Choquet 积分的模糊化扩展 II 型

杨 蓉 郑三元

(深圳大学机电与控制工程学院 深圳 518000)

摘 要 重点讨论 Choquet 积分的 II 型模糊化扩展形式。相对于 Choquet 积分的模糊化扩展 I 型，II 型扩展支持模糊数的被积函数，并给出精确数的积分结果。基于带符号的模糊测度，分别讨论了 Choquet 积分的模糊化扩展 II 型的计算方法和相关算法。并用实例说明：Choquet 积分的模糊化扩展 II 型作为一种聚合工具在处理涉及非精确数的数据挖掘问题（例如：推理和回归问题）上的实用价值。

关键词 Choquet 积分，模糊数，模糊测度，聚合

中图法分类号 TP14 **文献标识码** A

Type II Fuzzification on Choquet Integral

YANG Rong ZHENG San-yuan

(College of Mechatronics and Control Engineering, Shenzhen University, Shenzhen 518000, China)

Abstract This paper provided a detailed discussion on one fuzzification of Choquet integral which supports fuzzy-valued integrand and gave crisp-valued integration result. It is a generalized Choquet integral for fuzzy-valued integrand, interval-valued integrand, as well as the crisp-valued integrand. The presented generalized Choquet integral with respect to signed fuzzy measure can act as an aggregation tool which is especially useful in many information fusing and data mining problems (such as regression and decision making) where not only crisp data but also heterogeneous fuzzy data are involved.

Keywords Choquet integral, Fuzzy data, Fuzzy measure, Aggregation

1 引言

近年来，作为一种有效的非线性聚合工具，基于模糊测度(Fuzzy measure)[1,2]上的 Choquet 积分[3,4]常被应用在信息融合和数据挖掘等问题上。建立在各特征量上的模糊测度因具有非可加性及非单调性，故能够有效描述各特征量对目标量贡献之间的交互关系。因此，Choquet 积分被成功地应用在非线性多重回归问题[5]、分类问题以及决策管理等问题上[6]。在这些应用中，模糊测度的具体值常作为未知参数需要通过训练数据集及一定算法来得出。

Choquet 积分只支持精确数的被积函数，故只能处理精确数据的信息挖掘，对于在各种实际应用中广泛出现的非精确数据（模糊数据），则无能为力。基于以上原因，需要对 Choquet 积分进行模糊化建模，使其不但能处理精确数据，亦能处理模糊数据。该模糊化建模要求 Choquet 积分能够支持模糊数的被积函数，被称为 Choquet 积分的模糊化扩展。此扩展可作为 Choquet 积分的一种普遍性形式，因其不仅可处理传统 Choquet 积分能够处理的精确数据，亦可处理其它形式的信息，如区间数、模糊数和语义变量等。

Choquet 积分的模糊化扩展有多种扩展范畴。对于一个给定测度值为实数的模糊测度，当被积函数允许支持模糊数时，可以分别定义 Choquet 积分的积分结果为模糊数或精确实数的形态，即可构成两种不同的 Choquet 积分的模糊化扩展，分别称其为 I 型和 II 型。I 型的相关理论研究和扩展应用已在相关文献中介绍[14]。本文针对 II 型模糊化扩展进行论述，并在其基础上重点提出一个针对模糊数据的非模糊化的聚合模型。

2 模糊数

令 R 为实数集$(-\infty,\infty)$，通常利用模糊数来量化模糊的概念。

定义 1 一个规则的模糊数是一个定义在实数集 R 上的模糊子集，其隶属度函数 $m:R\to[0,1]$满足以下条件：

(1)至少存在一个 $a_0\in R$，使得 $m(a_0)=1$；

(2)$m(t)$在$(-\infty,a_0]$上单调上升，在$[a_0,\infty)$上单调下降；

(3)$m(t)$上半连续，即当 $t_0<a_0$ 时 $\lim\limits_{t\to t_0+}m(t)=m(t_0)$，当 $t_0>a_0$ 时 $\lim\limits_{t\to t_0-}m(t)=m(t_0)$；

(4) $\int_{-\infty}^{\infty}m(t)\mathrm{d}t<\infty$。

常规模糊数中最常用的是梯形模糊数[10]，其隶属度函数具有以下形式：

本文受国家自然科学基金项目(61105044)资助。

杨 蓉(1976－)，女，博士，副教授，主要研究方向为人工智能、模糊识别，E-mail：ryang@szu.edu.cn；**郑三元**(1963－)，女，硕士，副教授，主要研究方向为智能控制。

$$m(t)=\begin{cases}1, & \text{if } t=[a_b, a_c] \\ \dfrac{t-a_l}{a_b-a_l}, & \text{if } t\in[a_l, a_b) \\ \dfrac{t-a_r}{a_c-a_r}, & \text{if } t\in(a_c, a_r] \\ 0, & \text{otherwise}\end{cases}$$

