

非加法的測度による新しい積分の提案

本田あおい (九州工業大学・情報工学部)

1 はじめに

ファジィ測度は加法性を持たず, 単調性のみを仮定した測度である [1]. そのため, ファジィ測度による積分を行う際にルベーグ積分を適用することができない. いくつかのファジィ測度による積分 (ファジィ積分) が, これまで提案・発見されてきたが, 中でもショケ積分はその性質の良さからファジィ積分として最も広く利用されている [2]. 本論文では新しいファジィ積分として包除積分の概念を提案する. この積分のクラスはショケ積分も含まれる自由度の高いクラスであり, 目的に応じて柔軟に使い分けることができる. また, この積分を用いた主観的意思決定モデルを提案し映像評価の実データに適用し有効性を検証する.

2 準備

本論文を通して, 全体集合を $N := \{1, 2, \dots, n\}$, $2^N := \{A \mid A \subseteq N\}$ とする. また $\min$ 演算, $\max$ 演算をそれぞれ $\wedge, \vee$, 集合 A に対して A の要素の個数を $|A|$ で表すことにする.

定義 1. 集合関数 $v: 2^N \rightarrow [0, 1]$ が次の 2 条件を満たすとき, ファジィ測度とよぶ.

1. $v(\emptyset) = 0, v(N) = 1$,
2. $A, B \in 2^N, A \subseteq B$ ならば $v(A) \leq v(B)$.

定義 2. v を 2^N 上のファジィ測度とし f を N 上の非負関数とする. f の v についてのショケ積分は次式で定義される.

$$(C) \int f dv := \sum_{i=1}^n (f(\sigma(i)) - f(\sigma(i-1)))v(\{\sigma(i), \dots, \sigma(n)\}),$$

ただし σ は $f(\sigma(1)) \leq \dots \leq f(\sigma(n)), f(\sigma(0)) := 0$ を満たす N 上の置換.

定義 3. v を 2^N の集合関数とする. v のメビウス変換, m^v は次で定義される. 任意の $A \in 2^N$ に対して,

$$m^v(A) := \sum_{B \subseteq A} (-1)^{|A \setminus B|} v(B). \quad (1)$$

v と m^v は一対一対応, すなわち逆変換が存在し, 任意の $A \in 2^N$ に対して

$$v(A) = \sum_{B \subseteq A} m^v(B). \quad (2)$$

定義 4. v を 2^N 上のファジィ測度, k を正整数とする. 任意の $A \in 2^N$ に対して $|A| > k$ ならば $m^v(A) = 0$ が成り立つとき v を k -加法的測度とよぶ.

k -加法的測度は k 個の以下の要素間に対してのみ相互作用が働く測度とみなすことができる.

v についてのショケ積分は v のメビウス変換の線形結合で別表現できる.

命題 5 ([3]). v についてのショケ積分は v のメビウス変換 m^v を用いて次のように別表現できる.

$$(C) \int f dv = \sum_{A \subseteq N} \left(\bigwedge_{i \in A} f(i) \right) m^v(A). \quad (3)$$

定義 6 (T ノルム). 2項演算 $\otimes : [0, 1] \times [0, 1] \rightarrow [0, 1]$ が次の4条件をみたすとき T ノルムと呼ぶ:

1. 任意の $x > 0$ に対して $0 \otimes 0 = 0, x \otimes 1 = x$,
2. $x \leq y$ ならば $x \otimes z \leq y \otimes z$,
3. $x \otimes y = y \otimes x$,
4. $(x \otimes (y \otimes z)) = ((x \otimes y) \otimes z)$.

4 より $\otimes : [0, 1] \times [0, 1] \rightarrow [0, 1]$ は $\otimes : [0, 1]^n \rightarrow [0, 1]$ に拡張できる.

3 包除積分

本節では新しいファジィ積分として包除積分を提案する.

定義 7. v を 2^N 上のファジィ測度, $\otimes : \cup_{k=1}^n [0, 1]^k \rightarrow [0, 1]$ を T ノルム, $f, f(i) \leq 1$ を N 上の非負関数とする. f の v についての包除積分を次で定義する.

$$\otimes \int f dv := \sum_{A \subseteq N} \left\{ \sum_{B \supseteq A} \left((-1)^{|B \setminus A|} \otimes_{i \in B} f(i) \right) \right\} v(A),$$

ただし $\bigotimes_{i \in B} w_i = \otimes \{w_i \mid i \in B\}, B \subseteq N$.

