人　工　知　能 99−23
（1995．3．7）

数値データの論理的推論

月本 洋
(株) 東芝 研究開発センター

数値データは今まで論理的推論の対象ではなかった。本論文では数値データが直観主義論理で推論できることを示す。その概要は以下の通りである。多重線形関数空間は関数の定義域が $\{0,1\}$ の時には古典論理の代数モデルであるブール代数の拡張でありユークリッド空間となるが、定義域が [0,1] の時にもユークリッド空間になる。多重線形関数空間は、Heyting 代数になるので、直観主義論理のモデルになる。線形関数は多重線形関数空間に含まれるので、直観主義論理で論理的に推論できる。また数値データは線形関数で近似できるので、数値データも直観主義論理で論理的に推論できる。

Logical Reasoning of Numerical Data

Hiroshi Tsukimoto
Research & Development Center, Toshiba Corporation

This paper shows that numerical data can be reasoned by intuitionistic logic. The space of multi-linear functions, which is the extension of Boolean algebra, can be made into a Euclidean space. The space can be also made into a Euclidean space when the domain is [0,1]. The space is the model of intuitionistic logic. Therefore, multi-linear functions can be reasoned by intuitionistic logic. Linear functions can be also reasoned by intuitionistic logic. If numerical dara can be approximated by linear functions, the numerical data can be reasoned by intuitionistic logic.

1 はじめに

数値データは今まで論理的推論の対象ではなかった。本論文では数値データが直観主義論理で推論できることを示す。まず、多重線形関数空間（次節で説明）は関数の定義域が $\{0,1\}$ の時には古典論理の代数モデルであるブール代数の拡張でありユークリッド空間となるが、定義域が [0,1] の時にもユークリッド空間になる [4]。またブール関数の定義域を [0,1] に拡張することによって得られる連続値論理関数は古典論理の全公理をみたす [6]。従ってこの連続値論理関数も本論文ではブール関数と呼ぶことにする。多重線形関数空間はブール代数の拡張であるのだから、古典論理から何らかの公理を除去した非古典論理のモデルになっていると言える。本論文では多重線形関数空間（正確には多重線形関数のある部分集合）が、Heyting 代数であることを示すことによって,直観主義論理のモデルになっていることを示す。線形関数は多重線形関数空間に含まれるので、直観主義論理で論理的に推論できる。この推論の計算量は指数オーダーなので多項式オーダーの推論方式も提示する。数値データは線形関数で近似できるので、数値データも直観主義論理で論理的に推論できる。簡単な例で数値データの推論を示し、この推論を古典論理で近似することによってその推論の意味について述べる。

2 節では多重線形関数空間について述べる。3 節では多重線形関数空間が直観主義論理のモデルになることを述べる。4 節では非古典論理関数のブール関数近似について述べる。そして 5 節では線形関数、6 節では数値データの、論理的推論について述べる。記号に関しては次の通りである。 $f,g...$— 論理関数、 $x,y...$— 変数、 $\mathrm{f}((f_i)),\mathrm{g}((g_i))...$— 論理ベクトル

2 多重線形関数空間

定義 1 多重線形関数の定義は以下の通りである。

$\sum_{i=1}^{2^n} a_i x_1^{e_1} \cdot\cdot x_n^{e_n}$

但し a_i は実数、x_i は変数、e_i は 1 か 0、n は変数の数である。（以下、特に記述がない場合は、n は変数の数であり、$\sum$ は $\sum_{i=1}^{2^n}$ である。）

例 2 変数の多重線形関数は $axy+bx+cy+d$ である。

2.1 定義域が $\{0,1\}$ の場合

定理 2 変数の定義域を $\{0,1\}$ とする。多重線形関数空間はブール代数の原子で張られるユークリッド空間である。

証明 n 変数のブール代数の原子は以下の通りである。

$a_i=\prod_{j=1}^{n} e(x_j)\ (i=1,...,2^n)$

ここで $e(x_j)=\overline{x_j}$ または x_j である。例えば 2 変数のブール代数の原子は

$xy, x\overline{y}, \overline{x}y, \overline{xy}$

である。任意のブール関数はこの原子の線形結合で表現される。すなわち

$\sum e_i a_i$

ここで a_i は原子であり、e_i は係数（1 か 0）である。ここで係数を $\{1,0\}$ から実数に拡張すれば線形空間になる。即ちその線形関数はブール代数の原子の線形関数であり、