式中，$a_l, a_b, a_c, a_r \in R$ 且 $a_l \leqslant a_b \leqslant a_c \leqslant a_r$。在本论文中，梯形模糊数简示为$[a_l \quad a_b \quad a_c \quad a_r]$。当 $a_l \geqslant 0$ 时，称该梯形模糊数为非负的。一个三角模糊数可被看作是梯形模糊数的特殊形式（当 $a_b=a_c$）；一个区间数也可被看作是梯形模糊数的特殊形式（当 $a_l=a_b$ 且 $a_c=a_r$）；同样，一个精确实数可看作是当 $a_l=a_b=a_c=a_r=a$ 时的梯形模糊数。因此，以下的讨论及模型既适用于模糊数据，也适用于精确实数数据。

3 模糊集合的带符号模糊测度

令 $X=\{x_1, x_2, \cdots, x_n\}$，则 X 的幂集(Power set)用符号 $P(X)$ 来表示，而关于 X 上的所有模糊子集的集合用符号 $F(X)$ 来表示。任何一个 X 的模糊子集 $\widetilde{A}$，可被描述为

$$\widetilde{A}=\{d_1/x_1, d_2/x_2, \cdots, d_n/x_n\}$$

式中，d_i 是 $\widetilde{A}$ 在 x_i 处的隶属度，$i=1,2,\cdots,n$；亦可简洁描述为

$$\widetilde{A}=(d_1, d_2, \cdots, d_n)$$

定义 2 集函数 $\mu: P(X)\to R$，若满足条件 $\mu(\varnothing)=0$，则被称为带符号模糊测度。

一个定义在 $P(X)$ 上的带符号模糊测度可以用一个(2^n-1)维的向量来描述，即$(\mu_1, \mu_2, \cdots, \mu_{2^n-1})$。其中，对于每个下标值 $h=1,2,\cdots,2^n-1$，如果 h 被表示为一串二进制数符 $h_n h_{n-1}, \cdots, h_1$，则有

$$\mu_h=\mu(\bigcup_{h_j=1}\{x_j\})$$

例如，当 $n=3$ 时，下标 1,2,…,7 被分别描述为二进制数符串 001,010,…,111，则 $\mu_1=\mu(\{x_1\})$，$\mu_2=\mu(\{x_2\})$，$\mu_3=\mu(\{x_1, x_2\})$，$\mu_4=\mu(\{x_3\})$，$\mu_5=\mu(\{x_1, x_3\})$，$\mu_6=\mu(\{x_2, x_3\})$，$\mu_7=\mu(\{x_1, x_2, x_3\})$。

根据 Zadeh 的概率测度理论，任何一个建立在 $P(X)$ 上的带符号模糊测度亦能扩展到模糊幂集 $F(X)$ 上。

令 μ 是定义在 $P(X)$ 上的一个带符号模糊测度，对于任何一个具有隶属度函数 $m_{\widetilde{A}}(x): X\to[0,1]$ 的模糊集合 $\widetilde{A}\in F(X)$，定义

$$\tilde{\mu}(\widetilde{A})=\int m_{\widetilde{A}}\,\mathrm{d}\mu \tag{1}$$

则对于任何一个 $A\in P(X)$，有 $\tilde{\mu}(A)=\int \chi_A \mathrm{d}\mu=\mu(A)$，其中

$$\chi_A(x)=\begin{cases}1, & \text{if } x\in A \\ 0, & \text{if } x\notin A\end{cases}$$

是 A 的特征函数。实际上，在 $P(X)$ 上 $\tilde{\mu}$ 与 μ 是一致的，即 $\tilde{\mu}$ 是 μ 从 $P(X)$ 到 $F(X)$ 的扩展，因此，可称 $\tilde{\mu}$ 是模糊集合的带符号模糊测度。

例 1 令 $X=\{x_1, x_2, x_3\}$，设定义在 X 上的带符号模糊测度 μ 定义为 $\mu(\varnothing)=0$，$\mu(\{x_1\})=1$，$\mu(\{x_2\})=1$，$\mu(\{x_1, x_2\})=3$，$\mu(\{x_3\})=2$，$\mu(\{x_1, x_3\})=1$，$\mu(\{x_2, x_3\})=4$，$\mu(\{x_1, x_2, x_3\})=5$。设 X 的一个模糊集合 $\widetilde{A}=(0.5, 1, 0.25)$，即

$$m_{\widetilde{A}}(x)=\begin{cases}0.5, & \text{if } x=x_1 \\ 1, & \text{if } x=x_2 \\ 0.25, & \text{if } x=x_3\end{cases}$$

根据模糊集合的带符号模糊测度的定义，有

$$\begin{aligned}\mu(\widetilde{A}) &= \int m_{\widetilde{A}}\,\mathrm{d}\mu \\ &= m_{\widetilde{A}}(x_3)\mu(\{x_1, x_2, x_3\})+[m_{\widetilde{A}}(x_1) \; m_{\widetilde{A}}(x_3)]\mu(\{x_1, x_2\})+[m_{\widetilde{A}}(x_2) \; m_{\widetilde{A}}(x_1)]\mu(\{x_2\}) \\ &= 0.25\cdot 5+(0.5-0.25)\cdot 3+(1-0.5)\cdot(-1) \\ &= 1.5\end{aligned}$$

