展開法の確率モデルによる意味空間の分析

岡 本 安 晴*

ANALYSIS OF SEMANTIC SPACE BY PROBABILISTIC UNFOLDING MODEL

Yasuharu, OKAMOTO

A new probabilistic unfolding model is developed to analyze semantic space. This model operates on rating data of adjectives, which are not paired. Distance d is defined to be r-metric and related to rating probability p by $p = \exp(-cd^b)$. An example of analysis by this method shows good fit of the model to the data, which promises further investigation and development of the proposed method.

1. 序

言葉(概念)の意味の分析法の1つに Semantic Differential (SD) 法という方法がある。この方法によると、言葉は意味空間 (Semantic Space) という情緒的 (affective) に中性 (neutral) な点を原点とし、原点から遠ざかるほど情緒的意味付けの強くなる空間に配置される (cf. Osgood, 1967)。この意味空間においては、原点に関して対象な点は「よい-悪い」のように正反対の意味をもつが、SD 法における「意味」空間で表されているものは情緒的なものである (cf. Heise, 1971; Miller et al., 1988) ことに注意されねばならない。意味空間の設定は、分析対象として用意された複数個の言葉に対して、複数個の形容詞対による評定データを採用することにより行われる。形容詞対は「よい-悪い」のように意味的 (情緒的) に正反対のものが対にされている。すなわち、それぞれの各形容詞対に対応して意味空間内の原点を通る直線がそれぞれ想定されることになる。形容詞対による評定データの因子分析に基づいて、意味空間は3次元であり、各次元は評価 (evaluation)、力量 (potency)、活動 (activity) を表しているといわれている (cf. Osgood, 1967)。

上の従来の「意味空間」の因子分析的解釈に対して、言葉(概念)と形容詞の関係を展開法的に設定す

ることも可能である。「飛行機」と「速い」のように適合性の高い言葉と形容詞の間の距離は近く、「船」と「速い」のように適合性の低いものの距離は遠いとするのである。この展開法の場合には空間内に必ずしも意味的 (情緒的) に反対のものを対にして想定する必要はない。岡本(1991)はこの展開法の考え方による SD 法データの分析例を報告している。岡本(1991)の用いた展開法は数量化4類の考え方に従うものであった。これは、SD 法データから算出される非類似度データはそのままでは距離データとして扱えないこと、non-metric 変換による距離を推定する場合には degeneration という問題がある (cf. Heiser, 1987) ことのため、データに含まれる情報の大略を把握する (cf. 林ら, 1976) という数量化法の考え方が適当であると判断されたことによる。

展開法による分析を行うのであれば、評定用の形容詞は対にして用いられるより1つ1つ単独に言葉との適合性の評価が求められる方が自然である。又、この場合には適当な反対語の設定が難しい形容詞あるいは情緒語を評定用として用いることも可能になる。概念(言葉)と個々の形容詞との適合度の評定データが収集されるのであれば、それに適したデータ分析用のモデルが想定される。そのようなモデルの1つとして次節において説明されるような確率モデルを用いることにより、nonmetric 法に伴う degeneration の問題が避けられるとともに、統計的検定が可能になる。

* 金沢大学文学部 (Faculty of Letters, Kanazawa University)

2. モ デ ル

概念(言葉) i に対する形容詞 j の適合度の評定の確率モデルを次のように設定する。

適合度の評定が、「全く当てはまらない」をカテゴリ「0」とし、「非常によく当てはまる」をカテゴリ「 $K-1$ 」とする K カテゴリにより行われるとし、概念 i と形容詞 j の組み合わせに対する評定値を C_{ij} ($C_{ij}=0, \dots, K-1$) で表す。概念 i と形容詞 j の距離が d_{ij} であるとき、 C_{ij} が k を越える確率 $P(C_{ij} > k)$ を次式で与える。

$$P(C_{ij} > k) = \exp(-c_k \cdot d_{ij}^\beta)$$

ここで、 c_k はカテゴリ k に対応して決まる値であり $0 < c_0 < c_1 < \dots < c_{K-2}$ である。カテゴリに対する確率を上のように与える方法は系列カテゴリ法(高根, 1990)と呼ばれているものと本質的に同じである。 $\beta > 0$ はカテゴリ境界の明瞭さを表すもので、 β の値が大きい程 d_{ij} と C_{ij} を関係付ける確率が 0.5 に近い領域は小さくなる。

