Why Sugeno λ-Measures

Hung T. Nguyen[1,2], Vladik Kreinovich[3]
Joe Lorkowski[3], and Saiful Abu[3]

[1]Department of Mathematical Sciences
New Mexico State University
Las Cruces, NM 88003
[2]Faculty of Economics
Chiang Mai University, Thailand
hunguyen@nmsu.edu

[3]Department of Computer Science
University of Texas at El Paso
500 W. University
El Paso, Texas 79968
sabu@miners.utep.edu, vladik@utep.edu
lorkowski@computer.org

1. Traditional Approach: Probability Measures

- Traditionally, uncertainty has been described by probabilities.
- The probability $p(A)$ of a set A is usually interpreted as the frequency with which events from the set A occur.
- In this interpretation:
  - if we have two disjoint sets A and B with $A \cap B = \emptyset$,
  - then the frequency $p(A \cup B)$ with which the events from A or B happen
  - is equal to the sum of the frequencies $p(A)$ and $p(B)$ corresponding to each of these sets.
- This property of probabilities measures is known as *additivity*: if $A \cap B = \emptyset$, then

$$p(A \cup B) = p(A) + p(B).$$

2. Need to Go Beyond Probability Measures

- To adequately describe expert knowledge, we often need to go beyond probabilities.
- In general, instead of probabilities, we have the expert's *degree of confidence* $g(A)$ in A.
- Clearly, $g(\emptyset) = 0$ and $g(X) = 1$.
- Also, clearly, the larger the set, the more confident we are that an event from this set will occur:

$$A \subseteq B \text{ implies } g(A) \leq g(B).$$

- Functions $g(A)$ that satisfy these properties are known as *fuzzy measures.*

3. Sugeno λ-Measures

- M. Sugeno introduced a specific class of fuzzy measures which are now known as *Sugeno λ-measures*.
- If we know $g(A)$ and $g(B)$ for two disjoint sets, we can still reconstruct the degree $g(A \cup B)$.
- For Sugeno measure,

$$g(A \cup B) = g(A) + g(B) + \lambda \cdot g(A) \cdot g(B).$$

- When $\lambda = 0$, this formula transforms into additivity.
- Sugeno λ-measures are among the most widely used and fuzzy measures.

4. Problem

- The practical success of Sugeno measures is somewhat paradoxical:
  - The main point of using fuzzy measures is to go beyond probability measures.
  - On the other hand, Sugeno λ-measures are, in some reasonable sense, equivalent to probabilities.
- In this talk, we explain this seeming paradox: from the computational viewpoint,
  - processing Sugeno measure directly is much more computationally efficient
  - than using a reduction to a probability measure.
- We also analyze which other probability-equivalent fuzzy measures have this property.

5. Sugeno λ-Measure is Mathematically Equivalent to a Probability Measure

- In Sugeno measure, if we know $a = g(A)$ and $b = g(B)$ for $A \cap B = \emptyset$, then we can compute $c = g(A \cup B)$ as
$$c = a + b + \lambda \cdot a \cdot b.$$
- We would like to find a 1-1 function $f(x)$ for which $p(A) \stackrel{\text{def}}{=} f^{-1}(g(A))$ is a probability measure.
- This means that if $c = a + b + \lambda \cdot a \cdot b$, then $c' = a' + b'$,
$$\text{where } a' = f^{-1}(a), \quad b' = f^{-1}(b), \quad \text{and} \quad c' = f^{-1}(c).$$
- For $\lambda > 0$, this holds for $f(x') = \dfrac{1}{\lambda} \cdot (\exp(x') - 1)$.
- For $\lambda < 0$, this holds for $f(x') = \dfrac{1}{|\lambda|} \cdot (1 - \exp(-x'))$.
- So, a Sugeno λ-measure is indeed equivalent to a probability measure.

6. Processing Sugeno Measures Is More Computationally Efficient than Using Probabilities

- If we know $g(A)$ and $g(B)$, then we can compute

$$g(A \cup B) = g(A) + g(B) + \lambda \cdot g(A) \cdot g(B).$$

- This computation uses only hardware supported (thus, fast) $+$ and $\cdot$. Alternative is to:
  - compute $p(A) = f^{-1}(g(A))$ and $p(B) = f^{-1}(g(B))$;
  - add these probabilities $p(A \cup B) = p(A) + p(B)$;
  - finally, re-scale this resulting probability back into degree-of-confidence: $g(A \cup B) = f(p(A \cup B))$.
- In this approach, we compute logarithm (to compute $f^{-1}(x)$) and exponential function (to compute $f(x)$).
- These computations are much slower than $+$ and $\cdot$.
- Thus, the direct use of Sugeno measure is definitely much more computationally efficient.

7. How to Explain the Use of Sugeno Measure in Probabilistic Terms

- We are interested in expert estimates of probabilities of different sets of events.
- It is known that expert estimates of the probabilities are biased:
  - the expert's subjective estimates $g(A)$ of the corresponding probabilities $p(A)$
  - are equal to $g(A) = f(p(A))$ for an appropriate re-scaling function $f(A)$.
- In this case, a natural ideas seems to be:
  - to re-scale all the estimates back into the probabilities: $p(A) = f^{-1}(g(A))$, and
  - to use the usual algorithms to process these probabilities.

