from below. Indeed, if $a > b$,

$$\{x : f(x) > a\} \subset \{x : f(x) > b\},$$

and therefore $M_f(a) \leq M_f(b)$ (because the measure is monotone). On the other hand,

$$\lim_{\substack{a \rightarrow b \\ a > b}} \{x : f(x) > a\} = \{x : f(x) > b\},$$

due to the continuity of the measure from below, showing that

$$\lim_{\substack{a \rightarrow b \\ a > b}} M_f(a) = M_f(b).$$

The rearrangement of f is (essentially) the inverse of $M_f(y)$. See Ferreira (2000; 2001) for details. The only difficulty is that M_f , which for discontinuous functions can be constant over certain intervals, does not in general have an inverse. However, it is possible to find a generalized inverse using the following procedure: Draw a horizontal line of height x across the plot of M_f and search for the leftmost point y still satisfying $M_f(y) < x$ (or the rightmost point y still satisfying $M_f(y) > x$). Selecting the first hypothesis leads to the definition

$$\tilde{f}(x) := \inf_y \{M_f(y) < x\},$$

which is a left-continuous function. It can be readily verified that this function will coincide with the inverse whenever the inverse exists.

Let f satisfy a Lipschitz condition of order γ and constant α , as in Eq. (2). Then,

$$M_f(y) - M_f(y + \epsilon) \geq \left(\frac{\epsilon}{\alpha}\right)^{1/\gamma}, \quad (3)$$

provided that y has a neighborhood contained in the range of f . This ensures the existence of a $\delta > 0$ such that for all $\epsilon < \delta$ the point $y + \epsilon$ is in the range of f . To see that Eq. (3) holds, note that under the hypothesis there will be at least one x_0 such that $y = f(x_0)$ (the horizontal line of height y will intersect the (continuous) graphic of f at least once). When y increases to $y + \epsilon$, the intersection will move to some x_1 , with $y + \epsilon = f(x_1)$. But

$$|f(x_0) - f(x_1)| = \epsilon \leq \alpha |x_0 - x_1|^\gamma,$$

and so

$$|x_0 - x_1| \geq \left(\frac{\epsilon}{\alpha}\right)^{1/\gamma}.$$

The measure of the set $S_y := \{x : f(x) > y\}$, when y increases by ϵ , decreases at least by the amount indicated on the right-hand side, and Eq. (3) follows.

A slightly more general result can be obtained, using the modulus of continuity ω_f . It asserts that

$$M_f(y) - M_f(y + \epsilon) \geq \omega_f^{-1}(\epsilon).$$

To obtain a bound for the modulus of continuity of the rearrangement, rewrite Eq. (3) as

$$|M_f(a) - M_f(b)| \geq \left(\frac{|b - a|}{\alpha} \right)^{1/\gamma},$$

and set $x_0 = M_f(a)$ and $x_1 = M_f(b)$. Then,

$$|x_0 - x_1| \geq \left(\frac{\bar{f}(x_0) - \bar{f}(x_1)}{\alpha} \right)^{1/\gamma},$$

and therefore

$$|\bar{f}(x_0) - \bar{f}(x_1)| \leq \alpha |x_0 - x_1|^\gamma. \quad (4)$$

Thus, if f satisfies a certain Lipschitz condition, so will the rearrangement $\bar{f}$. A similar result holds for the modulus of continuity: $\omega_{\bar{f}}(\delta) \leq \omega_f(\delta)$, which reduces to Eq. (4) in the Lipschitz case.

When f is discontinuous, M_f can be constant throughout intervals, and such bounds cannot be expected to exist. For an application of these concepts in the context of analog median filtering see Ferreira (2000; 2001).

Before discussing the approximation results it is convenient to discuss the rearrangement of step functions.

3. REARRANGEMENTS OF STEP FUNCTIONS

Let Δ be an arbitrary positive real. Define

$$\chi_n(t) := \begin{cases} 1, & n\Delta < t \leq (n+1)\Delta, \\ 0, & \text{otherwise.} \end{cases}$$

Consider a piecewise constant left-continuous function, with possible discontinuities at the integers, that is,

$$f(x) = c_n, \quad x \in I_n := (n\Delta, (n+1)\Delta],$$

where the c_n , $n = 0, 1, 2, \dots$, are real numbers. Using the characteristic functions χ_n of the intervals I_n , we may write

$$f(x) = \sum_n c_n \chi_n(x).$$

It is easy to see that the rearrangement of f can be obtained from the sorted sequence $\bar{c}_n$:

$$\bar{f}(x) = \sum_n \bar{c}_n \chi_n(x).$$

In other words, the rearrangement of a step function can be obtained by sorting the steps in decreasing order. This is shown in Fig. 1.

