計量多次元展開法の変量モデル

足立 浩平

A Random Effect Model in Metric Multidimensional Unfolding

Kohei ADACHI

A metric multidimensional unfolding procedure is proposed to represent the rows (individuals) and columns (items) of a proximity data matrix as the points in a low-dimensional space. The procedure is based on a random effect model which allows us to avoid the problem of incidental parameters. In the model, the individual points are regarded as normally-distributed random variables, while the item points are regarded as fixed parameters. The probability density of the proximity data is derived from the assumption that the true proximity is a linear function of the distance from the individual point to the item point and the observed proximity is perturbed by the normally-distributed error. The marginal likelihood obtained by integrating out the random individual points is maximized using the EM algorithm with a generalized SMACOF algorithm. The proposed procedure is evaluated with a simulation study and is applied to a preference rating data.

1. 序 論

多次元展開法(展開法)は, 多次元尺度法(MDS)を変形した解析手法である。行と列に同じ対象が並ぶ親近性データ行列を解析して, 対象の空間布置を求めるのがMDSであるが, 展開法では, 行と列に異なる対象が並ぶ親近性データ行列が適用対象となり, 行および列の各対象が単一の空間内に位置づけられる(齋藤, 1980; 高根, 1980; 岡太・今泉, 1994)。親近性データ行列を $\mathbf{X}=(x_{ij})$ で表わし, 行および列対象の点の座標値ベクトルをそれぞれ $\mathbf{f}_i$ および $\mathbf{a}_j$ と表わせば, 展開法の目的は, 親近性データ x_{ij} の値が大きいほど, 行点と列点間の距離 $d(\mathbf{f}_i, \mathbf{a}_j)$ が短くなるような点の座標値パラメータ(一連の $\mathbf{f}_i, \mathbf{a}_j$)を求めることと定式化できる。ここで, $\mathbf{X}$ を行と列に同じ対象が並ぶ正方行列として, $\mathbf{f}_i=\mathbf{a}_i$ と制約すれば, 展開法はMDSと同等

になる点で, MDSの変形というより, むしろ拡張と位置づけることもできる。

展開法を創案したCoombs(1950, 1964)をはじめ, 初期の研究(例えば, Schönemann, 1970)の多くは誤差を伴わない親近性だけを考慮するものであったが, その後の研究によって, 誤差を伴う実際データに適用できる諸方法が考案されてきた。これらの方法は, 確率モデルに基づく確率的展開法と, 確率モデルを考慮しない記述的展開法に大別できる。記述的展開法には, 最小二乗基準(DeSarbo & Carroll 1985; Greenacre & Browne 1986; Heiser, 1987)やその他の基準(Hojo, 1994; 岡本, 1995)を利用する方法が含まれ, また, 最小二乗基準に基づくMDSのアルゴリズム(Kruskal, 1964; Carroll & Chang, 1972; de Leeuw, 1977; Takane, Young & de Leeuw, 1977)が展開法も実行できるように拡張されている。一方, 確率的展開法も種々の方法が提案されており(Zinnes & Griggs, 1974; Ramsay, 1980; DeSarbo & Hoffman, 1986; De Soete, Carroll & DeSarbo, 1986; 岡本, 1994; 足立, 1995; Mackay & Zinnes, 1995), いずれの方法も最尤法によってデータ行列の行点・列点を推定するものである。なお, 記述的展開法と確率的展開

甲子園大学人間文化学部 (College of Humanities,
Koshien University)
連絡先氏名: 足立浩平
郵便宛先: 〒665-0006 兵庫県宝塚市紅葉が丘10番
1号
E-mail: VYS03471@nifty.ne.jp

法の区別は完全なものではなく、例えば、連続変数データを対象にする場合には、最小二乗基準に基づく記述的展開法の解と、距離を平均とした正規分布をデータの分布とする確率的展開法の最尤解が同じになる点で、両方法は同等である¹⁾。

さて、親近性データの中でも特に選好データが、しばしば展開法の適用対象となる（例えば、Coombs, 1964; DeSarbo & Rao, 1989）。ここで、選好データとは、個体 i が項目 j を好む程度あるいは好むか否かを表すもので、データ行列 $\mathbf{X}=(x_{ij})$ は個体×項目の行列となる²⁾。このように $\mathbf{X}$ の行に観測単位といえる個体が並ぶ場合、個体および項目に対応する行点 $\mathbf{f}_i$ と列点 $\mathbf{a}_j$ を同時推定する展開法には、統計学的に望ましくない問題が生じる。すなわち、観測回数つまり個体数が増えるほど、推定すべき個体の点（座標値ベクトル $\mathbf{f}_i$ ）も増大していく。こうしたパラメータは付随パラメータ (incidental parameter) と呼ばれ、他方、項目点（列点 $\mathbf{a}_j$ ）のように個体に付随しないものは構造パラメータ (structural parameter) と呼ばれるが、一般に付随パラメータが存在する場合、観測数を増加させても構造パラメータの推定値の偏りがなくなることが知られている (Neyman & Scott, 1948)。つまり、個体数を増やせば、推定すべき未知パラメータ $\mathbf{f}_i$ の数が増大し、さらに、 $\mathbf{a}_j$ の推定精度の改善も望めないわけである。

以上の問題を回避する方策として、付随パラメータを、定数パラメータではなく、何らかの確率分布に従って変動する潜在変数と見なし、潜在変数が積分消去された周辺尤度の最大化によってパラメータを推定するアプローチがある。こうしたアプローチは、因子分析や項目反応理論などの分野で採られ（例えば、柳井・繁樹・前川・市川, 1990; 芝, 1991; 池田, 1994）、一般に、付随パラメータを定数パラメータとするモデルを母数モデル (fixed effect model) と呼ぶのに対して、それを変数と見なすモデルを変量モデル (random effect model) と呼ぶ。

展開法においても、個体点 $\mathbf{f}_i$ を潜在変数と見なす変量モデルは幾つか提案されている。Hojo (1997, 1998) は、被験者（個体）が選好度に従って項目を順位づけたデータを対象にして、変量モデルの展開法を提案している。また、Takane (1996, 1998) は、選好度の高い項目を被験者に選択させて得られる二値選択データを対象にして、変量モデルを提案している。さらに、Andrich (1998) および Hoijtink (1991) のアプローチも、個体と項目の一次元尺度構成に限定されるが、二値選択データのための変量モデル展開法と見なせる。以上のように幾つかのアプローチがあるが³⁾、いずれも、データ x_{ij} が順位や二値といった離散変数、言い換えれば、非計量データである場合の方法である。本研究では、データ x_{ij} が離散変数ではなく、連続変数と仮定される計量データの場合を考慮して、変量モデルに基づく展開法を提案する。

提案法の骨子と解法の詳細は、次節に記す。3 節ではシミュレーションによる提案法の評価および実データへの適用例を示し、4 節では、提案法の変形や拡張に関する議論を中心として考察を行う。