4 模糊值方程的 *α-cut*

设 X 为一有限集合，记为 $X=\{x_1, x_2, \cdots, x_n\}$。令 $\tilde{f}$ 是定义在 X 上的一个模糊值方程，其值域是所有规则模糊数集合的一个子集。方程 $\tilde{f}$ 可表示为$(m_1, m_2, \cdots, m_n)$，其中 m_i 是 $f(x_i)$ 的隶属度函数，$i=1,2,\cdots,n$。

定义 3 对任何 $\alpha\in R$，模糊值方程 $\tilde{f}=(m_1, m_2, \cdots, m_n)$ 的 *α-cut* 记为 $\widetilde{F}_\alpha$，是 X 的一个模糊子集，其隶属度函数 $m_{\widetilde{F}_\alpha}$ 在每个 x_i 处具有隶属度函数：

$$m_{\widetilde{F}_\alpha}(x_i)=\begin{cases}\dfrac{\int_\alpha^\infty m_i(t)\mathrm{d}t}{\int_{-\infty}^\infty m_i(t)\mathrm{d}t}, & \text{if } \int_{-\infty}^\infty m_i(t)\neq 0 \\ \max\limits_{t\geqslant\alpha} m_i(t), & \text{otherwise}\end{cases}$$

$i=1,2,\cdots,n$

从定义 3 可看出，一个模糊值方程 $\tilde{f}$ 在 x_i 处的 *α-cut* 是由 $\tilde{f}(x_i)$ 的隶属度函数曲线在该 α 右方的面积相对整个隶属度函数曲线总面积的百分比来确定的。定义 3 亦可扩展到实数值方程的 *α-cut*，此时，$\tilde{f}(x_i)$ 的隶属度函数曲线所占的面积是 0，因此上述提到的百分比会出现 0/0 的情况。在此特殊情况下，需定义 $m_{\widetilde{F}_\alpha}(x_i)=\max\limits_{t\geqslant\alpha} m_i(t)$，其中

$$m_i(t)=\begin{cases}1, & \text{if } t=c \\ 0, & \text{if } t\neq c\end{cases}$$

当 $\tilde{f}(x_i)$ 为一区间数$[a,b]$，关于 $\tilde{f}$ 在 x_i 处的 *α-cut* 的隶属度函数可具体描述为：

$$m_{F_\alpha}(x_i)=\begin{cases}1, & \text{if } \alpha<a \\ \dfrac{b-\alpha}{b-a}, & \text{if } \alpha\in[a,b] \\ 0, & \text{if } \alpha>b\end{cases}$$

例 2 令 $X=\{x_1, x_2, x_3\}$，$\tilde{f}$ 是定义在 X 上的模糊值方程，设 $\tilde{f}(x_1)=[1 \quad 1.5 \quad 2 \quad 2.5]$，$\tilde{f}(x_2)=[4 \quad 4.5 \quad 5 \quad 5]$，$\tilde{f}(x_3)=[3 \quad 3.5 \quad 4 \quad 4.5]$。如图 1 所示，根据模糊值方程的 *α-cut* 定义，当 $\alpha=2$ 时 $\widetilde{F}_\alpha=(0.25, 1, 1)$；当 $\alpha=3.5$ 时 $\widetilde{F}_\alpha=(0, 1, 0.75)$。假如存在一个定义在 X 上的带符号模糊测度，其取值如例 1 所示，则根据模糊集合的带符号模糊测度的定义，有

$$\mu(\widetilde{F}_2)=\int m_{\widetilde{F}_2}\,\mathrm{d}\mu=0.25\cdot 5+(1-0.25)\cdot 4+(1-1)\cdot(-1)=4.25$$

$$\mu(\widetilde{F}_{3.5})=\int m_{\widetilde{F}_{3.5}}\,\mathrm{d}\mu=0\cdot 5+(0.75-0)\cdot 4+(1-0.75)\cdot(-1)=2.75$$

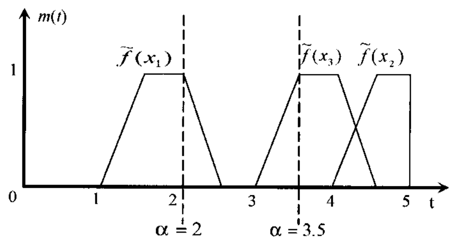


图 1　实例 2 中模糊值方程 $\tilde{f}$ 的隶属度函数

5　Choquet 积分的模糊化扩展 II 型

在对 Choquet 积分进行模糊化扩展之前，需先回顾传统 Choquet 积分的相关概念。

定义 4　设 $\langle X,P(X)\rangle$ 为可测空间，μ 是定义在 $P(X)$ 上的带符号的模糊测度，$f:X\to(-\infty,\infty)$ 为一实数方程，则 f 关于 μ 的 Choquet 积分定义为：

$$\int f\mathrm{d}\mu=\int_{-\infty}^{0}[\mu(F_\alpha)-\mu(X)]\mathrm{d}\alpha+\int_{0}^{\infty}\mu(F_\alpha)\mathrm{d}\alpha$$