$\sum r_i a_i$

となる。但し r_i は実数である。これは多重線形関数である。また関数間に自然に 2 乗距離が入れられる。即ち二つの関数を

$\sum r_i a_i, \sum s_i a_i$

とすればその距離は

$\sqrt{\sum(r_i-s_i)^2}$

である。従って多重線形関数空間はユークリッド空間になる。□

これは多重線形関数空間が、内積を

$<f,g>=\sum_{\{0,1\}^n} fg$

と定義するとユークリッド空間になると言うことである。

2.2 定義域が [0,1] の場合

定義域が [0,1] の場合でも多重線形関数空間はユークリッド空間になる [4]。

定理 3 定義域が [0,1] の場合、内積 $< f, g >$ を次の様に定義すると多重線形関数空間はユークリッド空間になる。

$< f, g >= 2^n \int_0^1 \tau(fg)dx$

但し τ は $\tau(x^n) = x$ と言う写像である。

証明 [4] を参照。

定義 4 論理積、論理和、否定を以下のように定義する。

1. 論理積 $\tau(fg)$
2. 論理和 $\tau(f + g - fg)$
3. 否定 $\tau(1 - f)$

ブール関数の定義域を $\{0, 1\}$ から $[0, 1]$ へ拡張した関数は、定義 4 の論理演算で古典論理の全公理を満たす [6] ので、本論文ではブール関数と呼ぶ。以降、定義域は特に断らない限り、 $\{0, 1\}$ か [0,1] である。多重線形関数を以降論理関数とも言う。この論理関数はブール関数と非古典論理関数に分類される。論理関数のベクトル表現を論理ベクトルと言う。またブール関数のベクトル表現をブールベクトルと言う。

2.3 論理演算のベクトル表現

ブール代数の論理積、論理和、否定をベクトル表現するとそれぞれ以下のようになる。

$\mathbf{f} \wedge \mathbf{g} = (Min(f_i, g_i))$

$\mathbf{f} \vee \mathbf{g} = (Max(f_i, g_i))$

$\overline{\mathbf{f}} = (1 - f_i)$

なお定義域が [0,1] の場合の論理積、論理和、否定のベクトル表現は定義 4 より以下のようになるが、ブール関数では、即ち f_i, g_i が $\{0, 1\}$ では、上記の Min, Max での定義と同じになる。

$\mathbf{f} \wedge \mathbf{g} = (f_i g_i)$

$\mathbf{f} \vee \mathbf{g} = (f_i + g_i - f_i g_i)$

$\overline{\mathbf{f}} = (1 - f_i)$

3 直観主義論理のモデル

3.1 Heyting 代数

直観主義論理の代数モデルである Heyting 代数の公理は以下の通りである [1],[3]。

1. $f \wedge f = f, f \vee f = f$
2. $f \vee g = g \vee f, f \wedge g = g \wedge f$
3. $f \wedge (g \wedge h) = (f \wedge g) \wedge h,\ f \vee (g \vee h) = (f \vee g) \vee h$
4. $f \wedge (f \vee g) = f, f \vee (f \wedge g) = f$
5. $f \wedge (g \vee h)) = (f \wedge g) \vee (f \wedge h),\ f \vee (g \wedge h) = (f \vee g) \wedge (f \vee h)$
6. $f \wedge f' = O$
7. $f \wedge O = O, f \vee O = f,\ f \wedge I = f, f \vee I = I$

但し f' は f の O に関する相対擬補元である。O, I は最小元、最大元である。

3.2 多重線形関数空間のモデル

定義 5 $\mathbf{f} \le \mathbf{g}, \mathbf{f} \wedge \mathbf{g}, \mathbf{f} \vee \mathbf{g}$ を以下のように定義する。

$\mathbf{f} \le \mathbf{g} \equiv \forall i (f_i \le g_i)$
$\mathbf{f} \wedge \mathbf{g} \equiv (Min(f_i, g_i))$
$\mathbf{f} \vee \mathbf{g} \equiv (Max(f_i, g_i))$

上記の定義はブール代数の論理演算と同じである。(2.3参照)