例 8. T ノルムとして積演算を採用すると, 次のようになる.

$${}^{(M)} \int f \, dv := \sum_{A \subseteq N} \left\{ \sum_{B \supseteq A} \left((-1)^{|B \setminus A|} \prod_{i \in B} f(i) \right) \right\} v(A)$$

これを積包除積分と呼ぶことにする.

命題 9. ファジィ測度 v についての包除積分は v のメビウス変換 m^v を用いて次のように別表現できる.

$$\otimes \int f \, dv = \sum_{A \subseteq N} \left(\bigotimes_{i \in A} f(i) \right) m^v(A).$$

この別表現を用いると, 包除積分は $\sum_{A \subseteq N} (\bigotimes_{i \in A} f(i))$ の線形結合で表わすことができる.

例 10. 積包除積分は次のように別表現できる.

$${}^{(M)} \int f \, dv = \sum_{A \subseteq N} \left(\prod_{i \in A} f(i) \right) m^v(A).$$

系 11. ショケ積分は $\otimes$ として $\min$ 演算 $\wedge$ を用いた包除積分である.

4 主観的評価問題への応用

商品等に対する人間の主観的評価のモデル化を考える. ある商品があり, その商品の評価を決定する n 個の要因 $X_1, X_2, \dots, X_n$ があるとする. この商品について $X_1, X_2, \dots, X_n$ の評点を少しずつ変えた M 個の商品サンプルがあり, これらの商品サンプルの主観的な評価がデータとして得られているとする (表 1). 商品サンプル $j, j = 1, \dots, M$ の各要因の評点を $x_1^j, x_2^j, \dots, x_n^j$, 商品の主観的な総合評価を y^j で表わしたものが 1 組のデータで, このデータがサンプルの個数分得られている. つまり $n+1$ 番目に商品の総合評価を加えた $n+1$ 個組のデータ $(x_1^1, x_2^1, \dots, x_n^1, y^1), (x_1^2, x_2^2, \dots, x_n^2, y^2), \dots, (x_1^M, x_2^M, \dots, x_n^M, y^M)$ から成っている. 1 組のデータのうち, 各要因の評点は何らかの測定値のような客観的な値であり, 商品の総合評価のみが主観的な値である. $x_1^j, \dots, x_n^j$ をそれぞれ最大値 1, 最小値 0 に正規化し, これらを $f^j(X_1), \dots, f^j(X_n)$ とする (表 2),

$$f^j(X_i) = \frac{x_i^j - \min_j(x_i^j)}{\max_j(x_i^j) - \min_j(x_i^j)}. \quad (4)$$

表 1: Crude data

No.	x_1	x_2	x_3	x_4	x_5	evaluation y
1	44.46	0.93	0.35	0.77	1.00	4.79
2	41.24	1.38	0.42	1.36	1.00	4.58
3	39.06	1.33	0.45	1.91	1.00	4.54
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
j	36.49	0.72	0.35	2.77	6.85	4.04
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
M	42.58	0.00	0.00	100.00	69.96	3.42

ただし要因のうち、値が小さいほど商品の評価がよくなるようなものに関しては大小を反転させることとする。

$$f^j(X_i) = 1 - \frac{x_i^j - \min_j(x_i^j)}{\max_j(x_i^j) - \min_j(x_i^j)}. \quad (5)$$

表 2: Normalized data

No.	$f(X_1)$	$f(X_2)$	$f(X_3)$	$f(X_4)$	$f(X_5)$	y
1	0.8272	0.7362	0.7127	0.9974	1.00	4.79
2	0.7270	0.6057	0.6751	0.9915	1.00	4.58
3	0.6591	0.6208	0.6751	0.9915	1.00	4.54
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
j	0.5789	0.7952	0.7821	0.9850	0.6735	4.04
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
M	0.7688	1.00	0.9771	0.00	0.8233	3.42

これらの正規化された n 個の客観的数値 及び 1 個の主観的数値の組を用いて解析を行う。

今回実際に用いたのは映像サービスに関するデータである。映像品質を決定する要因は次の 5 つである：

- X_1 := 映像全体の劣化量,
- X_2 := ブロック歪による劣化量,
- X_3 := ボケに伴う劣化量,
- X_4 := 局所的な空間劣化量算出,
- X_5 := フリーズ劣化量算出.