式中，$F_\alpha=\{x\mid f(x)\geqslant\alpha,x\in X\}$，$\alpha\in(-\infty,\infty)$。

当 X 是一有限集合，记为 $X=\{x_1,x_2,\cdots,x_n\}$ 且 $f:X\to(-\infty,\infty)$ 时，Choquet 积分的计算公式相应变为：

$$\int f\mathrm{d}\mu=\sum_{i=1}^{n}[f(x_i{}')-f(x_{i-1}{}')]\cdot\mu(\{x_i{}',x_{i+1}{}',\cdots,x_n{}'\})$$

式中，$(x_1{}',x_2{}',\cdots,x_n{}')$ 是 $\{x_1,x_2,\cdots,x_n\}$ 的某种排列，其满足条件 $f(x_1{}')\leqslant f(x_2{}')\leqslant\cdots\leqslant f(x_n{}')$，且 $f(x_0{}')=0$。

为计算方便，Wang 在参考文献[7]中将实数函数的 Choquet 积分的计算转化为两个维数 (2^n-1) 的向量之内积，如下所示：

$$\int f\mathrm{d}\mu=\sum_{j=1}^{2^n-1}z_j\mu_j \tag{2}$$

式中，$j=1,2,\cdots,2^n-1$。

$$z_j=\begin{cases}\min\limits_{\{i\mid j_i=1\}}f(x_i)-\max\limits_{\{i\mid j_i=0\}}f(x_i), & \text{若所得值}>0\\ & \text{或 } j=2^n-1\\ 0, & \text{其它情况}\end{cases} \tag{3}$$

式中，实数 j 以二进制数形式表述为 $j_nj_{n-1}\cdots j_1$。

当 X 是拥有 n 个元素的有限集合，它可被描述为 $X=\{x_1,x_2,\cdots,x_n\}$。为方便起见，$\mu(\{x_{j1},x_{j2},\cdots,x_{jm}\})$ 记作 μ_j，这里 $j=\sum_{k=1}^{m}2^{j_k-1}$，且 $\{j_1,j_2,\cdots,j_m\}$ 是 $\{1,2,\cdots,m\}$ 的一个子集。亦可直接表示为：$\mu_0=(\varnothing)$，$\mu_1=\mu(\{x_1\})$，$\mu_2=\mu(\{x_2\})$，$\mu_3=\mu(\{x_1,x_2\})$，$\mu_4=\mu(\{x_3\})$，$\mu_5=\mu(\{x_1,x_3\})$，$\mu_6=\mu(\{x_2,x_3\})$，$\mu_7=\mu(\{x_1,x_2,x_3\})$，$\cdots$，$\mu_{2^n-1}=\mu(\{x_1,x_2,\cdots,x_n\})$。

令 $\tilde{f}$ 是定义在集合 X 上的模糊值方程，其值域为所有梯形模糊数集合，令 μ 是定义在 $P(X)$ 上的带符号模糊测度。从第 3 节的论述可见，带符号模糊测度可被扩展成模糊集合的带符号模糊测度。因此，可定义 Choquet 积分的模糊化扩展 II 型（模糊值方程 $\tilde{f}$ 关于带符号模糊测度 μ 的 Choquet 积分）如下：

$$\int\tilde{f}\mathrm{d}\mu=\int_{-\infty}^{0}[\mu(\tilde{F}_\alpha)-\mu(X)]\mathrm{d}\alpha+\int_{0}^{\infty}\mu(\tilde{F}_\alpha)\mathrm{d}\alpha \tag{4}$$

由于无法对模糊值方程 $\tilde{f}$ 在集合 $X=\{x_1,x_2,\cdots,x_n\}$ 上的 $\tilde{f}(x_i)$ 进行精确的大小排序，因此，Choquet 积分的模糊化扩展 II 型虽然在定义上与传统的 Choquet 积分非常相似，但在计算上却不能延续传统 Choquet 积分基于排序的计算方法。为简化计算，需先分析一个 Choquet 积分的模糊化扩展 II 型的重要性质。

定理 1　令 μ 是定义在 $F(X)$ 上的带符号模糊测度，$\tilde{f}$ 为定义在集合 X 上的一个模糊值方程，对任何实数常数 c，

$$\int\tilde{f}\mathrm{d}\mu=\int(\tilde{f}-c)\mathrm{d}\mu+c\cdot\mu(X)$$

证明：令 $\tilde{g}=\tilde{f}-c$，则 $\tilde{g}$ 也是一个模糊值方程。对任何实数值 α，$\tilde{g}$ 的 α-cut，$\tilde{G}_\alpha$，满足 $\tilde{G}_\alpha=\tilde{F}_{\alpha+c}$，即 $\tilde{G}_{\alpha-c}=\tilde{F}_\alpha$。因此，令 $\beta=\alpha-c$，则有：