2. モデルと解法

2.1 モデルと周辺最尤法

個体 $i(=1, \dots, n)$ の項目 $j(=1, \dots, m)$ に対する親近性の程度を表した計量データを x_{ij} で表し、これらをまとめて n (個体) × m (項目) の行列 $\mathbf{X}=(x_{ij})$ で表す。展開法では、前述のように、個体および項目を同一の空間内の点で表す。空間の次元数を K とした上で、項目 j に対応する点の座標値ベクトルを $\mathbf{a}_j=[a_{j1}, \dots, a_{jK}]'$ で表し、さらに計 m 個の項目の座標値をまとめて、 $m \times K$ の行列 $\mathbf{A}=[\mathbf{a}_1, \dots, \mathbf{a}_m]'$ で表す。また、個体 i に対する点の座標ベクトルを $\mathbf{f}_i=[f_{i1}, \dots, f_{iK}]'$ で表す。ここで、項目の座標行列 $\mathbf{A}$ を未知の定数パラメータと見なすのに対して、個体点の座標値ベクトル $\mathbf{f}_i$ は潜在変数と見なす。個体点と項目点との距離を、ユー

¹⁾ 一連の x_{ij} が独立に平均 $d(\mathbf{f}_i, \mathbf{a}_j)$ で同一分散の正規分布に従うという仮定のもとでの $\mathbf{f}_i, \mathbf{a}_j$ の最尤解と、 $\sum_{i,j} [x_{ij} - d(\mathbf{f}_i, \mathbf{a}_j)]^2$ を最小にする $\mathbf{f}_i, \mathbf{a}_j$ は同じである。

²⁾ 選好データがしばしば扱われる心理学では個体を被験者、項目を刺激と呼ぶが、以下の議論は、必ずしも $\mathbf{X}$ の行が被験者（人間）であることを要求せず、刺激という用語も限定的であるので、より一般性が高い個体・項目という語を用いる。

³⁾ 他に、個体点を変数と見なすアプローチとして、個体が選好判断を反復するケースを考慮し、第 r 回目の反復における個体 i の点 $\mathbf{f}_{ir}$ が、 $\mathbf{f}_i$ をパラメータとする確率分布に従うと仮定する展開法があるが (De Soete, Carroll & DeSarbo, 1986; Jedidi & DeSarbo, 1991; Mackay & Zinnes, 1995)、これらの方法は、 $\mathbf{f}_{ir}$ が変数であるものの、 $\mathbf{f}_i$ が個体 i の付随パラメータである点で、本研究の目的とは関係しない。

クリッド距離によって

$$d_j(\mathbf{f}_i, \mathbf{A}) = \|\mathbf{f}_i - \mathbf{a}_j\| = \sqrt{\sum_{k=1}^K (f_{ik} - a_{jk})^2} \quad (1)$$

と定義する. なお, 距離を前節では $d(\mathbf{f}_i, \mathbf{a}_j)$ と表記したが, 以下では記述の便宜上 $d_j(\mathbf{f}_i, \mathbf{A})$ と表す. この距離を, 次の線形回帰モデルによって, 親近性データ x_{ij} に対応づける.

$$x_{ij} = c - bd_j(\mathbf{f}_i, \mathbf{A}) + e_{ij}. \quad (2)$$

ここで, 係数 b と切片 c は未知の定数パラメータであり, e_{ij} は誤差変数を表す. なお, 説明変数にあたる $d_j(\mathbf{f}_i, \mathbf{A})$ は確率変数である.

一連の個体点の座標値ベクトル $\mathbf{f}_i (i=1, \dots, n)$ は, 互いに独立に標準多変量正規分布に従うと仮定する. すなわち, 零ベクトルを $\mathbf{0}$, K 次の単位行列を $\mathbf{I}_K$ で表すと,

$$\mathbf{f}_i \sim N_K(\mathbf{0}, \mathbf{I}_K) \quad (3)$$

と表せ, $\mathbf{f}_i$ の確率密度は,

$$p(\mathbf{f}_i) = (2\pi)^{-\frac{K}{2}} \exp\left(-\frac{1}{2} \mathbf{f}_i^T \mathbf{f}_i\right) \quad (4)$$

と書ける. 以上の仮定は, 全ての個体が単一の母集団に属し, 個体差つまり個体点のばらつきが標準多変量正規分布によって表現されることを意味する. ここで, 分布の種類を正規分布としたのは, 連続変数を扱う多く統計手法が正規分布を採用することになったためである. さらに, 分布の平均ベクトルと共分散行列をそれぞれ $\mathbf{0}$, $\mathbf{I}_K$ としたのは, 以下の理由による. まず, 平均ベクトルを $\mathbf{0}$ としても, 距離(1)が原点に関する不定性をもつため, 一般性は損なわれない (この仮定によって原点が定まる). 次に, 共分散行列を $\mathbf{I}_K$ とおく仮定は, 個体点の散布の程度と形状を規定する幾分強い制約に見えるが, 散布の程度 (空間布置の大きさ) については, (2) の係数 b によって吸収 (調整) されるため, 実際には制約されない. 個体点の散布の形状は制約を受けるが, 同様の仮定は, 因子分析 (直交解・変量モデル) の因子得点をはじめ, 他の統計手法でも利用されるものである.

一連の誤差 $e_{ij} (i=1, \dots, n; j=1, \dots, m)$ は, 互いに独立に, 平均 0, 分散 σ_j^2 の正規分布に従う, つまり

$$e_{ij} \sim N(0, \sigma_j^2) \quad (5)$$

と仮定する. ここで, σ_j^2 に付く添え字 j は誤差の分散が項目 j に固有であることを表すが, 誤差分散を項目

間で不変 ($\sigma_j^2 = \sigma^2$) とする制約も後述の適用例で考慮する. なお, 誤差分散が個体に固有, または, 個体と項目に固有とみなす (つまり, 分散は σ_i^2 または σ_{ij}^2 と表される) モデルも考えられるが, このモデルでは誤差分散が個体に伴って増える付随パラメータになってしまうので, 本研究の目的にはそぐわない. さて, 一連の $\mathbf{f}_i$ と e_{ij} の独立性を仮定すると, $\mathbf{f}_i$ と一連のパラメータ $\boldsymbol{\theta} = [\mathbf{A}, b, c, \sigma_1^2, \dots, \sigma_m^2]$ が所与のもとで, 計 m 個の項目に対する個体 i の親近性データ $\mathbf{x}_i = [x_{i1}, \dots, x_{im}]$ の確率密度は,

$$p(\mathbf{x}_i | \mathbf{f}_i, \boldsymbol{\theta}) = \left[\prod_{j=1}^m (2\pi\sigma_j^2)^{-\frac{1}{2}} \right] \exp\left\{ -\frac{1}{2} \sum_{j=1}^m \frac{[x_{ij} - c + bd_j(\mathbf{f}_i, \mathbf{A})]^2}{\sigma_j^2} \right\} \quad (6)$$

と表せる.

