

基于 Schweizer 算子簇的柔性概率逻辑算子的研究^{*}

王万森^{1,2} 何华灿²

(首都师范大学信息工程学院 北京 100037)¹ (西北工业大学计算机学院 西安 710072)²

摘要 Schweizer 算子簇是泛逻辑学研究零级非相容 T/S 范数完整簇的数学基础,由它构造的与/或运算具有连续单调可变性。基于 Schweizer 算子簇构造的概率逻辑算子,既可满足概率测度的基本公理,又可实现概率逻辑运算的连续单调可变。

关键词 T/S 范数, Schweizer 算子, 概率逻辑, 泛逻辑学

Research on Flexible Probability Logic Operator Based on Schweizer Operator

WANG Wan-Sen^{1,2} HE Hua-Can²

(College of Information Engineering, Capital Normal University, Beijing 100037)¹

(College of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072)²

Abstract Schweizer operator cluster is mathematics foundation of the research on zero pole not-compatible T/S norms complete cluster in Universal logics. The AND/OR operation constructed by this cluster is series monotony and variable. The Probability logic operator constructed by Schweizer operator cluster is not only can satisfy the basic axiom of probability, but also realize the seriate monotone and variable operation of Probability logic.

Keywords T/S norms, Schweizer operator, Probability logic, Universal logics

1 引言

概率逻辑在不确定性推理方面具有重大的理论价值和应前景,但其目前还存在一定缺陷。如何解决这些缺陷,泛逻辑学提供了一条有效途径^[1]。泛逻辑学是一个可包容多种具体逻辑的柔性逻辑体系,按照其理论,概率逻辑仅是它的一个特例^[2,3]。

T/S 范数是泛逻辑学的主要数学工具。在众多的 T/S 范数簇中, Schweizer 算子簇是目前发现的惟一的一个 T/S 范数完整簇,由它所构造的概率逻辑算子,既可满足概率测度的基本公理,又可保持 Schweizer 算子簇的连续单调可变的特性,即实现了概率算子的柔性化^[3,4]。

2 非相容柔性概率逻辑算子的运算模型

利用 Schweizer T 范数和 Schweizer S 范数簇,对基于命题语言 L_p 的概率逻辑算子进行重新定义,可得到随 h 连续单调变化的新的概率逻辑算子。

2.1 基于 L_p 的概率逻辑基础

L_p 是经典命题逻辑所使用的形式语言,卡尔纳普(Carnap)概率逻辑和波普尔(Popper)概率逻辑是两个典型的基于 L_p 的概率逻辑模型^[5]。 L_p 上的概率测度可定义如下^[6]。

定义 1 设 F 为样本空间 Ω 上的 σ 代数, F 上的实值函数,如果具有

- (1) 非负性:任意 $A \in F$, 有 $0 \leq P(A)$;
- (2) 规范性: $P(\Omega) = 1$;
- (3) 可加性:任意 $A, B \in F, A \cap B = \emptyset$, 有 $P(A \cup B) = P(A) + P(B)$;

则称 P 为 (Ω, F) 上的概率测度,而 $P(A)$ 为事件 A 的概率。

2.2 Schweizer T 范数

三角范数(triangular norm, 又称 T-模或 T-范数)是一种能生成经典三角不等式的函数^[7],泛逻辑学中的许多运算模型都是建立在其基础之上的^[2,3]。

定义 2 设 T 是 $[0, 1]$ 上的二元函数,即 $T: [0, 1]^2 \rightarrow [0, 1]$, 对任意 $x, y, z \in [0, 1]$, 若 T 满足交换律、结合律、单调性和有单位元 1, 即满足以下 4 条公理:

- (T₁) $T(x, y) = T(y, x)$;
- (T₂) $T(x, T(y, z)) = T(T(x, y), z)$;
- (T₃) 当 $y \leq z$ 时, 有 $T(x, y) \leq T(x, z)$;
- (T₄) $T(x, 1) = x$;

则称 T 是一个三角范数(或称 T 模或 T 范数)。

用同样的方式,可定义三角余范(或称 T 余范数或 S 范数)。

2.3 非相容零级 T 范数簇和 S 范数簇

基于 Schweizer 算子可生成非相容的零级泛与运算模型 $T(x, y, h)$ 和零级泛或运算模型 $S(x, y, h)$ 。它们的定义分别如下:

$$T(x, y, h) = (\max(0^m, x^m + y^m - 1))^{1/m}$$

$$S(x, y, h) = 1 - (\max(0^m, (1-x)^m + (1-y)^m - 1))^{1/m}$$

其中, $m = (3-4h)/(4h(1-h))$, $h = ((1+m) - ((1+m)^2 - 3m)^{1/2})/(2m)$, $h \in [0, 1]$, $m \in R$ 。