定理 6 多重線形関数空間の $[0,1]^m$ の超立方体は定義 *5* により *Heyting* 代数となる。但し m は空間の次元である(以下同様)。証明は簡単なので省略する。この場合、I は 1 即ち $(1,...,1)$ となり、O は 0 即ち $(0,...,0)$ となる。また、O に関する相対擬補元は以下のようになる。

$\mathbf{f}' \equiv (f_i')$

$$f_i' = \begin{cases} 0(f_i > 0) \\ 1(f_i = 0) \end{cases}$$

定理 7 多重線形関数空間の $\prod_{i=1}^{m}[a_i, b_i]$ の超直方体は定義 *5* により *Heyting* 代数となる。証明は簡単なので省略。この場合 I は $(b_1, ..., b_m)$ で O は $(a_1, ..., a_m)$ である。また、O に関する相対擬補元は以下のようになる。

$\mathbf{f}' \equiv (f_i')$

$$f_i' = \begin{cases} a_i(f_i > a_i) \\ b_i(f_i = a_i) \end{cases}$$

従って多重線形関数空間の任意の超直方体が定義 *5* により *Heyting* 代数になる。

4 非古典論理関数のブール関数による近似

非古典論理関数をブール関数で近似することは非古典論理関数の意味を知るには有意義である [5]。

定理 8 非古典論理関数が最も近いブール関数で近似されることを考える。非古典論理関数の論理ベクトルを (f_i) とし、ブール関数の論理ベクトルを $(g_i)(g_i = 0, 1)$ とする。近似方法は以下の通りである。

$$g_i = \begin{cases} 1(f_i \ge 0.5) \\ 0(f_i < 0.5) \end{cases}$$

証明 非古典論理関数に最も近いブール関数は $\sum(f_i - g_i)^2$ を最小にする。各項は独立に最小化でき、さらにブール関数の論理ベクトルの要素は $g_i = 1$ または 0 なので、近似法は上記のようになる。□

5 線形関数の論理的推論

線形関数は多重線形関数空間に含まれる。従って線形関数を直観主義論理で推論することができる。但し定義域は $[0,1]^n$ か $\{0,1\}^n$ である。

5.1 線形関数の半順序関係

二つの線形関数 f, g を

$p_1x_1 + ... + p_nx_n + p_{n+1}$, $q_1x_1 + ... + q_nx_n + q_{n+1}$

とし、そのブール代数の原子による直交展開の係数を a_j, b_j とすると

$f \le g \equiv \forall j (a_j \le b_j)$

となる。上記の半順序関係をチェックするには指数オーダーの計算量が必要なので変数の数が多い場合には現実的でないので計算量が多項式オーダのチェック方法が必要になる。次項では計算量が多項式オーダーのチェック方法を提示する。

5.2 線形関数の半順序関係の多項式オーダーの判定方法

半順序関係の判定方法を指数オーダーから多項式オーダに落すことは、直交展開の係数 a_j, b_j で判定することから線形関数の係数 p_i, q_i で判定することへ変更することである。そこで直交展開の係数と線形関数の係数の関係をまず提示しておく。

ブール代数の原子による直交展開は以下の通りである。線形関数を

$p_1x_1 + ... + p_nx_n + p_{n+1}$

とし、それのブール代数の原子による直交展開式を

$a_1x_1 \cdot \cdot x_n + ... + a_{2^n}\overline{x_1} \cdot \cdot \overline{x_n}$

とする。

定理 9 各係数 a_i は次のようになる。

$a_1 = p_1 + ... + p_n + p_{n+1},$
$a_2 = p_1 + ... + p_{n-1} + p_{n+1},$
$...$
$a_{2^{n-1}} = p_1 + p_{n+1},$
$a_{2^{n-1}+1} = p_2 + p_3 + ... + p_{n-1} + p_n + p_{n+1},$
$...$
$a_{2^n-1} = p_n + p_{n+1},$
$a_{2^n} = p_{n+1}$

証明 a_i は

$p_1x_1 + ... + p_nx_n + p_{n+1} = a_1x_1 \cdot \cdot x_n + ... + a_{2^n}\overline{x_1} \cdot \cdot \overline{x_n}$