一連の未知パラメータ $\boldsymbol{\theta}$ の推定は, 以下に記す周辺尤度の最大化 (周辺最尤法) によって行う. まず, (6) から潜在変数 $\mathbf{f}_i$ を積分消去して得られる周辺密度

$$p(\mathbf{x}_i | \boldsymbol{\theta}) = \int p(\mathbf{x}_i | \mathbf{f}_i, \boldsymbol{\theta}) p(\mathbf{f}_i) d\mathbf{f}_i \quad (7)$$

を考慮する. 実際の計算では, 上式の積分を, 潜在変数の Q 個の離散点 $\mathbf{f}_q^* (q=1, \dots, Q)$ を用いて,

$$p(\mathbf{x}_i | \boldsymbol{\theta}) = \sum_{q=1}^Q p(\mathbf{x}_i | \mathbf{f}_q^*, \boldsymbol{\theta}) p^*(\mathbf{f}_q^*) \quad (8)$$

によって離散近似する. ここで, (4) の $\mathbf{f}_i$ を $\mathbf{f}_q^*$ に換えたものを $p(\mathbf{f}_q^*)$ と表すと, $p^*(\mathbf{f}_q^*) = p(\mathbf{f}_q^*) / \sum_{q=1}^Q p(\mathbf{f}_q^*)$ であり, $p(\mathbf{x}_i | \mathbf{f}_q^*, \boldsymbol{\theta})$ は (6) の $\mathbf{f}_i$ を $\mathbf{f}_q^*$ に換えたものである. 全データ $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_n]$ に対するパラメータ $\boldsymbol{\theta}$ の周辺尤度は, (8) の積,

$$L(\boldsymbol{\theta} | \mathbf{X}) = \prod_{i=1}^n p(\mathbf{x}_i | \boldsymbol{\theta}) = \prod_{i=1}^n \left[\sum_{q=1}^Q p(\mathbf{x}_i | \mathbf{f}_q^*, \boldsymbol{\theta}) p^*(\mathbf{f}_q^*) \right] \quad (9)$$

で与えられるが, これを最大にする $\boldsymbol{\theta}$ の推定値 $\hat{\boldsymbol{\theta}}$ を求める. この推定には次項に記す EM アルゴリズムを利用する.

推定値 $\hat{\boldsymbol{\theta}}$ が求めれば後, ベイズの定理を利用して, 事後的に潜在変数 $\mathbf{f}_i$ を算定する. まず, パラメータ $\boldsymbol{\theta}$ および個体 i のデータ $\mathbf{x}_i$ が所与のもとで, 潜在変数の離散点 $\mathbf{f}_q^* (q=1, \dots, Q)$ が個体 i の点である事後確率は

$$p(\mathbf{f}_q^* | \mathbf{x}_i, \boldsymbol{\theta}) = \frac{p(\mathbf{x}_i | \mathbf{f}_q^*, \boldsymbol{\theta}) p^*(\mathbf{f}_q^*)}{\sum_{q=1}^Q p(\mathbf{x}_i | \mathbf{f}_q^*, \boldsymbol{\theta}) p^*(\mathbf{f}_q^*)} \quad (10)$$

と表せる。この式の θ に推定値 $\hat{\theta}$ を代入して事後確率の推定値 $p(f_q^* | x_i, \hat{\theta})$ を求める。一連の $p(f_q^* | x_i, \hat{\theta})$ ($q=1, \dots, Q$) は、各 f_q^* が個体 i の点であることの言わば「確からしさ」を表すが、この $p(f_q^* | x_i, \hat{\theta})$ の分布の平均、つまり、ベイズ事後平均

$$\hat{f}_i = \sum_{q=1}^Q f_q^* p(f_q^* | x_i, \hat{\theta}) \quad (11)$$

を f_i の推定値とみなす。以上が、提案する方法の骨子である。

2.2 EM アルゴリズム

積分の離散近似に基づく周辺尤度 (9) の最大化には、EM アルゴリズム (Dempster, Laird & Rubin, 1977) が利用できる。EM アルゴリズムの基本原理は一般に知られているので、提案法における EM アルゴリズムの計算手続きだけを記す。この手続きは、一連のパラメータ $\theta = \{A, b, c, \sigma_1^2, \dots, \sigma_m^2\}$ の初期値 $\theta^{(0)}$ を定めた上で、E ステップとそれに引き続く M ステップを、収束するまで反復するものである。

まず、 t 回目の反復 ($t=0, 1, 2, \dots$) で得られたパラメータ θ の推定値を $\theta^{(t)} = \{A^{(t)}, b^{(t)}, c^{(t)}, \sigma_1^{2(t)}, \dots, \sigma_m^{2(t)}\}$ と表す。E ステップでは、データ x_i を所与として、 $\theta = \theta^{(t)}$ としたときの f_q^* の確率

$$h_{iq} = p(f_q^* | x_i, \theta^{(t)}) = \frac{p(x_i | f_q^*, \theta^{(t)}) p(f_q^*)}{\sum_{q=1}^Q p(x_i | f_q^*, \theta^{(t)}) p(f_q^*)} \quad (12)$$

を求める。次の M ステップでは、 θ と f_q^* の対数尤度の期待値

$$\begin{aligned} l = l(\theta) &= \sum_{i=1}^n \sum_{q=1}^Q h_{iq} \log p(x_i | f_q^*, \theta) p(f_q^*) \\ &= -\frac{n}{2} \sum_{j=1}^m \log \sigma_j^2 \\ &\quad - \frac{1}{2} \sum_{i=1}^n \sum_{q=1}^Q h_{iq} \sum_{j=1}^m \frac{[x_{ij} - c + b d_j(f_q^*, A)]^2}{\sigma_j^2} \\ &\quad + \text{Const} \end{aligned} \quad (13)$$

を増加させる θ を求め、これを $\theta^{(t+1)} = \{A^{(t+1)}, b^{(t+1)}, c^{(t+1)}, \sigma_1^{2(t+1)}, \dots, \sigma_m^{2(t+1)}\}$ とする。ここで、右辺では $\sum_q h_{iq} = 1$ を利用しており、 $\text{Const} = -(nm/2) \log 2\pi + \sum_i \sum_q h_{iq} \log p(f_q^*)$ はパラメータ θ に関係のない定数項である。通常の EM アルゴリズムでは M ステップで期待対数尤度 (13) を最大にする θ を求めるが、最大化までは行わずに、上記のように $l(\theta) \geq l(\theta^{(t)})$ を満たす θ を $\theta^{(t+1)}$ とする方法は、特に GEM

(一般化 EM) アルゴリズムと呼ばれる (例えば, McLachlan & Krishnan, 1997)。以上の M ステップ終了後に収束の判定を行い、収束していれば、 $\hat{\theta} = \theta^{(t+1)}$ として反復を修了し、収束していなければ、 $t := t+1$ と繰り返して E ステップに戻る。

さて、M ステップでは、(13) を増加させる $\theta = \{A, b, c, \sigma_1^2, \dots, \sigma_m^2\}$ を求めるため、次の 1), 2), 3) の手続きを順に行う。1) $b = b^{(t)}, c = c^{(t)}, \sigma_j^2 = \sigma_j^{2(t)}$ ($j=1, \dots, m$) と固定して (13) を増加させる A を求め、これを $A^{(t+1)}$ とする。2) $A = A^{(t+1)}, \sigma_j^2 = \sigma_j^{2(t)}$ ($j=1, \dots, m$) と固定して (13) を最大にする b, c を求め、これらを $b^{(t+1)}, c^{(t+1)}$ とする。3) $A = A^{(t+1)}, b = b^{(t+1)}, c = c^{(t+1)}$ と固定して (13) を最大にする σ_j^2 ($j=1, \dots, m$) を求め、これらを $\sigma_j^{2(t+1)}$ とおく。