基于 Schweizer 算子的零级运算模型是一种非相容运算模型,其相容差为

$$\Delta = x + y - S(x, y, h) - T(x, y, h)$$

2.4 独立性与广义相关性

独立性是概率中的一个概念,用来描述一个事件对另一个事件的概率的影响,即事件之间的相关性。概率虽然已经

^{*} 国家自然科学基金(60273087)资助;北京市自然科学基金(4032009)资助。王万森 教授,博士生,主要研究方向:人工智能、人工情感、不确定推理、概率逻辑等;何华灿 教授,博士生导师,主要研究方向:人工智能、泛逻辑、概率逻辑、不确定推理等。

发现了事件之间的独立性,但却没有对独立性进行深入研究,更没能建立起独立性与概率之间的有机联系,而只是简单地将事件之间的独立性分为相互独立和互不独立两种类型。

广义相关性是泛逻辑学中的概念,用来描述不同命题之间的关联性,由广义相关系数 $h \in [0.5, 1]$ 来刻画。独立性和广义相关性虽然是两个不同概念,但它们之间却存在一定的联系。对 h 的一些特殊值,其对应的概率描述为:

$h=1$,表示最大相吸(最小相斥),对应的概率表示为: $P(A \wedge B) = \min(P(A), P(B))$, $P(A \vee B) = \max(P(A), P(B))$, $P(A|B) = \min(P(A), P(B))/P(B)$ 。

$h=0.75$,表示独立相关(中度相容),对应的概率表示为: $P(A \wedge B) = P(A)P(B)$, $P(A \vee B) = P(A) + P(B) - P(A)P(B)$, $P(A|B) = P(A)$ 。

$h=0.5$,表示最大相斥(最小相吸),对应的概率表示为: $P(A \wedge B) = \max(0, P(A) + P(B) - 1)$, $P(A \vee B) = \min(1, P(A) + P(B))$, $P(A|B) = \max(0, P(A) + P(B) - 1)/P(B)$ 。

2.5 非相容柔性概率逻辑算子的定义

基于 Schwizer 算子簇的非相容柔性概率逻辑算子的定义方法是,利用泛逻辑学的零级泛与 $T(x, y, h)$ 、泛或 $S(x, y, h)$ 和泛非 $N(x)$ 模型来建立相应的概率逻辑运算模型^[8]。

定义3 设 P 是一个 L_p 语言的概率函数, $A, B \in L_p$, 且 $P(A) = x, P(B) = y$, 则基于 Schwizer 算子的柔性概率逻辑算子的运算模型分别为:

(1) 概率非命题连接词完整簇的运算模型,即柔性概率逻辑的非运算模型:

$$P(\neg A) = N(P(A)) = N(x, 0.5) = 1 - x$$

(2) 概率与命题连接词完整簇的运算模型,即柔性概率逻辑的与运算模型

$$P(A \wedge B, h) = T(P(A), P(B), h) = (\max(0^m, x^m + y^m - 1))^{1/m}$$

(3) 概率或命题连接词完整簇的运算模型,即柔性概率逻辑的或运算模型可以有两种定义方法:一种是按 $T(x, y, h)$ 定义,另一种是按 $S(x, y, h)$ 定义。如果按 $T(x, y, h)$ 定义,则有

$$P(A \vee B, h) = P(A) + P(B) - T(P(A), P(B), h) = x + y - (\max(0^m, x^m + y^m - 1))^{1/m}$$

(4) 概率条件命题连接词完整簇的运算模型,即柔性概率逻辑的条件运算模型:

$$P(A|B, h) = T(P(A), P(B), h) / P(B) = (\max(0^m, x^m + y^m - 1))^{1/m/y}$$

在以上定义中, $m = (3 - 4h) / (4h(1 - h))$, $h = ((1 + m) - ((1 + m)^2 - 3m)^{1/2}) / (2m)$, $h \in [0.5, 1]$, $m \in \mathbb{R}$ 。