$$\begin{aligned}\int\tilde{f}\mathrm{d}\mu&=\int_{-\infty}^{0}[\mu(\tilde{F}_\alpha)-\mu(X)]\mathrm{d}\alpha+\int_{0}^{\infty}\mu(\tilde{F}_\alpha)\mathrm{d}\alpha\\ &=\int_{-\infty}^{0}[\mu(\tilde{G}_{\alpha-c})-\mu(X)]\mathrm{d}\alpha+\int_{0}^{\infty}\mu(\tilde{G}_{\alpha-c})\mathrm{d}\alpha\\ &=\int_{-\infty}^{0}[\mu(\tilde{G}_{\alpha-c})-\mu(X)]\mathrm{d}(\alpha-c)+\\ &\quad\int_{0}^{\infty}\mu(\tilde{G}_{\alpha-c})\mathrm{d}(\alpha-c)\\ &=\int_{-\infty}^{-c}[\mu(\tilde{G}_\beta)-\mu(X)]\mathrm{d}\beta+\int_{-c}^{\infty}\mu(\tilde{G}_\beta)\mathrm{d}\beta\\ &=\int_{-\infty}^{-c}[\mu(\tilde{G}_\beta)-\mu(X)]\mathrm{d}\beta+\int_{-c}^{0}\mu(\tilde{G}_\beta)\mathrm{d}\beta+\\ &\quad\int_{0}^{\infty}\mu(\tilde{G}_\beta)\mathrm{d}\beta-\int_{-c}^{0}\mu(X)\mathrm{d}\beta+\int_{-c}^{0}\mu(X)\mathrm{d}\beta\\ &=\int_{-\infty}^{0}[\mu(\tilde{G}_\beta)-\mu(X)]\mathrm{d}\beta+\int_{0}^{\infty}\mu(\tilde{G}_\beta)\mathrm{d}\beta+\\ &\quad\int_{-a}^{0}\mu(X)\mathrm{d}\beta\\ &=\int\tilde{g}\mathrm{d}\mu+c\cdot\mu(X)\\ &=\int(\tilde{f}-c)\mathrm{d}\mu+c\cdot\mu(X)\end{aligned}$$

证毕。

如前所述，本文主要考虑梯形模糊数。由于梯形模糊数的支集（support set）是有限集，故对于任何 $\tilde{f}(x_i)$，都存在其支集的左边界 a_{il}，$i=1,2,\cdots,n$，因此可构造一个非负的模糊值方程 $\tilde{g}=\tilde{f}-c$，其中 $c=\min\limits_{1\leqslant i\leqslant n}a_{il}$，则 Choquet 积分的模糊化扩展 II 型的计算可简化为下式：

$$\int\tilde{f}\mathrm{d}\mu=\int_{0}^{\infty}\mu(\tilde{G}_\alpha)\mathrm{d}\alpha+c\cdot\mu(X)$$

式中，$\tilde{G}_\alpha$ 是 $\tilde{g}$ 的 α-cut。

下面考虑一般情况下的算法设计，即模糊值方程 $\tilde{f}$ 在每个特征量处取值模糊数 $\tilde{f}(x_i)$，$i=1,2,\cdots,n$。考虑一般的梯形模糊数，可以描述为 $\tilde{f}=([a_{1l}\ a_{1b}\ a_{1c}\ a_{1r}],[a_{2l}\ a_{2b}\ a_{2c}\ a_{2r}],\cdots,[a_{nl}\ a_{nb}\ a_{nc}\ a_{nr}])$ 对于任何 $\alpha\in R$，$\tilde{f}$ 的 α-cut 是一个 X 的模糊子集，其在特征量 x_i 处的隶属度为：

$$m_{\tilde{F}_\alpha}(x_i)=\begin{cases}1, & \text{当 }\alpha\leqslant a_{il}\\ 1-\dfrac{(\alpha-a_{il})^2}{(a_{ir}+a_{ic}-a_{il}-a_{ib})(a_{ib}-a_{il})}, & \text{当 }\alpha\in(a_{il},a_{ib}]\\ \dfrac{a_{ir}+a_{ic}-2\alpha}{a_{ir}+a_{ic}-a_{il}-a_{ib}}, & \text{当 }\alpha\in(a_{ib},a_{ic}]\\ \dfrac{(a_{ir}-\alpha)^2}{(a_{ir}+a_{ic}-a_{il}-a_{ib})(a_{ir}-a_{ic})}, & \text{当 }\alpha\in(a_{ic},a_{il}]\\ 0, & \text{当 }\alpha>a_{ir}\end{cases}$$

$i=1,2,\cdots,n$

在此情况下，因无法给出 $\mu(\tilde{F}_\alpha)$ 的解析式，故对 Choquet 积分的模糊化扩展 II 型的关键分量 $\int_{0}^{\infty}\mu(\tilde{F}_\alpha)\mathrm{d}\alpha$ 的计算造成一定困难。为此，本文提出以下的数值计算算法来逼近计算

Choquet 积分的模糊化扩展 II 型的积分值。

1)输入 n(X 中特征量数目)、K(α 值的分度数,默认值 100),模糊值方程 $\tilde{f}$ 取值 $\tilde{f}(x_i)=[a_{il}\quad a_{ib}\quad a_{ic}\quad a_{ir}]$,$i=1,2,\cdots,n$,带符号模糊测度 μ 取值 $\mu_j=\mu(\bigcup\limits_{i;frc(\frac{j}{2^i})\in[\frac{1}{2},1)}\{x_i\})$,$j=1,2,\cdots,2^n-1$。

2)检查 $a_{il}\leqslant a_{ib}\leqslant a_{ic}\leqslant a_{ir}$,$i=1,2,\cdots,n$ 是否成立。若成立,执行步骤 3);若不成立,则显示出错信息,需重新输入步骤 1)各项数值。