上記の 2) で求める b, c は、連立方程式 $dl/db=0$ かつ $dl/dc=0$ を解くことによって、

$$\bar{b} = \frac{x^* d^* - w^* g}{w^* D^* - d^{*2}}, \quad \bar{c} = \frac{1}{w^*} (x^* + \bar{b} d^*) \quad (14)$$

で与えられる。ここで、

$$\begin{aligned} w_{ijq} &= \frac{h_{iq}}{\sigma_j^2}, \quad w^* = \sum_{i,j,q} w_{ijq}, \quad x^* = \sum_{i,j,q} w_{ijq} x_{ij}, \\ d^* &= \sum_{i,j,q} w_{ijq} d_j(f_q^*, A), \\ D^* &= \sum_{i,j,q} w_{ijq} d_j^2(f_q^*, A), \quad g = \sum_{i,j,q} w_{ijq} x_{ij} d_j(f_q^*, A) \end{aligned} \quad (15)$$

である。3) で求める σ_j^2 ($j=1, \dots, m$) は、 $dl/d\sigma_j^2=0$ より

$$\bar{\sigma}_j^2 = \frac{1}{n} \sum_{i=1}^n \sum_{q=1}^Q h_{iq} [x_{ij} - c + b d_j(f_q^*, A)]^2 \quad (16)$$

で与えられる。手続き 1) は解析的には達成されず、次項に記す方法を採用。

2.3 一般化 SMACOF 法

期待対数尤度 (13) の右辺で、項目の座標行列 A の関数になる項は、

$$\bar{\eta}(A) = \sum_{i=1}^n \sum_{q=1}^Q h_{iq} \sum_{j=1}^m \frac{[x_{ij} - c + b d_j(f_q^*, A)]^2}{\sigma_j^2} \quad (17)$$

であり、これを展開して整理し、 A に関係のない項を除くと、

$$\eta(A) = \sum_{j=1}^m \sum_{q=1}^Q s_{jq} d_{jq}^2(A) - 2 \sum_{j=1}^m \sum_{q=1}^Q r_{jq} d_{jq}(A) \quad (18)$$

が得られる。ここで、距離が A の関数であることを強調するために $d_{jq}(A) = d_j(f_q^*, A)$ と表し、

$$s_{jq} = \frac{b^2}{\sigma_j^2} \sum_{i=1}^n h_{iq}, r_{jq} = \frac{b}{\sigma_j^2} \sum_{i=1}^n h_{iq}(c - x_{ij}) \quad (19)$$

である。期待対数尤度(13)を増加させる $\mathbf{A}$ を求めることと、(18)を減少させる $\mathbf{A}$ を求めることは同等であるが、このために、MDSの解法として知られる SMACOF (Scalling by MAXimizing a CONVex Function) 法 (de Leeuw, 1977, 1988) を拡張した Heiser (1991) の一般化 SMACOF 法が利用できる。一般化 SMACOF 法は、適当な初期値から始めて座標行列を更新する手続きを反復する方法で、(18)のような距離の二次関数の最小化を目的とする。一般化 SMACOF 法の概説は付録にまわし、ここでは、提案法における計算手続きだけを記す。

前項の EM アルゴリズムにおける t 回目の反復で得られた $\mathbf{A}^{(t)}$ が、一般化 SMACOF 法の反復の初期値となる。つまり、一般化 SMACOF 法の反復計算は、EM アルゴリズムの反復ループの下位ループとなるが、両ループを区別するため、ここでは $\mathbf{A}^{(t)}$ を $\mathbf{A}_{[0]}$ と表す。さて、一般化 SMACOF 法では、 $u(=0, 1, 2, \dots)$ 回目の反復で得られた座標行列 $\mathbf{A}_{[u]}$ を次式によって更新して、左辺を更新行列 $\mathbf{A}_{[u+1]}$ とする。

$$\mathbf{A} = \mathbf{A}^+(\Phi \mathbf{A}_{[u]} + \mathbf{V} \mathbf{F}^*). \quad (20)$$

ここで、 $\mathbf{F}^* = [\mathbf{f}_1^*, \mathbf{f}_2^*, \dots, \mathbf{f}_Q^*]$ であり、 $\mathbf{A}^+$, Φ , $\mathbf{V}$ は、以下のように定義される行列である。まず、 $\mathbf{A}^+$ は、

$$\lambda_{jq} = \begin{cases} s_{jq} & (r_{jq} \geq 0) \\ s_{jq} + |r_{jq}|/d_{jq}(\mathbf{A}_{[u]}) & (r_{jq} < 0 \text{ かつ } d_{jq}(\mathbf{A}_{[u]}) > 0) \\ s_{jq} + r_{jq}^2/\varepsilon & (r_{jq} < 0 \text{ かつ } d_{jq}(\mathbf{A}_{[u]}) = 0) \end{cases} \quad (21)$$

の和 $\lambda_{j+} = \sum_{q=1}^Q \lambda_{jq}$ を対角要素とする対角行列 $\mathbf{A} = \text{Diag}(\lambda_{1+}, \dots, \lambda_{m+})$ のムーア・ペンローズ逆行列である。すなわち、 $\mathbf{A}^+$ も m 次の対角行列であり、その第 j 対角要素は、 $\lambda_{j+} \neq 0$ のとき λ_{j+}^{-1} 、 $\lambda_{j+} = 0$ のとき 0 となる。なお、(21)に現れる ε は、(次の段落で記す)十分に小さい正の定数である。また、 $\Phi = \text{Diag}(\phi_{1+}, \dots, \phi_{m+})$ も m 次の対角行列であり、その対角要素は、

$$\phi_{jq} = \begin{cases} r_{jq}/d_{jq}(\mathbf{A}_{[u]}) & (r_{jq} \geq 0 \text{ かつ } d_{jq}(\mathbf{A}_{[u]}) > 0) \\ 0 & (\text{その他}) \end{cases} \quad (22)$$

の和 $\phi_{j+} = \sum_{q=1}^Q \phi_{jq}$ である。 $\mathbf{V} = (\mathbf{v}_{jq})$ は、各要素が

$$\mathbf{v}_{jq} = \lambda_{jq} - \phi_{jq} \quad (23)$$

となる $m \times Q$ の行列である。

以上の更新後の行列 $\mathbf{A}_{[u+1]}$ は、 $\eta(\mathbf{A}_{[u+1]}) \leq \eta(\mathbf{A}_{[u]}) + \omega \varepsilon$ を満たす (付録参照)。ここで、 ω は $r_{jq} < 0$ かつ $d_{jq}(\mathbf{A}_{[u]}) = 0$ になる j と q の対の数を表すが、 $d_{jp}(\mathbf{A}_{[u]}) = 0$ (これに伴って $\omega > 0$) となることは稀少であり、さらに、 ε を十分に小さい正の数に設定すれば、殆どの場合に目的関数(18)の減少 $\eta(\mathbf{A}_{[u+1]}) \leq \eta(\mathbf{A}_{[u]})$ が達成されるといえる。しかし、この減少が達成されないことも考慮し、次の方法を探る。 $\eta(\mathbf{A}_{[u+1]}) \leq \eta(\mathbf{A}_{[u]})$ となる限り、(20)による座標行列の更新を複数回(計 U 回)反復し、得られた $\mathbf{A}_{[u+1]}$ を $\mathbf{A}^{(t+1)}$ として、メインループである EM アルゴリズムの反復計算に戻る。もし、(反復が U 回に達する前に) $\eta(\mathbf{A}_{[u+1]}) > \eta(\mathbf{A}_{[u]})$ となれば、そこで反復を止め、その時点での更新前の $\mathbf{A}_{[u]}$ を $\mathbf{A}^{(t+1)}$ として EM アルゴリズムに戻る。以上の方法により、(18)を減少させる $\mathbf{A}$ が得られる。なお、次節の数値例の計算では $U=10$, $\varepsilon=10^{-9}$ とする。