モデルを簡単にするため、 $K=2$ の場合について考える。すなわち、データ C_{ij} が「当てはまる」「当てはまらない」の 1-0 型の場合を扱う。このとき c_0 を 1 とおくことができるので

$$P(C_{ij}=1) = P(C_{ij} > 0) = \exp(-d_{ij}^\beta) \quad (1)$$

となる。

d_{ij} の関数として $P(C_{ij}=1)$ の形は (1) 式以外のもも考えられる。

DeSarbo, Andrich, Böckenholt ら (DeSarbo et al., 1986; DeSarbo et al., 1987a, b; Andrich, 1988, 1989; Böckenholt et al., 1990, 1991) では次のような形が用いられている。

$$P(C_{ij}=1) = F(d_{ij}^\beta + b) \quad (2)$$

ここに、 $F(x)$ は正規分布あるいはロジスティック分布の累積分布関数である。上式においては、 $d_{ij}^\beta \geq 0$ であることにより $P(C_{ij})$ の値を小さくするためには b を負の値にとる必要がある。この b は threshold と呼ばれている。

$F(x)$ がロジスティック分布である場合に

$$x = \beta \log(d_{ij}^\beta)$$

とおけば次式 (3) の形になる。

$$P(C_{ij}=1) = \frac{1}{1 + d_{ij}^{2\beta}} \quad (3)$$

この形は Böckenholt ら (1990) で扱われている。

(1) 式は、単純な指数関数であり、特に対数尤度の計算に適しているが、さらに (2) 式のように $d_{ij} \geq 0$ による threshold パラメータ b の導入という不自然さ可以避免することができる。又、(3) の式の分数式に比べて d_{ij} が 0 あるいは ∞ に近づくとき $P(C_{ij}=1)$ はより速く 1 あるいは 0 に近づくという特徴がある。

距離 d_{ij} はユークリッド距離とされることが多い。しかし、例えば Shepard (1986) によれば、city-block 距離の方が適している場合もあるとされている。ここでは city-block も含めて Minkowski r -metric を採用して次のように置く。

$$d_{ij} = \left\{ \sum_{t=1}^p |x_{it} - y_{jt}|^r \right\}^{1/r}, \quad r \geq 1$$

ここに、 $\{x_{it}\}, \{y_{jt}\}; t=1, \dots, p$ は概念 i , 形容詞 j の p 次元空間における座標である。 $r=1$ のとき d_{ij} は city-block 距離となり各次元(特徴)の差異は単純に加算される。 $r=2$ のときはユークリッド距離であり各次元の差異は統合的(integral whole)に処理される。 $r=\infty$ のときは $d_{ij} = \max_t \{|x_{it} - y_{jt}|\}$ であり差異の最も大きい次元(特徴)のみが有効になり他の特徴の差異は無視される。

いま、概念 i に対して形容詞 j が当てはまる ($C_{ij}=1$) と評定された度数を n_{ij} , 当てはまらない ($C_{ij}=0$) と評定された度数を $\bar{n}_{ij}$ で表せば尤度 L は独立性の仮定のもとで次式により与えられる。

$$L = (\text{定数}) \prod_{i=1}^M \prod_{j=1}^N p_{ij}^{n_{ij}} \cdot (1 - p_{ij})^{\bar{n}_{ij}}$$

但し、 M と N は概念と形容詞の数であり $p_{ij} = P(C_{ij}=1)$ とおいた。

与えられたデータ $\{n_{ij}, \bar{n}_{ij}\}$ に対して L を最大にする $\{x_{it}\}, \{y_{jt}\}$, β を求める。モデルの適合度は次式の AIC 値によって表す。

$$\text{AIC} = -2 \cdot \max\{\log(L)\} + 2 \cdot df$$

但し、上式の計算において L の(定数)項は無視する。自由度 df は r の値により次のようになる。

$$df = \begin{cases} (M+N)p - p(p+1)/2 + 1 & r=2 \text{ のとき} \\ (M+N-1)p + 1 & r \neq 2 \text{ のとき} \end{cases}$$

これは原点の自由度に加えて、ユークリッド空間($r=$

2) の場合は回転による自由度が減ぜられるからである。βを自由パラメータとして扱うので自由度に+1される。

Lの最大化はその対数尤度 $l = \log(L)$ に対して行う。lを最大にするパラメータ値は automatic restart 付の conjugate gradient 法 (Powell, 1977) により求めるが、gradient の計算に使われる導関数は以下のようである。

$$\begin{aligned}\frac{\partial l}{\partial x_{ks}} &= \sum_j \left(\frac{n_{kj}}{p_{kj}} - \frac{\bar{n}_{kj}}{1 - p_{kj}} \right) \cdot \frac{\partial p_{kj}}{\partial x_{ks}} \\ \frac{\partial p_{kj}}{\partial x_{ks}} &= -\beta p_{kj} d_{kj}^{\beta-1} \cdot \frac{\partial d_{kj}}{\partial x_{ks}} \\ \frac{\partial d_{kj}}{\partial x_{ks}} &= \left(\frac{|x_{ks} - y_{js}|}{d_{kj}} \right)^{r-1} \text{signum}(x_{ks} - y_{js})\end{aligned}$$