で a_i に対応する直交関数の値のみ 1 にし、他の直交関数の値を 0 にするような変数の値の組合わせを代入することによって簡単に求められる。□

定理 10 線形関数が定数項がない場合は

$\forall i(p_i \leq q_i) \Leftrightarrow f \leq g$

である。

証明 左辺から右辺は定理 9 から明らかである。右辺から左辺は以下の通りである。

二つの線形関数 f, g を

$p_1x_1 + ... + p_nx_n,\ q_1x_1 + ... + q_nx_n$

とする。h を $g - f$ とおくと h は

$(q_1 - p_1)x_1 + ... + (q_n - p_n)x_n$

となる。

$f \leq g \Leftrightarrow 0 \leq h$

なので定義域 $[0,1]^n$、$\{0,1\}^n$ に関わりなく、任意の $\{0,1\}^n$ で $0 \leq h$ にならねばならない。従って

$\forall i(q_i - p_i \geq 0)$

であり、

$\forall i(p_i \leq q_i)$

とならねばならない。□

定理 11 線形関数が定数項がある場合は

$(q_{n+1} - p_{n+1}) + \sum_{1 \leq j \leq n, q_j - p_j < 0}(q_j - p_j) \geq 0 \Leftrightarrow f \leq g$

である。

証明 定理 10 の証明と同様に $h = g - f$ とすると h は

$(q_1 - p_1)x_1 + ... + (q_n - p_n)x_n + (q_{n+1} - p_{n+1})$

となる。

$f \leq g \Leftrightarrow 0 \leq h$

なので定義域$[0,1]^n$、$\{0,1\}^n$に関わりなく、任意の$\{0,1\}^n$で$0 \leq h$にならねばならない。即ち$\{0,1\}^n$でのhの最小値h_{min}が$h_{min} \geq 0$でなければならない。ところでh_{min}は定数項に負の係数を全て足した値になるので

$h_{min} = (q_{n+1} - p_{n+1}) + \sum_{1 \leq j \leq n, q_j - p_j < 0}(q_j - p_j)$

である。従って$h_{min} \geq 0$は

$(q_{n+1} - p_{n+1}) + \sum_{1 \leq j \leq n, q_j - p_j < 0}(q_j - p_j) \geq 0$

となる。□

5.3 線形関数の論理的推論の例

f, gを以下の通りとする。

$f = 0.40x - 0.25y + 0.25,\ g = 0.56x - 0.30y + 0.44$

$h = g - f$は以下のようになる。

$h = 0.16x - 0.05y + 0.19$

従ってh_{min}は

$0.19 - 0.05 = 0.14$

となるので$f \leq g$即ち$f \to g$である。そこで$f \to g$なので

$$\frac{f, f \to g}{g}$$

のように線形関数を論理的に推論することが可能になる。従って、いくつかの線形関数と線形関数間の論理的関係が与えられている時に、ある線形関数から別の線形関数が（直観主義論理で）演繹できるかどうか等の応用が考えられる。

6 数値データの論理的推論

数値データは回帰分析して線形関数で近似できる。線形関数は論理的に推論できるので、数値データを論理的推論の対象にすることができる。以下に例を示す。表1のデータ(F)を考える[2]。このデータは配向度が温度と時間でどのように変わるかを示している。即ち温度と時間が説明変数で配向度が被説明変数である。そこでこのデータFは

f(温度, 時間)$\to$ 配向度

を表していると考える。ただしfは論理関数である。このままでは扱えないので各変数の下限値、上限値で$[0,1]$に正規化すると表2となる。これを回帰分析すると

$f = 0.40x - 0.25y + 0.25$

となる。直交関数展開すると

$f = 0.40xy + 0.65x\overline{y} + 0.00\overline{x}y + 0.25\overline{x}\overline{y}$

になる。また同じ金属の配向度の別のデータ(G)を表3としよう。このデータを正規化すると表4になる。これを回帰分析すると

$g = 0.56x - 0.30y + 0.44$

となる。直交関数展開すると

$g = 0.70xy + 1.00x\overline{y} + 0.14\overline{x}y + 0.44\overline{x}\overline{y}$