3. 数 値 例

3.1 シミュレーション

提案法によるパラメータの推定精度を評価するため、シミュレーションを行った。まず、項目数 m を 9、空間の次元数 K を 2 とした上で、各項目点の座標値ベクトル $\mathbf{a}_j$ の真値を基盤目状に定めた (図 1 参照)。さらに、係数および切片の真値を $b=1.5$, $c=10.0$ とし、誤差分散 σ_j^2 の真値を、1.0~9.0 の値に設定した (表 1 参照)。これらのパラメータの真値に基づいて、人工データを次の手順で発生させた。1) 潜在変数つまり個体点の座標値ベクトル $\mathbf{f}_i$ を、(3)に従う独立二変量正規乱数によって計 n 個発生させる。2) (5)に基づく正規乱数および(2)によって、データ x_{ij} を発生させる。以上 1), 2)の手続きを、個体数 n を 50, 100, 200 とし、各 n について乱数パターンを変えながら 15 回繰り返した。つまり、 $n \times m (=9)$ のデータ行列 $\mathbf{X}$ を計 3 ($n=50, 100, 200$) $\times 15$ (乱数パターン) とおり生成した。

これらの人工データ行列のそれぞれを、提案法によって $K=2$ とし解析した。解析では離散点 $\mathbf{f}_q^*$ ($q=1, \dots, Q$) の数 Q は 9^2 とした。すなわち、区間 $[-4.0, 4.0]$ を 9 等分して得られる各区間の中央の値を $z_{\alpha}(z_1 = -3.556, z_2 = -2.667, \dots, z_9 = 3.556)$ とおき、計 9^2 とおりのベクトル $[z_{\alpha}, z_{\beta}]$ ($\alpha=1, \dots, 9$; $\beta=1, \dots, 9$) を $\mathbf{f}_q^*$ ($q=1, \dots, 9^2$) とした。

図1. 真の項目点 (+) と推定された項目点 (■) (括弧内の数値は標準化された二乗誤差)。

EM アルゴリズムの反復計算で必要となる初期値は、次のように設定した。項目点の座標行列 A の初期値 $A^{(0)}$ の設定には、解析的に解ける展開法を利用するのが適当だろう。こうした展開法として、Schönemann (1970), 柳井 (1994) および岡本 (1995) の方法が挙げられるが、ここでは、計算が簡便である岡本 (1995) の展開法を利用した。すなわち、この方法で得られる項目点および個体点の座標値ベクトルをそれぞれ $\hat{a}_j (j=1, \dots, m)$ および $\hat{f}_i (i=1, \dots, n)$ と表すと、 n 個の $\hat{f}_i$ の重心 $f^{\#} = n^{-1} \sum_i \hat{f}_i$ と散布度 $\tau = (nK)^{-1} \sum_i \|\hat{f}_i - f^{\#}\|^2$ を求め、 $a_j^{(0)} = \tau^{-1/2} (\hat{a}_j - f^{\#})$ を $A^{(0)}$ の第 j ($j=1, \dots, m$) 行とした。ここで、 $\hat{a}_j$ から $f^{\#}$ を減じて $\tau^{1/2}$ で除しているのは、同じ変換を受けた個体点 $f_i^{(0)} = \tau^{-1/2} (\hat{f}_i - f^{\#})$ ($i=1, \dots, n$) の重心が 0 、要素の二乗和 $\sum_i \|f_i^{(0)}\|^2$ が nK となって、(3) 式 (f_i の各要素の平均が 0 、分散の総和が K という条件) に対応するからである。係数 b と切片 c の初期値は次の方法で求めた。上記の $f_i^{(0)}$ 、 $a_j^{(0)}$ から算出されるの個体点・項目点間のユークリッド距離 $d_{ij}^{(0)} (i=1, \dots, n; j=1, \dots, m; \text{計 } nm \text{ 個})$ を説明変数、データ x_{ij} を従属変数とした回帰分析を行い、得られた回帰係数 $\times (-1)$ および切片をそれぞれ $b^{(0)}$ 、 $c^{(0)}$ とした。誤差分散 σ_j^2 の初期値は、全ての j について等しく、データの全分散 $(nm)^{-1} \sum_{i,j} (x_{ij} - \bar{x})^2$ (ここで $\bar{x} = (nm)^{-1} \sum_{i,j} x_{ij}$) を 0.2 倍した値に設定した。以上の初期値によって反復計算を始め、 M ステップから次の M ステップまでの周辺対数尤度の増加量が 10^{-5} 以下になったとき、反復を打ち切った。

提案法の解では、(個体・項目の計 $n+m$ 点の) 布置全体の原点および大きさは一義に定まるが、回転に関する不定性が残る。そこで、真値との比較を考慮して、以上の解析で得られた各点の布置を直交プロクラステ

ス法 (Schönemann, 1966) によって回転させた。すなわち、推定された項目点の座標行列 $\hat{A}$ の回転結果 $\hat{A}T$ と真値の行列 A^* との二乗誤差 $\|\hat{A}T - A^*\|^2$ を最小にする直交行列 T を求め、この T によって、全ての項目点および個体点 $\hat{f}_i$ を回転させた⁴⁾。以下、回転後の $\hat{A}T$ 、 $\hat{f}_i T$ をそれぞれ $\hat{A}$ 、 $\hat{f}_i$ と表し、推定値として言及する。

以上の手続きより、 $3(n=50, 100, 200) \times 15$ (乱数パターン) 組の解析結果 (パラメータの推定値) が得られたが、各推定値がどの程度真値を再現しているかを検討する。まず、項目点の座標行列 A に着目し、推定値 $\hat{A} = [\hat{a}_1, \dots, \hat{a}_m]$ と真値 $A^* = [a_1^*, \dots, a_m^*]$ の二乗誤差 $\sum_j \|\hat{a}_j - a_j^*\|^2$ を、真の項目点の散布度 $\sum_j \|a_j^* - \bar{a}^*\|^2$ (ここで $\bar{a}^*$ は m 個の a_j^* の重心) で除して標準化した値を求めた。各個体数 n について 15 組の解析結果の標準化二乗誤差を小さい順に並べた結果、最小、第 4 位、中央、第 12 位、最大の値はそれぞれ、 $n=50$ のとき .033, .046, .069, .100, .370, $n=100$ のとき .015, .035, .047, .072, .151, $n=200$ のとき .009, .017, .021, .031, .063 であった (第 4, 第 12 位の値はそれぞれ 25, 75 パーセントイルにほぼ対応する)。この結果より、個体数 n が多いほど誤差が小さく、 A の再現精度が向上するといえる。図 1 には、上記の結果の中で誤差の中央値を示した結果を取り上げ、推定値と真値の布置を示すが、 n の増大に伴う再現

⁴⁾ f_i , あるいは、 f_i と A 両者の推定値を真値に適合させる回転法を採らなかったのは、変量モデルがパラメータ A の推定に主眼をおく点に配慮したからである。また、 f_i および A を別個に回転させなかったのは、同一空間内での両者の関係に基づく展開法の趣旨を考慮したためである。

表1. 係数・切片・分散の真値と誤差. 絶対値が最小, 中央, 最大の値を示した誤差を表示.