3)找出 $a=\min\limits_{1\leqslant i\leqslant n}a_{il}$,$b=\max\limits_{1\leqslant i\leqslant n}a_{ir}$,令 $\delta=\dfrac{b-a}{K}$。

4)令 $a_{il}=a_{il}-a$,$a_{ib}=a_{ib}-a$,$a_{ic}=a_{ic}-a$,$a_{ir}=a_{ir}-a$,$i=1,2,\cdots,n$;初始化 $\alpha=0$ 及 $S=\mu_{2^n-1}/2$。

5)$\alpha+\delta\Rightarrow\alpha$。

6)判断 $\alpha>b-a$? 若是,则 $\delta\cdot(S-\dfrac{\Delta S}{2})+a\cdot\mu_{2^n-1}\Rightarrow S$,输出 S,S 即为 $\int\tilde{f}\mathrm{d}\mu$ 的积分结果,程序结束;若非,则继续。

7)当 $i=1,2,\cdots,n$ 时,分别计算

$$c_i=\begin{cases}1, & \text{当 } \alpha\leqslant a_{il}\\ 1-\dfrac{(\alpha-a_{il})^2}{(a_{ir}+a_{ic}-a_{il}-a_{ib})(a_{ib}-a_{il})}, & \text{当 } \alpha\in(a_{il},a_{ib}]\\ \dfrac{a_{ir}+a_{ic}-2\alpha}{a_{ir}+a_{ic}-a_{il}-a_{ib}}, & \text{当 } \alpha\in(a_{ib},a_{ic}]\\ \dfrac{(a_{ir}-\alpha)^2}{(a_{ir}+a_{ic}-a_{il}-a_{ib})(a_{ir}-a_{ic})}, & \text{当 } \alpha\in(a_{ic},a_{il}]\\ 0, & \text{当 } \alpha>a_{ir}\end{cases}$$

8)将方程 $h(c_1,c_2,\cdots,c_n)$ 看作是建立在 X 上的实数值方程,计算 $\Delta S=\int h\mathrm{d}\mu$,即 $\Delta S=\sum\limits_{j=1}^{2^n-1}z_j\cdot\mu_j$,其中

$$z_j=\begin{cases}\min\limits_{i;frc(\frac{j}{2^i})\in[\frac{1}{2},1)}c_i-\max\limits_{i;frc(\frac{j}{2^i})\in[0,\frac{1}{2})}c_i, & \text{if it is}>0\\ 0, & \text{otherwise}\end{cases},$$

$j=1,2,\cdots,2^n-1$

9)$S+\Delta S\Rightarrow S$,转步骤 5)。

例 3 某杂志通过 3 项指标,即原创性(originality)、重要性(significance)及文字表达(presentation)来评价投稿的总体质量。每项指标用模糊概念"很差(bad)"、"较弱(weak)"、"一般(fair)"、"良好(good)"、"优秀(excellent)"来描述,以实数区间[0,5]作为梯形模糊数的支集来定义上述 5 个模糊概念,即

$\tilde{a}_b=[0\ 0\ 1\ 1.5]$,$\tilde{a}_w=[1\ 1.5\ 2\ 2.5]$,$\tilde{a}_f=[2\ 2.5\ 3\ 3.5]$

$\tilde{a}_g=[3\ 3.5\ 4\ 4.5]$,$\tilde{a}_e=[4\ 4.5\ 5\ 5]$

其隶属度函数如图 2 所示。用 x_1,x_2,x_3 来代替 3 项指标。每项指标对投稿评价的重要性以及它们交互影响投稿评价的重要性用一个带符号模糊测度 μ 来描述,其中,μ 定义在集合 $X=\{x_1,x_2,x_3\}$ 的 $P(X)$ 上,其值设定为 $\mu_1=0.2$,$\mu_2=0.3$,$\mu_3=0.8$,$\mu_4=0.1$,$\mu_5=0.4$,$\mu_6=0.4$,$\mu_7=1$。

假设某篇投稿被审稿人评价为 originality="excellent",significance="fair",presentation="weak",将此审稿人的评价看作是建立在 $X=\{x_1,x_2,x_3\}$ 上的模糊值方程 $\tilde{f}=(\tilde{a}_e,\tilde{a}_f,\tilde{a}_w)$,则对于该篇投稿的总体评价值可以用 Choquet 积分的模糊化扩展 II 型,即基于 μ 的模糊值方程 $\tilde{f}$ 的 Choquet 积分 $\int\tilde{f}\mathrm{d}\mu$ 来求得。利用本节提供的算法,可得到 $\int\tilde{f}\mathrm{d}\mu$ 的相对精确的近似值:

$$\int\tilde{f}\mathrm{d}\mu\approx\begin{cases}2.92176, & \text{当 } K=100\\ 2.92222, & \text{当 } K=1000\end{cases}$$