FIGURE 1 When two functions f and g have disjoint supports, $M_{f+g} = M_f + M_g$. The addition of two conveniently translated simple step functions of height A and B leads to the step function h . The distribution M_h is the sum of distribution functions, and the sorting can be obtained from the inverse (properly defined) of M_h , or equivalently, by sorting. In general, the (non-increasing) sorting of a step function with values $(A, B, C, \dots)$ can be obtained by sorting the sequence $(A, B, C, \dots)$ by decreasing order.

The proof is as follows: rearranging the steps of f in the descending order leads to a left-continuous non-increasing function.

We maintain that this function is equidistributed with f . To see that this is true, note that the distribution function of two functions with disjoint supports is the sum of the corresponding distribution functions. Hence, the distribution of a step function is the sum of the distribution functions of each step, regardless of the position of each one (Fig. 1). As a result, the distribution functions of the original step function and of its sorted version are equal.

The result now follows from the uniqueness of the rearrangement of f . More precisely, $\tilde{f}$ is the only non-increasing left-continuous function equidistributed with f (for details, see Ferreira, 2001). Therefore, the non-increasing step function that has been built is in fact the rearrangement of f .

4. THE APPROXIMATION RESULTS

In the continuous domain case, a fuzzy structure (or possibility distribution) is characterized by a function $f : D \rightarrow [0, 1]$, with $\sup f = 1$. We restrict ourselves to the case $D = [0, 1]$, to which other bounded domains D can be readily reduced.

To simplify, we begin by imposing one more condition upon f . We require that $f(x) = 1$ on an interval of Lebesgue measure δ , $\delta > 0$. Note that this implies $\tilde{f}(x) = 1$ for $0 \leq x \leq \delta$.

The condition is not too restrictive, since δ is arbitrarily small. Nevertheless, we will subsequently replace it by a weaker requirement, to allow a class of functions satisfying $f(x) = 1$ for exactly one value of x .

The first step is to build a step function f_Δ from the function f . We subdivide the unit interval into N intervals of width $\Delta = 1/N$. The step function is obtained by sampling f at the points $k\Delta$, $0 \leq k < N$. More precisely, we set $f_\Delta(x) = f(k\Delta)$, for x in $(k\Delta, (k+1)\Delta)$. It is clear that

$$\sup |f(x) - f_\Delta(x)| \leq \omega_f(\Delta),$$

where ω_f denotes the modulus of continuity of f .

Now we consider $I(f)$ and $I(f_\Delta)$, the information measures associated with f and f_Δ , defined by Eq. (1):

$$I(f) = \int_0^1 \frac{1 - \tilde{f}(x)}{x} dx,$$

$$I(f_\Delta) = \int_0^1 \frac{1 - \tilde{f}_\Delta(x)}{x} dx.$$

Note that $\tilde{f}$ is the rearrangement of f , and that $\overline{f_\Delta}$ is the rearrangement of the step function f_Δ obtained from f . The hypothesis $f(x) = 1$ over an interval of measure δ implies $\tilde{f}(x) = 1$ for $0 \leq x \leq \delta$. Moreover, since $\tilde{f}(0) = 1$ we also have $\overline{f_\Delta}(x) = 1$ for $0 \leq x \leq \delta$.

We wish to relate $I(f)$ and $I(f_\Delta)$ as $\Delta \rightarrow 0$. Assume (without loss of generality) that $\Delta < \delta$. Both $\tilde{f}$ and $\overline{f_\Delta}$ will be equal to unity on $(0, \Delta)$. Thus,

$$I(f_\Delta) = \int_{\Delta}^1 \frac{1 - \overline{f_\Delta}(x)}{x} dx = -\log \Delta - \int_{\Delta}^1 \frac{\overline{f_\Delta}(x)}{x} dx = \log N - \sum_{i=2}^N \int_{(i-1)\Delta}^{i\Delta} \frac{\overline{f_\Delta}(x)}{x} dx,$$

since $N\Delta = 1$. But $\overline{f_\Delta}$ is constant in each interval $((i-1)\Delta, i\Delta)$. Denoting that value by $\tilde{f}_i$, we have

$$I(f_\Delta) = \log N - \sum_{i=2}^N \tilde{f}_i \int_{(i-1)\Delta}^{i\Delta} \frac{dx}{x} = \log N - \sum_{i=2}^N \tilde{f}_i [\log i - \log(i-1)].$$