3 非形容柔性概率逻辑算子满足基本概率测度

非相容柔性概率逻辑算子满足如下概率测度的基本公理。

公理1 设 P 是一个 L_p 语言的概率函数, $A, B \in L_p$, Ω 为整个概率空间, A, B 为任何两两互不相容的事件, 且 $P(A) = x, P(B) = y$, 基于 Schwizer 算子的非相容柔性概率逻辑满足定义1所给出的概率测度的基本公理:

(1) 非负性: $P(A) \geq 0$;

(2) 规范性: $P(\Omega) = 1$;

(3) 可加性: 任意 $A, B \in F, A \cap B = \emptyset$, 有 $P(A \cup B) = P(A) + P(B)$ 。

证明:

(1) 由概率的定义 $P(A) = x$, 由于 $P(A) \geq 0$, 因此 $P(A)$

$= x \geq 0$ 。

(2) 由柔性概率非运算模型 $P(\Omega) = P(\Phi) = N(P(\Phi)) = N(\Phi, 0.5) = 1 - 0 = 1$

(3) 由概率的相容性和独立性可知, 不相容必定是不独立的。再由广义相关性的定义, 不相容是一种最大相斥的情况, 即 $h = 0.5$ 。

由前面分析, 当 $h = 0.5$ 时, 有 $P(A \vee B, 0.5) = \min(1, P(A) + P(B)) = P(A) + P(B)$

4 非形容柔性概率逻辑算子的单调性

在零级柔性概率逻辑算子中, $P(A \wedge B)$ 和 $P(A \vee B)$ 主要是通过 $T(x, y, h)$ 定义的。因此, 下面主要利用 $T(x, y, h)$ 来讨论零级柔性概率逻辑算子的性质。

定理1 对 $h \in (0.75, 1)$, $T(x, y, h)$ 是随 h 单调递增的。

证明: 根据对 m 和 h 的定义, 要证明 $T(x, y, h)$ 关于 h 是递增的, 需要分别证明 $T(x, y, h)$ 关于 m 总是递减的, 并且 m 关于 h 也总是递减的。

(1) 先证明 $T(x, y, h)$ 关于 m 总是递减的。由于

$$T(x, y, h) = (\max(0, x^m + y^m - 1))^{1/m}, \\ m = \frac{3 - 4h}{4h(1 - h)}$$

当 $0.75 < h < 1$ 时, 有 $m < 0$,

即 $x^m + y^m > 1, x^m + y^m - 1 > 0$

因此有

$$T(x, y, h) = (x^m + y^m - 1)^{1/m}$$

$$\text{即: } \ln T(x, y, h) = \frac{1}{m} \ln(x^m + y^m - 1)$$

两边关于 m 求导, 有

$$\frac{\partial T(x, y, h)}{\partial m} = \left(-\frac{1}{m^2} \right) \ln(x^m + y^m - 1) + \frac{1}{m} \times \frac{x^m \ln x + y^m \ln y}{x^m + y^m - 1}$$

由于 $T(x, y, h) > 0$, 即要证明 $\frac{\partial T(x, y, h)}{\partial m} < 0$

它等价于证明 $\frac{1}{m^2} \ln(x^m + y^m - 1) > \frac{1}{m} \frac{x^m \ln x + y^m \ln y}{x^m + y^m - 1}$

两边同乘以 $(x^m + y^m - 1)m^2$, 得

$$(x^m + y^m - 1) \ln(x^m + y^m - 1) > m(x^m \ln x + y^m \ln y)$$

即证 $(x^m + y^m - 1) \ln(x^m + y^m - 1) > x^m \ln x + y^m \ln y$ 。

只要能够证明上式对所有 $m < 0$ 都成立即可。

令 $x^m = a > 1, y^m = b > 1$, 上式等价于 $(a + b - 1) \ln(a + b - 1) > a \ln a + b \ln b$, 即 $(a + b - 1) \ln(a + b - 1) - a \ln a - b \ln b > 0$ 。

记不等式左边为函数为 $F(a, b) = (a + b - 1) \ln(a + b - 1) - a \ln a - b \ln b$, 只要能证明对任意 $a, b > 1$, 都有 $F(a, b) > 0$ 即可。

当 $a = 1$ 时, $F(1, b) = (1 + b - 1) \ln(1 + b - 1) - 1 \times \ln 1 - b \ln b = b \times \ln b - b \times \ln b = 0$