パラメータ		真値	誤差＝推定値－真値								
			$n=50$			$n=100$			$n=200$		
			最小	中央	最大	最小	中央	最大	最小	中央	最大
係数	b	1.5	0.004	0.121	0.236	－0.004	0.079	0.174	－0.003	－0.059	0.144
切片	c	10.0	－0.012	－0.123	1.383	－0.004	0.107	0.790	0.034	0.151	0.861
分散	σ_1^2	3.0	0.001	－0.284	0.705	－0.004	0.260	－1.005	－0.047	0.139	－0.569
	σ_2^2	9.0	0.004	－0.649	2.196	0.041	－0.319	－1.331	0.022	0.244	0.904
	σ_3^2	2.0	0.080	－0.391	1.215	－0.008	0.204	－0.500	－0.021	0.049	0.421
	σ_4^2	6.0	－0.030	0.511	－1.208	0.038	0.464	－1.044	－0.019	－0.133	－0.709
	σ_5^2	7.0	0.007	－0.586	－1.146	0.005	－0.112	－1.478	0.166	－0.428	0.647
	σ_6^2	1.0	0.012	0.348	－0.939	0.017	－0.107	0.603	－0.003	－0.086	0.436
	σ_7^2	4.0	－0.001	－0.489	－1.664	－0.051	0.297	0.509	－0.030	0.238	0.688
	σ_8^2	8.0	－0.038	0.340	－1.198	－0.059	－0.373	－1.536	0.012	0.203	－0.749
	σ_9^2	5.0	－0.116	0.436	0.734	0.049	－0.240	－0.779	－0.004	0.321	－0.827

図2. 生成された個体点 (○) と推定された個体点 (●) (括弧内の数値は標準化された二乗誤差).

度の向上が視察できる. ただし, 図1をみると, 個体数が最少の $n=50$ の場合でも再現が特に悪いわけではなく, 項目間のおおよその位置関係は再現されている.

次に, 係数 b , 切片 c および各項目の誤差分散 σ_j^2 に着目し, 各パラメータごとに解析結果の推定値から真値を減じた誤差 (残差) を求めた. 各 n について15組の解析結果の誤差の絶対値を順位づけ, 最小・中央の順位・最大の絶対値を示した誤差を表1に示す (順位づけは差の絶対値に基づくのに対し, 表には差を示している). 表の数箇所でも n が増大しても誤差が減少しない結果が見られるが, 総じてみれば, n の増大とともにパラメータの再現精度が向上することが窺える. さらに, 各 n につき11種のパラメータ (b, c と9個の

σ_j^2) $\times 15$ 組の解析結果, つまり, 延べ $11 \times 15 = 165$ 個の誤差の二乗和を平均した値は, .349 ($n=50$), .184 ($n=100$), .099 ($n=200$) となり, n の増加に伴うパラメータ再現度の向上を示す. なお, 表1をみると, $n=50$ でかつ最大誤差の結果でも, 推定誤差は特に著しくはないと思われる.

正規乱数によって生成された個体点 f_i と, ベイズ事後平均(11)による推定値 $\hat{f}_i$ との対応関係にも着目し, 両者の二乗誤差の総和 $\sum \| \hat{f}_i - f_i \|^2$ を, f_i の散布度 $\sum \| f_i - \bar{f} \|^2$ (ここで $\bar{f}$ は n 個の f_i の重心) で除して標準化した値を求めた. 各 n について15組の解析結果の中で, (昇順に) 最小, 第4位, 中央, 第12位, 最大の標準化二乗誤差はそれぞれ, $n=50$ のとき .266, .300, .340, .396, .469, $n=100$ のと

き.251, .291, .317, .324, .351, $n=200$ のとき.250, .265, .286, .302, .329であり, 個体数の増大に伴って誤差が減少することを示す. これらの結果の中で誤差の中央値を示した結果を取り上げ, 各 f_i と対応する $\hat{f}_i$ を線で結んだのが図2である. 図および標準化された二乗誤差を, 項目の座標行列 A の結果と比べると, f_i の再現度が劣る傾向が見られる. こうした傾向の理由として, 1) A がその関数の最適化によって推定されるのに対し, f_i はこの最適化の後に事後的に算定されること, 並びに, 2) 解の回転手続きにおいて, f_i ではなく A の推定値を真値に適合させる回転

法を採ったことが挙げられる. しかしながら, 図2を見渡すと, 例えば「布置の左の f_i に対して, 推定値 $\hat{f}_i$ が右に位置する」といった両者が大幅に隔たる例は見当たらず, こうした点では満足できる個体点の推定精度であると思われる.

3.2 選好評定データの解析例

提案法を実データに適用した. データは表2に示すとおりで, 計60名の個体(大学生の被験者)に7種類の項目(飲料水カテゴリー)の選好度を, 0から10までの値で評定させた結果である. ここで, 評定値の大

表2. 飲料水の選好評定データ(行 i は被験者番号, 列は項目: AJ=アップルジュース, BT=紅茶, CF=コーヒー, CL=コーラ, GT=日本茶, OJ=オレンジジュース, OT=ウーロン茶).

i	AJ	BT	CF	CL	GT	OJ	OT	i	AJ	BT	CF	CL	GT	OJ	OT
1	8	10	4	6	10	7	7	31	10	10	8	6	5	9	7
2	9	8	10	6	7	9	8	32	9	7	2	6	7	9	5
3	6	8	6	4	8	7	7	33	7	6	7	4	5	6	5
4	3	5	3	0	10	5	6	34	10	10	5	5	10	9	10
5	5	10	4	8	10	5	6	35	7	7	3	6	5	8	8
6	6	8	8	10	5	6	8	36	10	6	3	6	10	10	10
7	10	6	8	10	5	10	10	37	8	9	7	3	10	5	9
8	4	7	5	8	6	6	6	38	8	10	5	5	10	5	10
9	7	7	4	3	3	8	1	39	8	9	7	10	7	7	7
10	10	9	2	1	9	10	9	40	7	10	0	6	4	5	8
11	7	7	7	5	3	9	7	41	7	7	7	4	7	9	7
12	6	10	9	8	8	7	7	42	6	5	8	10	5	7	6
13	5	9	10	7	6	5	7	43	7	8	8	0	8	9	3
14	9	10	2	7	8	8	8	44	8	9	2	6	7	9	8
15	3	10	3	6	10	7	8	45	6	10	8	1	9	5	10
16	10	3	0	10	8	10	6	46	8	5	0	8	9	6	7
17	6	9	8	10	10	6	10	47	6	10	8	7	4	6	10
18	7	9	9	2	3	6	6	48	7	10	0	8	9	10	3
19	9	8	10	10	8	8	8	49	7	9	7	5	10	10	7
20	7	0	10	0	3	7	5	50	8	5	10	9	7	2	5
21	10	10	0	5	6	10	10	51	5	5	0	5	10	10	10
22	8	8	8	8	8	7	4	52	7	7	8	9	7	9	4
23	9	9	8	10	9	4	9	53	5	8	10	6	6	6	10
24	10	5	8	7	9	9	9	54	7	8	10	10	5	7	4
25	9	8	9	2	10	9	9	55	7	6	7	10	3	3	3
26	7	6	2	5	5	7	4	56	8	8	7	3	9	8	9
27	10	9	10	6	10	10	10	57	8	8	4	6	8	7	6
28	8	10	8	9	10	10	9	58	8	7	8	1	3	7	3
29	7	9	1	6	7	7	5	59	7	8	5	6	5	8	5
30	7	8	9	3	6	9	3	60	3	8	10	10	7	0	6

表3. 各種のモデルのもとでの AIC (括弧内は有効パラメータ数).