ここで

$$\text{signum}(z) = \begin{cases} +1 & z > 0 \\ 0 & z = 0 \\ -1 & z < 0 \end{cases}$$

である。

$$\begin{aligned}\frac{\partial l}{\partial y_{ks}} &= \sum_i \left(\frac{n_{ik}}{p_{ik}} - \frac{\bar{n}_{ik}}{1 - p_{ik}} \right) \cdot \frac{\partial p_{ik}}{\partial y_{ks}} \\ \frac{\partial p_{ik}}{\partial y_{ks}} &= -\beta p_{ik} d_{ik}^{\beta-1} \cdot \frac{\partial d_{ik}}{\partial y_{ks}} \\ \frac{\partial d_{ik}}{\partial y_{ks}} &= -\left(\frac{|x_{is} - y_{ks}|}{d_{ik}} \right)^{r-1} \cdot \text{signum}(x_{is} - y_{ks}) \\ \frac{\partial l}{\partial \beta} &= \sum_{i,j} \left(\frac{n_{ij}}{p_{ij}} - \frac{\bar{n}_{ij}}{1 - p_{ij}} \right) \cdot \frac{\partial p_{ij}}{\partial \beta} \\ \frac{\partial p_{ij}}{\partial \beta} &= -p_{ij} d_{ij}^{\beta} \log(d_{ij})\end{aligned}$$

β ≥ 0 の制約条件があるので

$$\beta = u^2$$

とき、lの最大化においてはuの関数とみなす。このとき

$$\frac{\partial l}{\partial u} = \frac{\partial l}{\partial \beta} \cdot \frac{d\beta}{du} = 2u \frac{\partial l}{\partial \beta}$$

3. 適用法

岡本 (1991) は「実験とテスト=心理学の基礎: 解説編」に載せられている SD 法のデータを数量化4類の基準による展開法で分析した結果を報告している。そこで SD 法データは形容詞が対にして用いられているが、これを対にせず1つ1つ個別の適合度の評定

データを収集し、前章のモデルを適用してみる。用いられた形容詞は「実験とテスト」において「評価」「力量」「活動」の3つの因子に対応するものとして用いられているもの9対から得られた18個であり、評定対象(言葉)は同じく「実験とテスト」で用いられている職業名20個である(表1参照)。個々の職業名に対して18個の形容詞1つ1つの適合性の評定が「よくあてはまる」「あてはまる」「どちらかといえばあてはまる」「どちらかといえばあてはまらない」「あてはまらない」「まったくあてはまらない」の6カテゴリーにより求められた。評定は授業内での「評定法のデモンストレーション」として受講生(大学生19名)に質問紙の冊子を配布して集団で行った。質問紙は、表紙に例が示されていた。例は、「コメディアン」という職業名に対して「明るい」という形容詞の適合度の評定を示すものであり、「よくあてはまる」のカテゴリーに○印が付けられていた。1頁目からが評定用紙であり、頁ごとに左上に職業名が示され、その下に18個の形容詞が縦に並べられていた。各形容詞の右側に1から6までの数字が順にカテゴリー「よくあてはまる」から「まったくあてはまらない」に対応するものとして並べられていた。評定者は、最も適切なカテゴリーに対応する数字を丸で囲むことが求められた。形容詞はランダムに並べられているが、1つの冊子内での順序はどの頁も同じである。一頁内の形容詞の並びの順序は最後のものが先頭に置かれる逆順のものと2種類用意した。職業名の順序はどの冊子も1頁目は練習用に用意された「漁師」であるが、練習用であることは評定者には知らされなかった。「漁師」は「実験とテスト」において練習用として用いられているものである。以下の分析において「漁師」のデータは除かれている。2頁目から21頁目までの職業名の順序も、2頁目のものが最後の21頁目にくるように2頁から21頁までの提示順序を逆にしたもの2種類を用意した。冊子は、形容詞と職業名のそれぞれ2種類の順序を組み合わせて4種類が用意された。1つの冊子の評定は約20分以内に全員終了した。