当 $b = 1$ 时, $F(a, 1) = (a + 1 - 1) \ln(a + 1 - 1) - a \ln a - 1 \times \ln 1 = a \times \ln a - a \times \ln a = 0$

若能证明关于 a, b 严格递增, 即可说明对任意 $a, b > 1$ 都有 $F(a, b) > 0$ 。

对 $F(a, b)$ 关于 a 求偏导数, 有

$$\frac{\partial F(a, b)}{\partial a} = \ln(a + b - 1) + \frac{a + b - 1}{a + b - 1} - \ln a - \frac{a}{a} \\ = \ln(a + b - 1) - \ln a$$

由于 $b > 1$, 因此有 $\ln(a + b - 1) > \ln a$, 即 $\frac{\partial F(a, b)}{\partial a} > 0$ 。

同理可证 $\frac{\partial F(a,b)}{\partial b} = \ln(a+b-1) - \ln a > 0$ 。

所以, $F(a,b)$ 关于 a, b 严格递增, 即 $T(x, y, h)$ 是随 m 单调递减的。

由于 $T(x, y, h)$ 关于 m 是递减的, 如果能证明 m 关于 h 总是递减的, 则说明 $T(x, y, h)$ 关于 h 是递增的。

(2) 再证明 m 关于 h 总是递减的

$$\begin{aligned}\frac{\partial m}{\partial h} &= \frac{-4(4h(1-h)) - (3-4h)(4-8h)}{(4h(1-h))^2} \\ &= \frac{-16h + 16h^2 - (12 - 16h - 24h + 32h^2)}{(4h(1-h))^2} \\ &= \frac{-(16h^2 - 24h + 12)}{(4h(1-h))^2}\end{aligned}$$

要证明 $\frac{\partial m}{\partial h} < 0$, 就是要证明 $16h^2 - 24h + 12 > 0$, 即要证明 $4h^2 - 6h + 3 > 0$

由于判别式 $\Delta = (-6)^2 - 4 \times 4 \times 3 = -12 < 0$, 因此总有 $4h^2 - 6h + 3 > 0$, 即 m 关于 h 是递减的。

综上所述, T 关于 h 是递增的。(证毕)

定理 2 对 $h \in (0.5, 0.75)$, $T(x, y, h)$ 是随 h 单调递增的。

证明: 根据对 m 和 h 的定义, 要证明 $T(x, y, h)$ 关于 h 是递增的, 需要分别证明 $T(x, y, h)$ 关于 m 总是递减的, 并且 m 关于 h 也总是递减的。

(1) 先证明 $T(x, y, h)$ 关于 m 总是递减的。

由于 $T(x, y, h) = (\max(0, x^m + y^m - 1))^{\frac{1}{m}}$,

$$m = \frac{3-4h}{4h(1-h)}$$

考虑 $\frac{\partial T(x, y, h)}{\partial m}$, 若 $\frac{\partial T(x, y, h)}{\partial m} < 0$, 则 $T(x, y, h)$ 是随 m 单调递减的。

∵ 当 $0.5 < h < 0.75$ 时, 有 $m > 0$, $0 < x^m < 1$, $0 < y^m < 1$ 。

当 $x^m + y^m - 1 \leq 0$ 时, 有 $T(x, y, h) = 0$ 。

当 $x^m + y^m - 1 > 0$ 时, 有 $T(x, y, h) = (x^m + y^m - 1)^{\frac{1}{m}}$ 。

下面考虑 $x^m + y^m - 1 > 0$ 的情况。

对 $T(x, y, h) = (x^m + y^m - 1)^{\frac{1}{m}}$ 两边取对数, 有 $\ln T(x, y, h) = \frac{1}{m} \ln(x^m + y^m - 1)$ 。

两边关于 m 求导, 有

$$\begin{aligned}\frac{\partial T(x, y, h)}{\partial m} &= \left(-\frac{1}{m^2}\right) \ln(x^m + y^m - 1) + \\ &\quad \frac{1}{m} \times \frac{x^m \ln x + y^m \ln y}{x^m + y^m - 1}\end{aligned}$$

要证明 $\frac{\partial T(x, y, h)}{\partial m} < 0$, 即等价于证明

$$\frac{1}{m^2} \ln(x^m + y^m - 1) > \frac{1}{m} \times \frac{x^m \ln x + y^m \ln y}{x^m + y^m - 1}$$