	異分散	等分散
1 次元	1911.0 (16)	1929.3 (10)
2 次元	1900.6 (22)	1930.5 (16)
3 次元	1906.9 (27)	1930.6 (21)

表4. 異分散2次元解における誤差分散の推定値 (項目名の略記は表2参照).

項目	推定値	項目	推定値
AJ	1.634	GT	2.902
BT	2.526	OJ	2.409
CF	9.320	OT	3.296
CL	6.836		

きさが選好度の高さに対応する.

次元数 K を 1, 2, 3 のそれぞれに設定して解析を行った. 離散点 $f_q^*(q=1, \dots, Q)$ の設定法は, $K=2$ のときは前項と同一で, 他のときも同様である. すなわち, $K=1$ のときは前項で定義した $z_a(a=1, \dots, 9)$ を離散点とし, $K=3$ のときはベクトル $[z_a, z_\beta, z_\gamma](a, \beta, \gamma=1, \dots, 9)$ を離散点とした(つまり $Q=9^K$ となる). また, 誤差の項目間等分散つまり $\sigma_j^2=\sigma^2(j=1, \dots, m)$ の制約条件のもとでの解析も行った. 項目間等分散の制約を導入しても, 2節で記した計算手続きが使える. すなわち, 一連の式で σ_j^2 を σ^2 に置き換え, (16)を $\sigma^2=(nm)^{-1}\sum_{i,q}h_{iq}\sum_j[x_{ij}-c+bd_j(f_q^*, A)]^2$ に換えれば, 他の手続きは同様である. 以上に記した以外の解析手続きは, 前項と同じである.

上記のように, 複数モデル(複数の次元数 K および制約の有無)のもとで解析を行う場合には, モデル間の適合度の比較が要求されるが, 適合度の指標として, Akaike (1974)の情報量規準(AIC)が利用できる. AICは, 各モデルごとに,

$$AIC = -2 \times \log(\text{最大尤度}) + 2 \times (\text{有効パラメータ数}) \quad (24)$$

によって定義され, その値が小さいモデルほどデータへの適合度が高いことを表す. ここで, 提案法では右辺の最大尤度が, 解析によって達成された最大周辺尤度 $L(\hat{\Theta}|\mathbf{X})$ となる. 誤差の項目間異分散を仮定するモデルの有効パラメータ数は, $\Theta=\{A, b, c, \sigma_1^2, \dots, \sigma_m^2\}$

図3. 異分散2次元解の項目(■)および被験者(●)の布置 (項目名の略記と被験者番号は表2参照).

に含まれるパラメータ数から, 項目点の回転の不定性に関する数を減じて, $mK+2+m-K(K-1)/2$ となる. また, 項目間等分散の制約をおくときの有効パラメータ数は, $mK+3-K(K-1)/2$ となる.

表3に, 各モデルのもとでの AIC の値を示す. まず, いずれの次元数についても, 誤差の項目間等分散の制約をおくモデルより, 制約を設けない異分散モデルの方が AIC が小さく, 高い適合度を示した. すなわち, 飲料水の種類によって評定値の誤差分散が異なるという, 項目固有の分散パラメータを導入することの意義が例証された. もちろん, 他のデータでは, パラメータ数の少ない等分散モデルの方が適合することもある. 次に, 適合度が高かった異分散モデルに限定すると, $K=2$ としたときの AIC が最小であった. つまり, 取り上げたモデルの中では異分散の2次元解が最も妥当であり, この解を採用して, その詳細を図3および表4に表示する.

図3には, 推定された項目(飲料水)点および個体(被験者)点の布置を示す. 項目点と個体点の近接関係から, 各被験者が選好する飲料水の種類が視覚的に把握でき, 項目点間の位置関係から空間布置の解釈ができる. まず, オレンジジュース(OJ)とアップルジュース(AJ), および, 日本茶(GT)とウーロン茶(OT)が近接し, ジュースおよび茶について, その詳細が異

なっても、親近性が高いことが窺える。ジュース類が布置の上部に位置し、他が下部に位置することから、布置の上下が「果汁・非果汁」の次元に対応するといえる。また、茶・ジュース類が布置の右に位置するのに対して、コーヒー (CF) とコーラ (CL) が左下に位置し、これらの間に紅茶 (BT) が位置する。この右から左下へ方向に対応する次元の命名は難しいが、飲料水に関する専門的知識に基づけば、有意義な解釈が可能かもしれない。なお、提案法のモデルには、(3) より回転の不定性があるため、次元による解釈をする場合、空間布置のどの方向に有意義な次元を見出しても構わない。

異分散 2 次元解の係数 b および切片 c の推定値は、それぞれ 1.610 および 10.234 であり、(2) より、個体・項目点間の距離と評定値との間に、評定値 $\cong 10.234 - 1.610 \times \text{距離}$ の近似関係があるといえる。表 4 には、各項目の誤差分散 σ_j^2 の推定値を示すが、コーヒー (CF) やコーラ (CL) は誤差分散が大きいに対して、ジュース類 (AJ, OJ) はそれが低いことが窺える。

4. 考 察

本研究では、変量モデルに基づく計量多次元展開法を提案した。提案法は、個体と項目との数量的な親近性データを対象にし、データが、個体・項目点間のユークリッド距離の線形関数に誤差が加わったものであると仮定する。ただし、このモデルは、項目点を定数パラメータとするのに対して、個体点を潜在変数とみなす変量モデルであり、解析では個体点を積分消去した周辺尤度を最大化する。この変量モデルの採用により、個体点が付随パラメータとなって生じる問題を回避できる。以上の提案法によって満足できる精度でパラメータを推定できることが、シミュレーションを通して例証された。選好評定データの解析で例示されたように、提案法で得られる少数次元の空間布置から、項目間関係および個体と項目との相互関係が容易に把握できる。

提案法を、以下に記すように、変形または拡張することができる。提案法のモデル (2) は、 $b \neq 0$ として、正規誤差変数 e_{ij} の正負を反転すると、

$$y_{ij} = d_j(\mathbf{f}_i, \mathbf{A}) + e_{ij} \quad (25)$$

と書き換えられる。ここで、 $y_{ij} = b^{-1}(c - x_{ij})$ であり、 $b > 0$ と仮定すれば、 y_{ij} は個体 i と項目 j の非親近性を