になる。fとgの直交展開の係数を比較すると

$0.40 \leq 0.70,\ 0.65 \leq 1.00,\ 0.00 \leq 0.14,\ 0.25 \leq 0.44$

とgの方が全て大きいため、

$f \leq g$、即ち$f \to g$

が言える。線形関数の係数では定数項があるため、この半順序関係はわからない。具体的にはyの係数が$-0.25 > -0.3$とgの方が小さい。従って

$(g \to z) \to (f \to z)$

となる。これをデータで表現すると

$G \to F$

となる。

上記の $G \to F$ の意味を理解するために古典論理で近似してみる。f, g を最も近いブール関数で近似すると、f は

$x\overline{y}$

となり、g は

$xy \vee x\overline{y} = x$

となり、

$G \to F$

は古典論理で近似すると

$(x \to z) \to (x\overline{y} \to z)$

となる。定義域が [0,1] なので、x, z は正比例を表し、$\overline{y}$ は反比例を表している。従って

$x \to z$

は「温度を上げれば配向度が上がる。」という命題を意味する。この命題は配向度は温度だけに依存し、温度以外の要因（この場合には「時間」）に関係しないということをも意味している。また

$x\overline{y} \to z$

は「温度を上げて時間を短くすれば配向度が上がる。」という命題を意味することになる。したがって

$(x \to z) \to (x\overline{y} \to z)$

は「温度を上げれば配向度が上がるのだから、温度を上げて時間を短くしたら配向度が上がる。」という命題を意味することになる。データ G は配向度は時間に関係ないということをも意味しているので、この命題は理にかなったものといえる。

そこで $G \to F$ を用いて

$$\frac{G, G \to F}{F}$$

のように数値データを論理的に推論することが可能になる。従って、いくつかの数値データと数値データ間の論理的関係が与えられている時に、ある数値データから別の数値データが（直観主義論理で）演繹できるかどうか等の応用が考えられる。但しこの $\to$ は数値データを回帰分析しているので近似的である。

表 1: ある金属の配向度（その１）

試料番号	温度 (℃)(x)	時間 (分)(y)	配向度 (%)(z)
1	1700	30	36.0
2	1800	25	41.5
3	1800	20	42.9
4	1850	30	42.0
5	1900	10	49.8
6	1930	10	51.0

表 2: ある金属の配向度（その１）（正規化後）

試料番号	温度 (℃)(x)	時間 (分)(y)	配向度 (%)(z)
1	0.00	1.00	0.00
2	0.44	0.75	0.24
3	0.44	0.50	0.30
4	0.65	1.00	0.26
5	0.87	0.00	0.60
6	1.00	0.00	0.65

表 3: ある金属の配向度（その 2）

試料番号	温度 (℃)(x)	時間 (分)(y)	配向度 (%)(z)
1	1700	30	38.3
2	1800	25	50.3
3	1800	20	46.8
4	1850	30	46.1
5	1900	10	56.7
6	1930	10	59.0

表 4: 別の金属の配向度（その 2）（正規化後）

試料番号	温度 (℃)(x)	時間 (分)(y)	配向度 (%)(z)
1	0.00	1.00	0.10
2	0.44	0.75	0.62
3	0.44	0.50	0.47
4	0.65	1.00	0.44
5	0.87	0.00	0.90
6	1.00	0.00	1.00

7 おわりに

数値データは今まで論理的推論の対象ではなかった。本論文では数値データが直観主義論理で推論できることを示した。その概要は次の通りであった。多重線形関数空間は関数の定義域が $\{0,1\}$ と [0,1] の時には古典論理の代数モデルであるブール代数の拡張でありユークリッド空間となる。多重線形関数空間は Heyting 代数になるので、直観主義論理のモデルになる。線形関数は多重線形関数空間に含まれるので、直観主義論理で論理的に推論できる。また数値データは線形関数で近似できるので、数値データも直観主義論理で論理的に推論できる。なお本論文では数値データの論理的推論の簡単な例を提示した。

参考文献

[1] D.V. Dalen: Intuitionistic logic, *Handbook of Philosophical Logic III, D. Gabbay and F.Guenthner eds.*, pp.225-339, D.Reidel, 1984.

[2] 有馬 哲, 石村 貞夫: 多変量解析のはなし, 東京図書, 1987.

[3] 小野 寛晰: 情報科学における論理, 日本評論社, 1994.

[4] 月本 洋: 命題論理の幾何的モデル, 情報処理学会論文誌, Vol.31, pp.783-791, 1990.

[5] 月本 洋: 数値データからの論理命題の発見, 人工知能学会誌, Vol.8, No.6, pp.752-759, 1993.

[6] 月本 洋: 古典論理の全ての公理を満たす連続値論理について, 電子情報通信学会論文誌, Vol.J77-D-I No.3, pp.247-252, 1994.