其证明过程与定理 1 相似, 故省略。因此 $T(x, y, h)$ 是随 m 单调递减的。

由于 $T(x, y, h)$ 关于 m 是递减的, 如果能证明 m 关于 h 总是递减的, 则说明 $T(x, y, h)$ 关于 h 是递增的。

(2) 再证明 m 关于 h 总是递减的

其证明过程与定理 1 类似, 故略。

综上所述, T 关于 h 是递增的。(证毕)

结论 基于 Schweizer 范数的柔性概率逻辑算子, 可以通过 h 的变化, 使其值在 $[0, 1]$ 区间连续单调变化。实际上, 由定理 1 和定理 2 还可以推出, 当 $h \in (0.75, 1)$ 时有 $P(A|B) > P(A)$, 当 $h \in (0.5, 0.75)$ 时有 $P(A|B) < P(A)$ 。这一研究, 为实现逻辑框架内的概率逻辑不确定推理, 提供了一条有效途径。笔者下一步的研究重点是要建立基于 Frank 算子的柔性概率逻辑算子及其运算模型。

参考文献

(上接第 163 页)

另外, 本文实验过程中手工确定最佳学习率、聚类数、隐层神经元数和最大训练次数的工作非常烦琐, 且增加不确定性。因此, 确定一种自动确定上述参数的机制显得非常迫切。也有一些参数的确定已有现成方法, 在今后的工作中, 我们将会在模型中集成这些算法, 并作进一步探索。

参考文献

- Sharkey A J C. Combining Artificial Neural Nets: Ensemble and Modular Multi-Net Systems. Springer-Verlag New York, Inc. Secaucus, NJ, USA, 1999
- Haykin S. Neural Networks: A Comprehensive Foundation. Prentice Hall PTR Upper Saddle River, NJ, USA, 1998
- Chen C H, Lin Z S. A committee machine with empirical formulas for permeability prediction. Computers & Geosciences, 2006, 32: 485~496
- Bullinaria J A. Lecture 15: Committee Machines. 2004
- Long J, Jian J, Cai Y. A Short-Term Climate Prediction Model Based on a Modular Fuzzy Neural Network. Advances in Atmospheric Sciences, 2005, 22: 428~435

- 王万森, 何华灿. 基于泛逻辑学的逻辑关系柔性化研究[J]. 软件学报, 2005, 16(5): 754~760
- 何华灿, 等. 泛逻辑学原理[M]. 北京: 科学出版社, 2001
- He Huacan, Wang Hua, Liu Yonghuai, et al. Princip of Universal Logics [M]. China: Science Press, 2006
- 王万森, 何华灿. 基于泛逻辑学的柔性命题逻辑研究[J]. 小型微型计算机系统, 2004, 25(12): 2116~2119
- Roeper P, Leblanc H. Probability theory and probability logic [M]. university of toronto press incorporated, 1999
- 张文修, 梁怡. 不确定性推理原理[M]. 西安: 西安交通大学出版社, 1996
- Klement E P, Mesiar R, Pap E. Triangular Norms. Dordrecht: Kluwer Academic Publishers, 2000
- 王万森, 何华灿. 基于泛逻辑学的概率命题逻辑研究与分析[J]. 计算机研究与发展, 2005, 42(7): 1204~1209
- McCulloch W S, Pitts W. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biology, 1943, 5: 115~133
- Morshed J, Kaluarachchi J J. Application of artificial neural network and genetic algorithm in flow and transport simulations. Advances in Water Resources, 1998, 22: 145~158
- Tresp V. Committee Machines. In: Handbook for Neural Network Signal Processing, Yu Hen Hu and J.-N. Hwang, eds. CRC Press, 2001
- Bezdek J C. Pattern Recognition with Fuzzy Objective Function Algorithms. Kluwer Academic Publishers Norwell, MA, USA, 1981
- Shitong W. Fuzzy System, Fuzzy Neural Networks and the Design of Application Program Shanghai. Shanghai Scientific & Technical Press, 1998
- Jin J, Wang J, Huang G, et al. Study on Influence of Effectiveness Factors on Modular Fuzzy Neural Network Based Prediction Model. In: Sixth World Congress on Intelligent Control and Automation, Conference Proceedings (WCICA 2006). Dalian, China, 2006
- Dongjun Y, Shitong W. A New Fuzzy Neural Network Model and Its Modelling Application Research. Journal of East China Shipbuilding Institute, 1998, 12