表1は、「よくあてはまる」「あてはまる」「どちらかといえばあてはまる」を1つのカテゴリー「あてはまる」にまとめ、「まったくあてはまらない」「あてはまらない」「どちらかといえばあてはまらない」を1つのカテゴリー「あてはまらない」にまとめて頻度を数え直したものを示すものである。(「あてはまる」の頻度/「あてはまらない」の頻度)の形式で示されている。両

表1. 職業名に対する形容詞の適切さの評定データ. 職業名 i に対して形容詞 j が「あてはまる」と評定された頻度 n_{ij} と「あてはまらない」と評定された頻度 $\bar{n}_{ij}$ が $n_{ij}/\bar{n}_{ij}$ の形式で示されている. その下の () 内の数値は6カテゴリーによる評定の平均値である.

表1 (その1)

職業名	立派な A	役立つ B	よい C	大きい D	力のある E	強い F
僧 侶 a	16/3 (3.68)	13/6 (2.84)	16/3 (3.32)	5/14 (2.00)	7/12 (2.26)	6/13 (2.26)
ヤクザ b	0/19 (0.89)	0/19 (0.89)	0/19 (1.05)	14/5 (3.00)	16/3 (3.47)	18/1 (3.68)
音楽家 c	19/0 (3.84)	15/4 (3.32)	17/2 (3.68)	7/12 (2.37)	6/13 (2.21)	9/10 (2.47)
銀行員 d	14/5 (2.95)	18/1 (3.53)	16/3 (3.05)	1/18 (1.58)	2/16 (1.72)	1/18 (1.68)
マンガ家 e	12/7 (2.63)	13/6 (2.68)	15/4 (2.95)	3/16 (1.79)	3/16 (1.63)	2/17 (1.63)
デザイナー f	17/2 (3.26)	15/4 (2.95)	17/2 (3.11)	6/13 (2.21)	7/12 (2.21)	5/14 (1.95)
保 母 g	17/2 (3.47)	19/0 (4.11)	19/0 (3.84)	4/15 (1.84)	13/6 (2.74)	8/11 (2.21)
大学教授 h	18/1 (3.53)	15/4 (3.32)	18/1 (3.32)	5/14 (2.00)	8/11 (2.32)	8/11 (2.26)
医 師 i	19/0 (4.11)	19/0 (4.37)	18/1 (3.84)	9/10 (2.53)	16/3 (3.26)	15/4 (3.11)
警察官 j	17/2 (3.37)	19/0 (3.79)	17/2 (3.32)	12/7 (2.84)	16/3 (3.42)	16/3 (3.68)
ホステス k	3/16 (2.00)	10/9 (2.47)	4/15 (2.16)	2/17 (1.58)	2/17 (1.53)	5/14 (2.00)
新聞記者 l	17/2 (3.53)	19/0 (3.89)	16/3 (3.11)	6/13 (2.05)	11/8 (2.58)	12/7 (2.74)
八百屋 m	15/4 (3.05)	19/0 (3.68)	14/5 (3.16)	6/13 (1.95)	12/7 (3.00)	13/6 (2.89)
船のり n	17/2 (3.53)	17/2 (3.26)	16/3 (3.32)	15/4 (3.00)	19/0 (3.89)	19/0 (3.89)
政治家 o	10/9 (2.47)	11/8 (2.53)	7/12 (2.05)	9/10 (2.42)	16/3 (3.53)	13/6 (3.11)
プロスポーツ選手 p	18/1 (3.63)	15/4 (2.89)	17/2 (3.21)	18/1 (3.89)	19/0 (4.16)	19/0 (4.21)
作 家 q	18/1 (3.58)	17/2 (3.26)	19/0 (3.42)	3/16 (1.74)	7/12 (2.11)	1/18 (2.00)
俳 優 r	15/4 (2.95)	14/5 (2.89)	14/5 (2.95)	13/6 (3.00)	12/7 (2.89)	10/9 (2.53)
大 工 s	18/1 (3.68)	19/0 (4.05)	17/2 (3.47)	15/4 (3.26)	17/2 (3.95)	19/0 (3.95)
スチュワーデス t	18/1 (3.68)	19/0 (3.79)	17/1 (3.50)	4/15 (1.74)	6/13 (2.16)	4/15 (2.00)

カテゴリーの頻度の和が評定者数 19 にならない場合があるのは評定者の記入漏れ等のためである。「よくあてはまる」から「まったくあてはまらない」までのカテゴリーに 5 点から 0 点までを与え、評定の平均値を

求めたものも表 1 に示されている. これらの平均値に対して数量化 4 類の基準による展開法による解を求め、前章のモデルによる分析の初期値とした.