假设另一篇投稿被审稿人评价为 originality="bad",significance="good",presentation="excellent",将此审稿人的评价看作是建立在 $X=\{x_1,x_2,x_3\}$ 上的模糊值方程 $\tilde{g}=(\tilde{a}_b,\tilde{a}_g,\tilde{a}_e)$,则

$$\int\tilde{g}\mathrm{d}\mu\approx\begin{cases}1.96618, & \text{当 } K=100\\ 1.96611, & \text{当 } K=1000\end{cases}$$

结果说明第一篇投稿比第二篇投稿在总体质量上优胜。

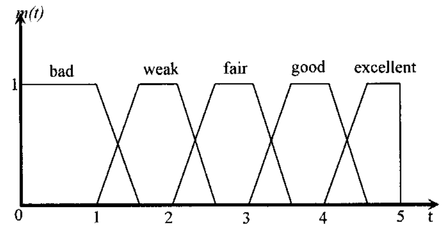


图 2 实例 3 中 $\tilde{a}_b,\tilde{a}_w,\tilde{a}_f,\tilde{a}_g,\tilde{a}_e$ 的隶属度函数

结束语 本文讨论了最常用的非线性积分 Choquet 积分的模糊化问题,即提出了 Choquet 积分的模糊化扩展 II 型模型。该模型作为一种直接处理模糊数据的聚合工具,可以用来解决模糊数据的非线性分类以及非线性回归等数据处理问题。建立在带符号模糊测度上的 Choquet 积分的模糊化扩展 II 型模型,其被积函数为模糊值函数,其积分结果为精确实数。尽管丢失了信息中的模糊度,但显然在处理决策问题时,聚合模型提供精确的聚合结果要比提供模糊的聚合结果有效得多,这一点可以从例 3 得到说明。

本文利用 Choquet 积分,将带符号模糊测度从全集(universal set)的幂集(power set)扩展到全集上所有模糊子集的集合(class of all subsets of universal set)上。实际上,任何一种建立在带符号模糊测度上的非线性积分都可以用来处理这样的扩展,比如 Wang's integral[8]。这将在我们的后续工作中进行进一步讨论。

参 考 文 献

[1] Choquet G. Theory of Capacities [J]. Annales de l'Institut Fourier,5:131-295

[2] Halmos P R. Measure Theory [M]. New York: Springer-Verlag New York Inc,1974:324

[3] Murofushi T, Sugeno M, Machida M. Non Monotonic Fuzzy Measures and the Choquet Integral [J]. Fuzzy Sets and Systems,1994,64:73-86

[4] Wang Z, Klir G J. Fuzzy Measure Theory [M]. New York: Springer-Verlag New York Inc,1992:363

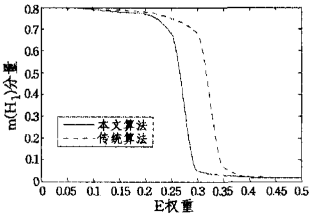
[5] Leung K S,Wong M L,Lam W,et al. Learning Nonlinear Multiregression Networks Based on Evolutionary Computation [J]. IEEE Transaction on Systems, Man and Cybernetics, Part B, 2002,32:630-644

[6] Xu K,Wang Z,Leung K S. Classification by Nonlinear Integral Projections [J]. IEEE Transaction on Fuzzy Systems,2003,11: 187-201

[7] Wang Z. A new genetic algorithm for nonlinear multiregressions based on generalized Choquet integrals [C] // Proceedings on 12th IEEE International Conference on Fuzzy Systems, 2003. 2003,2:819-821

(下转第 124 页)

果的突变要早于虚线表示结果的突变。通过与实际样本数据和初始 BPA 分布的对比，实线所表达的融合结果更为合理，同时也说明，本文采用的算法收敛速度快，判别能力强。



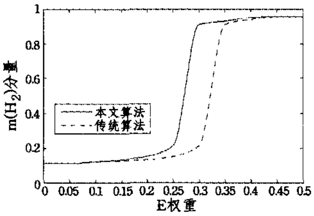


图 10 证据 E 与 $m(H_1)$ 的关系图　图 11 证据 E 与 $m(H_2)$ 的关系图

结束语　本文通过对 D-S 证据理论现有算法和信息融合过程的分析，提出了一种基于 D-S 证据理论的层次式信息融合算法，给出了算法的模型。该算法实现了信息融合的系统性，通过正交向量在三维信度区间的映射，解决了在层次式融合中确定高层概率分布的问题，并针对证据冲突情况做出了相应的算法修正。仿真结果表明，该算法在虚警概率较低的情况下具有较高的检测概率，且收敛速度快，优于目前现有的传统 D-S 证据理论算法。

参 考 文 献

[1] 张池平. 多传感器信息融合方法及其在空间目标识别中的应用[D]. 哈尔滨：哈尔滨工业大学，2006

[2] Findeisen R L, Allgower F, Foss B A. State and Output Feedback Nonlinear Model Predictive Control: An Overview[J]. European Journal of Control, 2003(9): 179-195

[3] 孙锐. 基于 D-S 证据理论的信息融合及在可靠性数据处理中的应用研究[D]. 成都：电子科技大学，2011

[4] 龚本刚. 基于证据理论的不完全信息多属性决策方法研究[D]. 合肥：中国科学技术大学，2007

[5] 蒋黎明，何佳浪，张宏. D-S 证据理论中一种新的冲突证据融合方法[J]. 计算机科学，2011，38(4)：236-238

[6] Heeyoul C, Seungjin C, Yoonsuck C. Probabilistic Combination of Multiple Evidence[C] // Proceedings of the 16th International Conference on Neural Information Proceesing. Bangkok, Thailand, 2009: 302-311