これらの5つの量を少しずつ変えた映像サンプル 1140 件に対して、それぞれ主観的に総合的な映像の評価を行っている。表 1,2 は実は今回実際に扱ったデータの一部である。\$X_1\$ は (4) 式を用いて正規化、\$X_2, X_3, X_4, X_5\$ は値が小さいほど映像品質は高くなるので (5) を用いて正規化している。

4.1 線形重回帰分析

このようなデータの解析には線形重回帰分析が広く用いられている。

$$y = \sum_{i=1}^n f(X_i) a_i + e$$

線形重回帰モデルではモデルの決定にはパラメータとして係数 \$a_1, \dots, a_n\$ と切片 \$e\$ を求めればよい。パラメータの決定には最小二乗法を用いる。

このモデルの利点は以下の通りである：

1. 最小二乗法でパラメータを簡単に決定でき、推定されたパラメータは最尤推定値となる。
2. 古くから広く使われているモデルであり、専用のツールやソフトウェアが多数存在し手軽に利用することができる。

また短所は以下の通りである。

1. このモデルでは説明できないような事例が多数存在することがわかっている。具体的には、総合評価には各要因間の相互作用がある場合が一般的であるが、このモデルでは各要因間の相互作用を表現することができない。実際、今回の映像サービス評価では線形重回帰モデルでは不十分であるとされており様々なモデルが提案されている。([6])

4.2 提案モデル – 包除モデル–

\$X = \{X_1, \dots, X_n\}\$ を要因全体の集合とし、\$v\$ を \$2^X\$ 上のファジィ測度とする。\$v\$ は要因の重視度を表す。モデルは次で与えられる。

$$\begin{aligned} z &= \otimes \int f dv + e \\ &= \otimes \int f dv = \sum_{A \subseteq N} \left(\bigotimes_{i \in A} f(i) \right) m^v(A) + e. \end{aligned}$$

このモデルの利点は以下の通りである。

1. このモデルでは人の主観をファジィ測度で表している. 要因間の相互作用を表現することができる. 必要ならば, 得られた m^v をメビウス逆変換することにより v を求めることが可能である. この v を用いて各項目の重視度や各項目間の相互関係を調べるができる.
2. 別表現を用いることにより, パラメータの線形結合で表わすことができる. その為, 線形重回帰分析のツールやソフトウェアを利用して誰にでも簡単にパラメータを決定することができる.

短所は以下の通りである.

1. 記述力が増した分, パラメータの個数は増大し, パラメータは 2^n 個となる. X の要素の個数が増えると爆発的に増大する. これらのパラメータを適切に決定するためには膨大な組のデータが必要となる.
2. T -ノルムとして $\wedge$ を採用した場合, つまりシヨケ積分を採用した場合は各パラメータに対するデータの独立性が保ちにくい. 例えば $f(i_0)$ が全体的に大きな値を取りやすいとする. このとき $\bigwedge_{i \neq i_0} f(i)$ と $\bigwedge_i f(i)$ は一致してしまう. このような場合には最小二乗法でパラメータを決定することができない.

問題 1 を解消するために, ファジィ測度として k -加法的測度を利用することが考えられる. この場合モデルは

$$z = \sum_{\substack{A \subseteq N \\ |A| \leq k}} \left(\bigotimes_{i \in A} f(i) \right) m^v + e$$

となりパラメータの個数を, 一般のファジィ測度を遣う場合の 2^n 個から $nC_1 + nC_2 + \dots + nC_k + 1$ 個に減らすことができる. これは問題 2 の解消にもつながり, パラメータの個数が減った分, データの独立性も保ちやすくなる. さらに問題 2 を解消するためには, データの独立性を保ちやすい T ノルムを用いることが考えられる.

今回の実験は包除積分モデルとして, $\wedge$ 包除積分モデル (シヨケ積分モデル), 及び積包除モデルを用いる. またファジィ測度は 2 加法的測度を採用する.

5 実験結果

前節で述べた映像評価データに対して,

1. 線形重回帰モデル

2. 2 加法的 $\wedge$ 包除積分モデル (シヨケ積分モデル)

3. 2 加法的積包除積分モデル

を適用した. それぞれパラメータを決定後, y と $\hat{y}$ の推定値の相関係数 (R 値) と二乗誤差平均 (RMSE) を求めた. それぞれ以下の式で求められる:

$$R := \frac{\sum_{i=1}^M (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^M (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^M (\hat{y}_i - \bar{\hat{y}})^2}},$$

$$\text{RMSE} := \sqrt{\frac{1}{M} \sum_{i=1}^M (y_i - \hat{y}_i)^2}.$$

ただし y_i は真の値, $\hat{y}_i$ は y_i の推定値, $\bar{y}$ は $y_i, i = 1, \dots, M$ の平均値, $M = 1140$ である. R は 1 に近いほどよい推定値が得られていることを意味する. シヨケ積分モデル, 積包除積分ともに, 線形重回帰モデルよりよ

表 3: R 値と二乗誤差平均

モデル	R 値	二乗誤差平均
線形重回帰モデル	0.753	0.664
シヨケ積分モデル	0.814	0.464
積包除積分モデル	0.838	0.435

い推定値が得られている.