すなわち、概念 i の形容詞 j に対する平均評定値を

表1 (その2)

職業名	速い G	騒がしい H	若い I	ひどい J	役立たぬ K	わるい L
僧 侶	1/18	2/17	1/18	4/15	4/15	5/14
a	(1.26)	(1.00)	(1.42)	(1.74)	(1.68)	(1.47)
ヤクザ	12/7	17/2	12/7	19/0	13/6	19/0
b	(2.42)	(3.68)	(2.68)	(3.84)	(2.79)	(4.11)
音楽家	5/14	1/18	10/9	4/15	2/17	0/19
c	(1.95)	(1.42)	(2.58)	(1.53)	(1.05)	(0.95)
銀行員	16/3	4/15	16/3	4/15	0/19	2/17
d	(3.37)	(1.84)	(3.00)	(1.84)	(1.16)	(1.58)
マンガ家	7/12	9/10	17/2	7/12	5/14	3/16
e	(2.21)	(2.68)	(3.32)	(2.21)	(1.95)	(1.84)
デザイナー	10/9	10/9	14/5	5/14	4/15	4/15
f	(2.47)	(2.42)	(3.26)	(1.89)	(1.79)	(1.32)
保 母	9/10	12/7	18/1	7/12	0/19	0/19
g	(2.32)	(2.95)	(3.79)	(2.00)	(0.89)	(0.68)
大学教授	1/18	1/18	1/18	4/15	2/17	2/17
h	(1.42)	(1.47)	(1.32)	(1.74)	(1.58)	(1.53)
医 師	10/9	2/17	8/11	5/14	0/19	4/15
i	(2.58)	(1.42)	(2.26)	(2.11)	(0.74)	(1.47)
警察官	16/3	6/13	16/3	10/9	1/18	3/16
j	(3.32)	(2.16)	(3.16)	(2.74)	(1.16)	(1.68)
ホステス	5/14	15/4	16/3	14/5	6/13	10/9
k	(1.89)	(3.63)	(3.74)	(3.00)	(2.21)	(2.47)
新聞記者	19/0	16/3	17/2	11/8	1/18	2/17
l	(4.37)	(3.63)	(3.37)	(2.79)	(1.37)	(1.53)
八百屋	14/5	18/1	9/10	5/14	0/19	2/17
m	(3.26)	(3.89)	(2.47)	(1.84)	(1.05)	(1.11)
船のり	13/6	11/8	13/6	8/11	2/17	4/15
n	(3.05)	(2.79)	(3.00)	(2.47)	(1.58)	(1.58)
政治家	2/17	13/6	0/19	13/6	11/8	15/3
o	(1.53)	(3.00)	(0.84)	(3.21)	(2.63)	(3.50)
プロスポーツ選手	17/2	12/7	19/0	7/12	4/15	3/16
p	(4.16)	(2.89)	(4.26)	(2.00)	(1.84)	(1.37)
作 家	2/17	3/16	5/14	3/16	1/18	2/17
q	(1.37)	(1.32)	(1.84)	(1.74)	(1.58)	(1.37)
俳 優	8/11	13/6	13/6	9/10	6/13	5/14
r	(2.21)	(2.95)	(2.84)	(2.47)	(2.00)	(1.89)
大 工	10/9	15/4	7/12	8/11	0/19	1/18
s	(2.79)	(3.26)	(2.26)	(2.32)	(0.89)	(1.26)
スチュワーデス	14/5	3/16	19/0	7/12	1/18	0/19
t	(2.89)	(1.53)	(4.37)	(1.95)	(1.26)	(1.00)

e_{ij} とするとき、数量化の基準 Q を次のように置く。

$$Q = -\sum e_{ij}(x_i - y_j)^2$$

Q の最大化を $\sum x_i^2 = \sum y_j^2 = 1$ の条件のもとで行い、得られた解は対応する特異値の平方根でスケールングをしておく、詳しくは岡本 (1993) を参照されたい。

前章のモデルによる、次元数と r -metric の各組合わせに対する AIC 値を表 2 に示す。 $r=4$ に対する 5 次元空間における解の AIC 値が最小である。この解とデータとの適合の様子を図 1 と図 2 に示す。

図 1 は、データ値 $\hat{p}_{ij} = n_{ij}/(n_{i\cdot} + \bar{n}_{i\cdot})$ を横軸に、そのモデルによる予測値 $p_{ij} = \exp(-d_{ij}^2)$ を縦軸にとり各

表 1 (その 3)