The sum on the right side is $U(\pi^{(N)})$, the U -uncertainty associated with the N possibility values $\tilde{f}_i$, obtained by sorting the sampled values of $\tilde{f}$. We have therefore established that

$$I(f_\Delta) = \log N - U(\pi^{(N)}). \quad (5)$$

The condition that we have imposed on f implies that $\tilde{f}(x) = 1$ for $0 \leq x \leq \Delta$, and consequently

$$|I(f_\Delta) - I(f)| \leq \int_{\Delta}^1 \frac{|\tilde{f}(x) - \overline{f_\Delta}(x)|}{x} dx \leq \int_{\Delta}^1 \frac{\omega_{\tilde{f}}(\Delta)}{x} dx = \omega_{\tilde{f}}(\Delta) \log N \leq \omega_f(\Delta) \log N. \quad (6)$$

Using Eq. (5), we conclude that

$$|\log N - U(\pi^{(N)}) - I(f)| \leq \omega_f(\Delta) \log N.$$

We thus see that $\log N - U(\pi^{(N)})$ approaches $I(f)$ as $N \rightarrow \infty$, that is, as $\Delta \rightarrow 0$, provided that

$$\omega_f(\Delta) \log N \rightarrow 0.$$

Assume, for example, that f satisfies a Lipschitz condition with exponent α . Then,

$$\omega_f(\Delta) \log N \leq \Delta^\alpha \log \frac{1}{\Delta} = \frac{\log N}{N^\alpha}$$

which does indeed converge to zero as $\Delta \rightarrow 0$ (or $N \rightarrow \infty$).

The expressions above confirm that $I(f)$ can be obtained as a limit of discrete distributions $\pi^{(N)}$ approximating f , and give approximation rates in terms of N and the smoothness of f . However, they were obtained under the hypothesis that $f(x) = 1$ on an interval of measure δ , precluding fuzzy numbers from the discussion. To circumvent this, we now turn to the general situation in which there is exactly one x_0 such that $f(x_0) = 1$.

We have, just as before,

$$I(f_\Delta) - I(f) = \int_0^1 \frac{\tilde{f}(x) - \overline{f_\Delta}(x)}{x} dx$$

and so

$$\begin{aligned} |I(f_\Delta) - I(f)| &\leq \int_0^\Delta \frac{|\tilde{f}(x) - \bar{f}_\Delta(x)|}{x} dx + \int_\Delta^1 \frac{|\tilde{f}(x) - \bar{f}_\Delta(x)|}{x} dx \\ &\leq \int_0^\Delta \frac{|\tilde{f}(x) - \bar{f}_\Delta(x)|}{x} dx + \omega_f(\Delta) \log N, \end{aligned}$$

where use has been made of Eq. (6). It remains true that $\tilde{f}(0) = 1$, and consequently $\bar{f}_\Delta(x) = 1$ for $0 \leq x \leq \Delta$. Hence,

$$|I(f_\Delta) - I(f)| \leq \int_0^\Delta \frac{|1 - \tilde{f}(x)|}{x} dx + \omega_f(\Delta) \log N. \quad (7)$$

Assume that there exists a neighborhood of x_0 and a constant $c > 0$ such that f satisfies

$$f(x) \geq 1 - c|x - x_0|.$$

This is shown in Fig. 2. The rearrangement of the function

$$g(x) := 1 - c|x - x_0|$$

can easily be computed in closed form, and is given by

$$\bar{g}(x) = 1 - \frac{c}{2}x.$$

But since $f(x) \geq g(x)$, it follows that

$$\tilde{f}(x) \geq 1 - \frac{c}{2}x$$

as $x \rightarrow 0$, and therefore

$$\int_0^\Delta \frac{|1 - \tilde{f}(x)|}{x} dx \leq \frac{c}{2}\Delta.$$

We see that the first term in Eq. (7) is $O(\Delta)$, and consequently

$$|I(f_\Delta) - I(f)| \leq O(\Delta) + \omega_f(\Delta) \log N.$$

This shows that the approximation result that we have derived above under more restrictive conditions remains valid under more general conditions. In fact, as

FIGURE 2 The condition imposed upon f : there exists a $c > 0$ such that, in a neighborhood of x_0 , f is above the function $g(x) = 1 - c|x - x_0|$. Note that f does not need to be differentiable at x_0 , nor does it need to be continuous throughout the neighbourhood.

the interested reader may confirm, the conditions considered could be weakened even further, to allow functions that are discontinuous at x_0 .