8. Sugeno Measure in Prob. Terms (cont-d)

- If we know the expert's estimates $g(A)$ and $g(B)$ for $A \cap B = \emptyset$, to predict the $g(A \cup B)$, we:
  - re-scale $g(A)$ and $g(B)$ into probabilities:
  $$p(A) = f^{-1}(g(A)) \text{ and } p(B) = f^{-1}(g(B));$$
  - compute $p(A \cup B) = p(A) + p(B)$; and
  - estimate $g(A \cup B)$ as $g(A \cup B) = f(p(A \cup B))$.
- For some biasing functions $f(x)$, it is computationally more efficient
  - *not* to re-scale into probabilities,
  - but to store and process the original values $g(A)$.
- This is, in effect, the essence of applications of a Sugeno λ-measure are about.

9. Which Fuzzy Measures Have This Property

- If we know the expert's estimates $a = g(A)$ and $b = g(B)$ for $A \cap B = \emptyset$, to predict the $g(A \cup B)$, we:
  - re-scale a and b into probabilities:
    $$p(A) = f^{-1}(a) \text{ and } p(B) = f^{-1}(b);$$
  - compute $p(A \cup B) = f^{-1}(a) + f^{-1}(b)$; and
  - estimate $g(A \cup B)$ as $F(a, b) = f(f^{-1}(a) + f^{-1}(b))$.
- One can check that $F(a, b)$ is commutative, associative, and $F(0, a) = a$.
- We want to find all such $F(a, b)$ for which direct computation is faster than this 3-stage approach.
- Computation is fast it consists of a sequence of hardware supported elementary operations: $+$, $-$, $\cdot$, $/$.

10. Analysis of Fuzzy Measures (cont-d)

- We are interested in functions

$$F(a,b) = f(f^{-1}(a) + f^{-1}(b)).$$

- These functions are commutative, associative, and $F(0,a) = a$.
- We want to find all such $F(a,b)$ for which direct computation is faster than this 3-stage approach.
- Computation is fast it consists of a sequence of hardware supported elementary operations: $+$, $-$, $\cdot$, $/$.
- Functions computed by a sequence of such operations are *rational* – fractions of polynomials.
- Thus, we look for rational commutative associative functions $F(a,b)$ for which $F(0,a) = a$.

11. Main Result

- We are looking for fuzzy measures:
  - which are equivalent to probability measures, but
  - for which direct computations are faster than reductions to probabilities.
- This leads to a search for rational commutative associative functions $F(a,b)$ for which $F(0,a)=a$.
- We prove that each such operation has one of the two forms:

$$F(a,b)=\frac{a+b+2B\cdot a\cdot b}{1+B^2\cdot a\cdot b};$$

$$F(a,b)=\frac{a+b+(2B+A)\cdot a\cdot b}{1-B\cdot(B+A)\cdot a\cdot b}.$$

- For $B=0$, the second formula leads to Sugeno measure.

12. Auxiliary Result

- We look for operations for which computing $F(a, b)$ directly is faster.
- The requirement that $F(a, b)$ is computable by elementary arithmetic operations leads to

$$F(a, b) = \frac{a + b + 2B \cdot a \cdot b}{1 + B^2 \cdot a \cdot b};$$

$$F(a, b) = \frac{a + b + (2B + A) \cdot a \cdot b}{1 - B \cdot (B + A) \cdot a \cdot b}.$$

- Out of elementary arithmetic operations, division is the slowest.
- Sugeno measure is the only one that does not use division and is, thus, the fastest.
- This explains why Sugeno measure is widely used.

13. Proof

- A classification of all possible rational commutative associative $F(a,b)$ is known (Brawley et al. 2001).
- For each such $F(a,b)$, there exists a fractional-linear $t(a)$ for which $F(a,b)=t^{-1}(t(a)+t(b))$ or

$$F(a,b)=t^{-1}(t(a)+t(b)+t(a)\cdot t(b)).$$

- The requirement $F(0,a)=a$ implies $t(0)=0$.
- A general fractional-linear function has the form

$$t(a)=\frac{p+q\cdot a}{r+s\cdot a}.$$

- The fact that $t(0)=0$ implies that $p=0$, so we get

$$t(a)=\frac{q\cdot a}{r+s\cdot a}.$$

14. Proof (cont-d)

- We have shown that $t(a) = \dfrac{q \cdot a}{r + s \cdot a}$.
- Here, we must have $r \neq 0$, because otherwise, $t(a)$ is a constant.
- Dividing the numerator and the denominator of $t(a)$ by r, we get:

$$t(a) = \frac{A \cdot a}{1 + B \cdot a}, \text{ where } A \stackrel{\text{def}}{=} \frac{q}{r}, B \stackrel{\text{def}}{=} \frac{s}{r}.$$

- We know that $F(a, b) = t^{-1}(t(a) + t(b))$ or

$$F(a, b) = t^{-1}(t(a) + t(b) + t(a) \cdot t(b)).$$

- Substituting this expression for $t(a)$ into the above formulas for $F(a, b)$, we get the desired expressions.

15. Acknowledgments

- This work was supported in part:
  - by the Faculty of Economics of Chiang Mai University and
  - by the National Science Foundation grants:
    - HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence), and
    - DUE-0926721.
- The authors are thankful to George J. Klir and Michio Sugeno for valuable discussions.

16. References

- J. V. Brawley, S. Gao, and D. Mills, "Associative rational functions in two variables", In: D. Jungnickel and H. Niederreiter, *Finite Fields and Applications 2001*, Proceedings of the Fifth International Conference on Finite Fields and Applications F_q5, Augsburg, Germany, August 2–6, 1999, Springer, Berlin, Heidelberg, 2001, pp. 43–56.
- D. Kahneman, *Thinking, Fast and Slow*, Farrar, Straus, and Giroux, New York, 2011.
- M. Sugeno, *Theory of Fuzzy Integrals and Its Applications*, Ph.D. Dissertation, Tokyo Institute of Technology, 1974.