職業名	小さな M	力のない N	弱い O	遅い P	静かな Q	老いた R
僧 侶 a	6/13 (1.95)	7/12 (2.37)	8/11 (2.11)	11/8 (2.32)	17/2 (4.00)	19/0 (3.95)
ヤクザ b	2/17 (1.32)	3/16 (1.63)	7/12 (2.05)	2/17 (1.42)	3/16 (1.42)	4/15 (2.11)
音楽家 c	3/16 (1.53)	5/14 (1.95)	4/15 (1.95)	3/16 (1.53)	11/8 (2.84)	11/8 (2.53)
銀行員 d	6/13 (2.11)	7/12 (2.37)	9/10 (2.26)	0/19 (1.11)	14/5 (2.95)	1/18 (1.63)
マンガ家 e	8/11 (2.37)	10/9 (2.53)	12/7 (2.53)	9/10 (2.63)	9/10 (2.32)	1/18 (1.47)
デザイナー f	3/16 (1.68)	7/12 (2.21)	5/14 (2.05)	4/15 (1.95)	9/10 (2.32)	2/17 (1.68)
保 母 g	11/8 (2.47)	5/14 (2.16)	7/12 (2.21)	1/18 (1.53)	2/17 (1.68)	1/18 (1.26)
大学教授 h	4/15 (1.89)	6/13 (2.26)	8/11 (2.37)	12/7 (2.63)	16/3 (3.37)	19/0 (3.95)
医 師 i	0/19 (1.32)	1/18 (1.37)	2/17 (1.58)	3/16 (1.53)	16/3 (3.16)	15/4 (2.95)
警察官 j	1/18 (1.47)	1/18 (1.32)	4/15 (1.32)	3/16 (1.37)	6/13 (2.21)	4/15 (2.05)
ホステス k	8/11 (2.37)	12/7 (3.00)	13/6 (2.84)	4/15 (1.84)	3/16 (1.58)	1/18 (1.00)
新聞記者 l	3/16 (1.74)	4/15 (1.89)	2/17 (1.63)	1/18 (0.89)	2/16 (1.56)	1/18 (1.58)
八百屋 m	6/13 (2.00)	3/16 (1.58)	3/16 (1.42)	0/19 (1.11)	0/19 (1.05)	9/10 (2.47)
船のり n	1/18 (1.37)	0/19 (1.05)	1/18 (1.21)	4/15 (1.63)	6/13 (2.32)	8/11 (2.47)
政治家 o	6/13 (2.11)	6/13 (2.05)	2/17 (1.47)	13/6 (2.89)	7/12 (2.11)	18/1 (4.16)
プロスポーツ選手 p	0/19 (1.00)	2/17 (1.32)	0/19 (1.00)	0/19 (0.89)	4/15 (1.79)	0/19 (0.74)
作 家 q	8/11 (2.00)	11/8 (2.68)	9/10 (2.37)	14/5 (3.11)	17/2 (3.68)	15/4 (3.05)
俳 優 r	4/15 (1.79)	6/13 (2.26)	8/10 (2.06)	5/14 (2.00)	7/12 (2.32)	7/12 (2.26)
大 工 s	2/17 (1.58)	1/18 (0.89)	0/19 (1.05)	1/18 (1.53)	3/16 (1.95)	14/5 (2.84)
スチュワーデス t	4/15 (2.00)	7/12 (2.32)	4/14 (1.89)	0/19 (1.16)	11/8 (2.58)	0/19 (0.63)

i, j の組み合わせについてプロットしたものである。
 d_{ij} は解 $\{x_{it}\}, \{y_{jt}\}$ に対して算出されたものである。
 図中, 2 本の折れ線はサンプル数 $N=19$ のときのデータ値 $\hat{p}_{ij}$ に対する 80% 信頼性区間を示すものである。

図 2 は, 横軸に距離の推定値 d_{ij} , 縦軸にデータ値 $\hat{p}_{ij}$ をとり, 各 i, j の組み合わせに対する点をプロットし

たものである。図中の太い曲線は確率 p を距離 d の関数 $p = \exp(-d^\beta)$ として表すものである ($\beta \approx 7.6$)。上下の細い折れ線は, 確率 p の 2 項分布においてサンプル数 $N=19$ のときの上下 10% の領域を示す。すなわち, 確率 p の 2 項分布においては, この 2 本の曲線の間にデータ値 $n_{ij}/(n_{ij} + \bar{n}_{ij})$ がある確率は 80%

表2. AIC の値

次元数	$r=2$	$r=3$	$r=4$	$r=5$	$r=6$
3	6575	6566	6592	6611	6612
4	6452	6452	6450	6453	6458
5	6422	6414	6412	6415	6418
6	6429	6443	6421	6445	6457
7	6459	6456	6468	6479	6483

分点数が 15×15 の場合の推定値

図1. データ値 $\hat{p}_{ij} = n_{ij} / (n_{ij} + \bar{n}_{ij})$ と予測値 $p_{ij} = \exp(-d_{ij}^\beta)$ の散布図. 2本の折れ線はデータ値 $\hat{p}_{ij}$ に対する80%信頼性区間 ($N=19$) を結んだものである.