We close the section with a remark concerning possible generalizations to unbounded domains (such as the real line). The values of the rearrangement of a function $f : \mathbb{R} \rightarrow [0, 1]$ will clearly range over $[0, 1]$ as x runs through $[0, \infty)$. The natural definition for $I(f)$ would involve an integral such as

$$\int_0^\infty \frac{1 - \bar{f}(x)}{x} dx.$$

As in the bounded domain case, the hypothesis $\sup f = 1$ implies $\bar{f}(0) = 1$, and the singularity at zero will not cause any problems. However, $1 - \bar{f}(x)$ is a non-decreasing function, and therefore $(1 - \bar{f}(x))/x$ will not be integrable over $[0, \infty)$. Therefore, the generalization of the present results to unbounded domains, if at all possible, requires a different approach.

5. CONCLUSION

We have considered measures of information based on fuzzy sets, for finite and for continuous universes of discourse (Ramer, 1990). The finite case is well understood, but the definition of the measure of information $I(f)$ in the continuous case depends on the concept of rearrangement. Although it has been observed that $I(f)$ can be obtained as a limit of discrete distributions $\pi^{(N)}$ approximating f , the corresponding quantitative approximation problem was, to the best of our knowledge, open.

We studied the convergence of $I(\pi^{(N)})$ (which is related to the U -uncertainty) to $I(f)$ in terms of the smoothness properties of f itself. Assuming that $f(x) = 1$ on some set of positive (but arbitrarily small) measure, and that the Lipschitz exponent of f is α , it follows that

$$|\log N - U(\pi^{(N)}) - I(f)| \leq \frac{\log N}{N^\alpha}.$$

When $f(x) = 1$ for a unique x_0 , and f satisfies the condition summarized in Fig. 2, we have shown that

$$|\log N - U(\pi^{(N)}) - I(f)| \leq O(1/N) + \frac{\log N}{N^\alpha}.$$

These results quantitatively solve the approximation problem considered, and shed some light on the behaviour of the approximation error as $N \rightarrow \infty$.

References

- Ferreira, P.J.S.G. (2000) "Sorting continuous-time signals and the analog median filter", *IEEE Signal Process. Lett.* **7**(10), 281–283.
- Ferreira, P.J.S.G. (2001) "Sorting continuous-time signals: analog median and median-type filters", *IEEE Trans. Signal Process.* **49**(11), 2734–2744.
- Hardy, G.H., Littlewood, J.E. and Pólya, G. (1952) *Inequalities* (Cambridge University Press, Cambridge).
- Klir, G.J. and Wierman, M.J. (1999) *Uncertainty-based information: elements of generalized information theory* (Physica-Verlag, Heidelberg).
- Ramer, A. (1990) "Concepts of fuzzy information measures on continuous domains", *Int. J. Gen. Syst.* **17**, 241–248.

Paulo J.S.G. Ferreira received his Ph.D. degree in Electrical Engineering from the University of Aveiro, Portugal, where he is currently Professor Catedrático in the Department of Electronics and Telecommunications. He was an Associate Editor of the journal *IEEE Transactions on Signal Processing*, and is on the Editorial Board of *Sampling Theory in Signal and Image Processing*. He co-edited with John J. Benedetto the book *Modern Sampling Theory: Mathematics and Applications*. His research interests are in the area of linear and non-linear signal processing, and include modelling and approximation, and problems such as sampling, interpolation and signal reconstruction.

Manuel J. Cabral S. Reis received his Ph.D. degree in Electrical Engineering from the University of Aveiro, Portugal. Currently, he is Professor Auxiliar in the Department of Engineering of University of Trás-os-Montes e Alto Douro, where he is the Director. His research interests are in the area of signal processing, and include modelling and approximation, and problems such as sampling, interpolation and signal reconstruction.