[7] 徐琰珂，梁晓庚，贾晓洪. 利用模糊证据理论的信息融合方法及其应用[J]. 哈尔滨工业大学学报，2012，44(3)：107-111

[8] Xu Xiao-bin, Wen Cheng-lin, Li Zhi-liang. A new method for constructing fuzzy evidence based on the non-consonantrandom set[J]. Journal of electronics, 2009, 26(1): 31 -37

[9] 何小飞. 基于贝叶斯网络和 D-S 证据理论的电网故障诊断模型研究[D]. 成都：西南交通大学，2010

[10] Thorp J S, Phadke A G. Expose Hidden Failures to Prevent caseading outages[J]. IEEE TranSonpower Delive, 2008, 10: 54-55

[11] 方阳. 基于层次分析法和 D-S 证据理论的电信网网络安全风险评估模型的研究与应用[D]. 北京：北京邮电大学，2010

[12] 李月，徐余法，陈国初，等. D-S 证据理论在多传感器故障诊断中的改进及应用[J]. 东南大学学报，2011，41(9)：102-106

[13] 李艳娜，乔秀全，李晓峰. 基于证据理论的上下文本体建模以及不确定性推理方法[J]. 电子与信息学报，2010，32(8)：1806-1811

[14] 刘勇生. 基于证据理论的网络安全风险组合评估方法[J]. 计算机仿真，2012，29(1)：106-109

[15] 肖文. 基于证据理论的多属性决策关联问题研究[D]. 南昌：江西财经大学，2010

[16] 黄蕊. 下一代电信网安全评估指标体系的研究[D]. 北京：北京邮电大学，2010

[17] Murphy C. Combining of Belief Functions When Evidence Conflicts [J]. Decision Support Systems, 2000, 29(1): 1-9

[18] Han Li, Shi Li-ping. Approach to evidence combination based on rough set[C] // International Conference on Electronic Computer Technology. Macau, China, 2009: 693-697

[19] Yager R. On the Dempster Shafer Framework and New Combination Rules[J]. Information Sciences, 1987, 41(2): 93-137

[20] Dubois D, Prade H. Representation and Combination of Uncertainty with Belief Functions and Possibility Measures[J]. Computational Intelligence, 1988, 4(3): 244-264

[21] Smarandaceh F, Dezert J. Proportional conflict redistribution rules for information fusion[M]. Smarandache F, Dezert J, eds. Vol. 2. Rehoboth: American Research Press, 2006: 3-68

[22] 李建平. 面向异构数据源的网络安全态势感知模型与方法研究[D]. 哈尔滨：哈尔滨工业大学，2010

[23] Boujelben M A, Smet Y D, Frikha A, et al. Building abinary outranking relation in uncertain, imprecise and multi-experts contexts: The application of evidence theory [J]. International Journal of Approximate Reasoning, 2009, 50(8): 1259-1278

[24] Chin K S, Yang J B, Guo M, et al. Evidential-Reasoning-Interval-Based Method for New Producet Design Assessment[J]. IEEE Transaetions on Engineering Management, 2009, 56(1): 142-156

[25] 张彦峰，何佩琨. 一种改进的 AdaBoost 算法——M-Asy AdaBoost[J]. 北京理工大学学报，2011，31(1)：64-73

[26] 冷宜兵，王平，张立. 证据理论合成准则的一种新算法及其验证[J]. 计算机仿真，2010，27(2)：162-165

[27] Moore A W, Zuev D. Internet traffic classification using bayesian analysis techniques[C] // Internet Traffic Classification Using Bayesian Analysis Techniques in the Proceedings of the ACM SIGMETRICS. Banff, Canada, 2005

（上接第 108 页）

[8] Buckley J J, Eslami E. An Introduction to Fuzzy Logic and Fuzzy Sets[M]. Physica-Verlag Heidelberg, 2002: 296

[9] 陶长琪，凌和良. 基于 Choquet 积分的模糊数直觉模糊数多属性决策方法[J]. 控制与决策，2012，27(9)：1381-1386

[10] 陈亚婷，吴博，张国春. 基于一种推广的 Choquet 积分的回归模型[J]. 河北大学学报：自然科学版，2010，30(4)：353-355

[11] 安相华，冯毅雄，谭建荣. 基于 Choquet 积分与证据理论的产品方案协同评价方法[J]. 浙江大学学报：工学版，2012，46(1)：163-169

[12] 张磊，樊治平，乐琦. 基于 Choquet 积分的综合风险测评方法[J]. 东北大学学报：自然科学版，2010，31(11)：1665-1668

[13] 王坚强，聂荣荣. 准则关联的直觉模糊多准则决策方法[J]. 控制与决策，2011，26(9)：1348-1352

[14] 郑三元，杨蓉. Choquet 积分的模糊化扩展 I 型 [J]. 深圳大学学报：理工版