である.

図1及び図2から, データに対して前章のモデルは良くあっているといえる.

図3は, $r=4$ に対する5次元空間の解 $\{x_{it}\}, \{y_{jt}\}$ を第1次元と第2次元について図示したものである. 数量化の方法に従った展開法による解を初期値としているので, 確率モデルによる解の座標軸にはその初期値の影響が残されている. このため第1次元と第2次元のように始めの方の次元に注目して解釈することができる. 図3.Aに形容詞 j の座標値 $\{y_{jt}\}$ を, 図3.Bに職業名 i の座標値 $\{x_{it}\}$ が示されている. 座標軸の単位は両図にわたって共通である. 図3.Aから第2次元が「よい-わるい」など (A, B, CとJ, K, L) の評価

個人パラメタ β の推定値

図2. データ値 $\hat{p}_{ij}$ と距離の予測値 d_{ij} の散布図. 太い曲線は距離 d と確率 $p = \exp(-d^\beta)$; $\beta \approx 7.6$ の関係を示す. 2本の細い折れ線はサンプル数 $N=19$ に対する確率 p の2項分布における上下10%の領域を示すものである. サンプル値 $n_{ij} / (n_{ij} + \bar{n}_{ij})$ が2本の折れ線の間にある確率は80%である.

を表す軸であると解釈されるが, 第1次元は「力量」(D, E, FとM, N, O)と「活動」(G, H, IとP, Q, R)の2つのグループが重なって表されている.

4. 考 察

図1及び図2に示されるデータとモデルの予測値との関係はよいといえる. 両図とも80%の確率に対応する折れ線が描かれているが, これらとの関係でデータと予測値の同時分布を考えてみればモデルはデータと「過度に」よく合っているといえる. これは, 図1においては確率80%の信頼性区間は個々の観測値 $\hat{p}_{ij}$ に対して算出され, 図2においても確率80%の領域は個々のモデル値 p に対して求められており, データとモデルの関係を全体としてとらえたときの確率ではないことによると思われる. しかし, 少なくとも図1及び図2に示されているモデルのデータへの適合性から, 2章のモデルによる(情緒的)意味空間の分析の可能性を今後さらに検討していくことの意義は十分にあ

図 3. A

図 3. B

図 3. 形容詞と職業名の解の布置. 形容詞の布置を図 3.A に, 職業名を図 3.B に示す. 形容詞と職業名の位置は表 1 に示されている英字を囲む小円により表され, 対応する形容詞と職業名が添えられている. 例えば, A は「立派な」, a は「僧侶」を表す. 座標軸の単位は両図において共通であるので, 形容詞と職業名の同時分布が, 図 3.A と図 3.B を重ねることにより得られる.

と思われる. 2 章のモデルの場合, データは言葉(概念)に対する形容詞の「適切さ」の評定が個々の形容詞について単独に求められるという形で収集される. このため, 対の構成が難しい形容詞, さらに情緒語の使用も可能となる.

モデルのデータに対する「適切さ」はその適合度の高さだけではなく, 分析結果の解釈についても検討されねばならない. これには図 3 に示されるような布置の解釈の問題が密接に関連してくるが, 3 章のデータはサンプル数が少ないことに加え, 他のデータとの比較という点で不十分である. 3 章のデータは岡本 (1991) の発表をふまえて集められているが, 今後は他の研究との比較が可能なデータの収集が望まれる.

分析結果の解釈の問題にも密接に関連しているのが距離 d_{ij} , 及び d_{ij} と確率 p_{ij} の関係の設定の仕方である. d_{ij} と p_{ij} の関係については, 岡本 (1992) の報告を参考にすれば細かい差異は無意味であると思われる. d_{ij} は, 差異の次元間の交互作用も考慮して r -metric を採用した. しかし, 情緒の心理学的な関係を表す「意味空間」における「距離」は常に r -metric でなければならないという根拠はない. さらに, d_{ij} を数学的な距離と考えることの妥当性も疑問である. むしろ, より一般的に心理的な「近さ-遠さ」を表す非類似度と考える方がよいように思われる. Takane ら (1983) では d_{ij} として Tversky (1977) の類似度のモデルに基づくものなどが扱われている. DeSarbo ら (1992) においても Tversky (1977) のモデルを基にして非類似度 d_{ij} が与えられている. 2 章のモデルの場合においても, 「意味空間」における「非類似度」 d_{ij} がどのように設定されるべきかということは今後の検討の課題の 1 つである.