映像主観評価に関して次のモデルが国際電気通信連合電気通信標準化部門にて ITU-T 勧告として策定されている (Telecommunication Standardization Sector of International Telecommunication Union (ITU-T) J.247 Annex A [6]):

$$\begin{aligned}
 y &= a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 \log_{10}(x_5) \\
 &\quad + a_6(a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4) \log_{10}(x_5) + e \\
 &= a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x'_5 \\
 &\quad + (a_1 a_6) x_1 x'_5 + (a_2 a_6) x_2 x'_5 + (a_3 a_6) x_3 x'_5 + (a_4 a_6) x_4 x'_5 + e,
 \end{aligned}$$

ただし $x'_5 := \log_{10}(x_5)$. このモデルは, 係数に制約があるが $x_2 x_3, x_2 x_4, \dots, x_3 x_4$ の係数を 0 とした 2 加法的積包除モデルと見なすことができる. フリーズ劣化量算出 (x_5) は 1 以上の値をとり, 小さな値に集中しているので非減少関数 $\varphi(x) = \log_{10} x$ で変換すると適当に分散させることができる. この x'_5 を用いて, 線形重回帰モデル, シヨケ積分モデル, 積包除モデル, ITU-T 勧告 J.247 Annex A の 4 つのモデルを適用した結果が表 4 である.

表 4: R 値と二乗誤差平均

モデル	R 値	二乗誤差平均
線形重回帰モデル	0.807	0.596
シヨケ積分モデル	0.856	0.522
積包除モデル	0.850	0.532
ITU-T R. J.247 Annex A	0.844	0.542

x_5 をそのまま使用したときに比べ全体的に良い結果が得られており、特に2つの包除積分モデルはどちらも ITU-T 勧告 J.247 Annex A よりもよい結果となった。

6 まとめ

包除積分の提案と主観的映像評価問題への応用を行った。実データを用いて既存のモデルよりよい結果を得ることができ有効性が確認できたと言える。

今回は2加法的ファジィ測度を用いた。2加法的を外し、一般のファジィ測度を適用したところ、シヨケ積分の場合にデータの独立性が失われ係数を決定することができなかった。積包除ではパラメータを決定でき2加法的モデルよりよい推定結果が得られた。しかしながら今回のデータ件数1140件に対してこの場合のパラメータ32個は多すぎるであろうと考える。

また、映像評価問題に特化したモデルである国際電気通信連合電気通信標準化部門 ITU-T 勧告 J.247 Annex A よりよい推定値を得ることができることもわかった。このモデルが映像に関する主観評価を客観的によく近似できることはわかっているが、モデルの意味するところはわかっていたが、このモデルは積包除積分であると見なすことができ、パラメータを主観的評価尺度を表わすファジィ測度とすることで各パラメータの意味づけすることが可能である。これを用いて人間の主観評価プロセスの説明、さらには詳細な分析を行うことができるであろう。映像評価の他にも、各分野で使われているモデルに包除積分で説明できるものがあるかもしれない。

参考文献

- [1] M. Sugeno: Fuzzy measures and fuzzy integrals—a survey, *In: Gupta MM, Saridis GN, Gaines BR (eds) Fuzzy automata and decision processes*, pp. 89–102, North Holland, 1977.
- [2] G. Choquet, Theory of capacities, *Ann Inst Fourier*, Vol. 5, pp. 131–295, 1953.
- [3] M. Grabisch and C. Labreuche: A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid, *A Quarterly Journal of Operations Research*, Vol. 6, pp. 1–44, Springer, 2008.
- [4] 中島信之: t-ノルムの全て, 三恵社, 2001.
- [5] P.L. Hammer and S. Rudeanu: *Boolean Methods in Operations Research and Related Areas*, Springer, 1968.
- [6] ITU-T Recommendation J.247: Objective perceptual multimedia video quality measurement in the presence of a full reference, (2008).