参考文献

- Andrich, D. (1988) The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, **12**, 35-51.
- Andrich, D. (1989) A probabilistic IRT model for unfolding preference data. *Applied Psychological Measurement*, **13**, 193-216.
- Böckenholt, U. and Böckenholt, I. (1990) Modeling individual differences in unfolding preference data: A restricted latent class approach. *Applied Psychological Measurement*, **14**, 257-269.

- Böckenholt, U. and Böckenholt, I. (1991) Constrained latent class analysis: Simultaneous classification and scaling of discrete choice data. *Psychometrika*, **4**, 699-716.
- Desarbo, W.S. and Hoffman, D.L. (1986) Simple and weighted unfolding threshold models for the spatial representation of binary choice data. *Applied Psychological Measurement*, **10**, 247-264.
- DeSarbo, W.S. and Hoffman, D.L. (1987a) Constructing MDS joint spaces from binary choice data: A multidimensional unfolding threshold model for marketing research. *Journal of Marketing Research*, **24**, 40-54.
- DeSarbo, W.S., Johnson, M.D., Manrai, A.K., Manrai, L.A. and Edwards, E.A. (1992) TSCALE: A new multidimensional scaling procedure based on Tversky's contrast model. *Psychometrika*, **57**, 43-69.
- Desarbo, W.S., Lehmann, D.R., Holbrook, M.B., Havlena, W. and Gupta, S. (1987b) A stochastic three-way unfolding model for asymmetric binary data. *Applied Psychological Measurement*, **11**, 397-418.
- 林知己夫・鮑戸 弘 (共篇) (1976) 多次元尺度解析法—その有効性と問題点—. サイエンス社
- Heise, D.R. (1971) The semantic differential and attitude research. In G.F. Summers (Ed.), *Attitude measurement*. Rand McNally and Company.
- Heiser, W.T. (1987) Joint ordination of species and sites: The unfolding technique. In P. Legendre and L. Legendre (Eds.) *Developments in numerical ecology*. Springer-Verlag.
- Miller, G.A. and Glucksberg, S. (1988) Psycholinguistic aspects of pragmatics and semantics. In R.C. Atkinson, R.J. Herrnstein, G. Lindzey and R.D. Luce (Eds.) *Stevens' handbook of experimental psychology*, 2nd ed., vol. 2: *Learning and cognition*. John Wiley and Sons.
- 岡本安晴 (1991) 数量化 IV 類の基準による展開法. 日本行動計量学会第 19 回大会発表論文抄録集.
- 岡本安晴 (1992) 正規分布とロジスティック分布のシュミレーションによる比較. 金沢大学文学部論集・行動科学科篇, 第 12 号, 33-45.
- 岡本安晴 (1993) 数量化 4 類の基準による制約条件について. 金沢大学文学部論集・行動科学科篇, 第 13 号, 35-54.
- Osgood, C.E. (1967) Cross-cultural comparability in attitude measurement via multilingual semantic differentials. In M. Fishbein (Ed.), *Attitude theory and measurement*. John Wiley and Sons.
- Powell, M.J.D. (1977) Restart procedures for the conjugate gradient method. *Mathematical Programming*, **12**, 241-254.
- Shepard, R.N. (1986) Discrimination and generalization in identification and classification: Comment on Nosofsky. *Journal of Experimental Psychology: General*, **115**, 58-61.
- 心理学実験指導研究会 (編) (1985) 実験とテスト＝心理学の基礎: 解説編. 培風館.
- 高根芳雄 (1990) 理想点判別分析法とその応用. 柳井晴夫・岩坪秀一・石塚智一 (編) 人間行動の計量分析—多変量データ解析の理論と応用. 東京大学出版会.
- Takane, Y. and Sergeant, J. (1983) Multidimensional scaling models for reaction times and same-different judgments. *Psychometrika*, **48**, 393-423.
- Tversky, A. (1977) Features of similarity. *Psychological Review*, **84**, 327-352.

—1993 年 4 月 5 日受付,
1994 年 1 月 24 日最終修正—