表す。すなわち、(25) と (5) より、提案法は、非親近性 y_{ij} の分布として、ユークリッド距離 $d_j(\mathbf{f}_i, \mathbf{A})$ を平均とした正規分布を採用していると位置づけられる。さて、こうした非親近性の分布として、Ramsay (1977, 1982) は、正規分布ではなく、対数正規分布を仮定することを提唱している。この仮定に従えば、 y_{ij} の対数が、 $d_j(\mathbf{f}_i, \mathbf{A})$ の対数を平均とした正規分布に従うことになる、

$$\log y_{ij} = \log d_j(\mathbf{f}_i, \mathbf{A}) + e_{ij} \quad (26)$$

と書くことができる。この仮定のもとで、 $\mathbf{f}_i$ を潜在変数とみなす変量モデルを考えることもできる。以上の対数正規分布に基づく方法は、 $y_{ij} > 0$ が要求されるなどの点で、提案法に比べて幾分複雑なアルゴリズムを要すると思われるが、興味深いアプローチの一つである。

また、提案法では各項目に固有の誤差分散を導入したが、分散と同様に切片を各項目に固有のものと仮定することもできる。つまり、(2) において c に添え字 j をつけ、

$$x_{ij} = c_j - b d_j(\mathbf{f}_i, \mathbf{A}) + e_{ij} \quad (27)$$

と拡張するわけである。(3)、(5) 等の分布の仮定をそのままにして、上式の距離 $d_j(\mathbf{f}_i, \mathbf{A}) = \|\mathbf{f}_i - \mathbf{a}_j\|$ を内積 $\mathbf{f}_i^T \mathbf{a}_j$ に置き換えると、因子分析 (直交因子) の変量モデルが得られる点で (柳井他, 1990)、モデル (27) は因子分析と展開法の比較といった理論的観点からは興味深いモデルである。

付 録

一般化 SMACOF 法の原理の概説を兼ね、座標行列 $\mathbf{A}$ の更新式 (20) の導出法を以下に記す。まず、項目と離散点の対 (j, q) を要素とする集合を、 r_{jq} と $d_{jq}(\mathbf{A}_{[u]})$ の値に基づいて、3 つの部分集合 $\mathcal{Q}_1 = \{(j, q); r_{jq} \geq 0\}$ 、 $\mathcal{Q}_2 = \{(j, q); r_{jq} < 0, d_{jq}(\mathbf{A}_{[u]}) > 0\}$ 、 $\mathcal{Q}_3 = \{(j, q); r_{jq} < 0, d_{jq}(\mathbf{A}_{[u]}) = 0\}$ に分けると、目的関数 (18) は、

$$\begin{aligned} \eta(\mathbf{A}) = & \sum_{(j,q)} s_{jq} d_{jq}^2(\mathbf{A}) - 2 \sum_{(j,q) \in \mathcal{Q}_1} r_{jq} d_{jq}(\mathbf{A}) \\ & + 2 \sum_{(j,q) \in \mathcal{Q}_2} |r_{jq}| d_{jq}(\mathbf{A}) + 2 \sum_{(j,q) \in \mathcal{Q}_3} |r_{jq}| d_{jq}(\mathbf{A}) \end{aligned} \quad (A1)$$

と書き換えられる。次に、(A1) の右辺の第 2, 3, 4 項を置き換えた関数

$$\mu(\mathbf{A}, \mathbf{A}_{[u]}) = \sum_{(j,q)} s_{jq} d_{jq}^2(\mathbf{A})$$

$$\begin{aligned}
& -2 \sum_{(j,q) \in \Omega_1} \phi_{jq} (\mathbf{f}_q^* - \mathbf{a}_j) (\mathbf{f}_q^* - \mathbf{a}_j^{[u]}) \\
& + \sum_{(j,q) \in \Omega_2} |\gamma_{jq}| \left[\frac{d_{jq}^2(\mathbf{A})}{d_{jq}(\mathbf{A}_{[u]})} + d_{jq}(\mathbf{A}_{[u]}) \right] \\
& + \omega \varepsilon + \sum_{(j,q) \in \Omega_3} \frac{\gamma_{jq}^2 d_{jq}^2(\mathbf{A})}{\varepsilon} \quad (\text{A2})
\end{aligned}$$

を導入する。ここで、 $\mathbf{a}_j^{[u]}$ は $\mathbf{A}_{[u]}$ の第 j 行ベクトルを表し、 ϕ_{jq} は (22) で定義される。また、 ε は任意の小さい正の数で、 ω は Ω_3 の要素数である。(A2) を最小にする $\mathbf{A}$ を $\mathbf{A}_{[u+1]}$ として、 $\mathbf{A}_{[u+1]}$ および $\mathbf{A}_{[u]}$ を (A1)、(A2) の $\mathbf{A}$ に代入すると、

$$\begin{aligned}
\eta(\mathbf{A}_{[u+1]}) & \leq \mu(\mathbf{A}_{[u+1]}, \mathbf{A}_{[u]}) \leq \mu(\mathbf{A}_{[u]}, \mathbf{A}_{[u]}) \\
& = \eta(\mathbf{A}_{[u]}) + \omega \varepsilon, \quad (\text{A3})
\end{aligned}$$

つまり、 $\eta(\mathbf{A}_{[u+1]}) \leq \eta(\mathbf{A}_{[u]}) + \omega \varepsilon$ となることが、Heiser (1991) に詳述される議論から導かれる。さて、(21)、(23) で定義される λ_{jq} , ν_{jq} を用いて、(A2) の右辺を整理し、 $\mathbf{A}$ に関係のない項を $Const^*$ にまとめると、(A2) は次のように書き換えられる。

$$\begin{aligned}
\mu(\mathbf{A}, \mathbf{A}_{[u]}) & = \sum_{(j,q)} \lambda_{jq} \|\mathbf{a}_j\|^2 \\
& - 2 \sum_{(j,q)} \phi_{jq} \mathbf{a}_j' \mathbf{a}_j^{[u]} - 2 \sum_{(j,q)} \nu_{jq} \mathbf{a}_j' \mathbf{f}_q^* + Const^* \\
& = \text{tr} \mathbf{A}' \mathbf{A} \mathbf{A} - 2 \text{tr} \mathbf{A}' \boldsymbol{\Phi} \mathbf{A}_{[u]} \\
& - 2 \text{tr} \mathbf{A}' \mathbf{V} \mathbf{F}^* + Const^*. \quad (\text{A4})
\end{aligned}$$

$\mu(\mathbf{A}, \mathbf{A}_{[u]})$ を最小にする $\mathbf{A}$ つまり $\mathbf{A}_{[u+1]}$ は、(A4) を $\mathbf{A}$ で偏微分したものを $\mathbf{0}$ とおいて得られるので、更新式 (20) が与えられる。

参 考 文 献

- 足立浩平 (1995). 混同行列の展開法モデル. 心理学研究, **66**, 58-62.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716-723.
- Andrich, D. (1988). The application of an unfolding model of the PIRT type to the measurement of attitude. *Applied Psychological Measurement*, **12**, 33-51.
- Carroll, J.D., & Chang, J.J. (1970). Analysis of individual differences in multidimensional scaling via an N -way generalization of 'Eckart-Young' decomposition. *Psychometrika*, **35**, 282-319.
- Coombs, C.H. (1950). Psychological scaling without a unit of measurement. *Psychological Review*, **57**, 145-158.
- Coombs, C.H. (1964). *A theory of data*. New

- York: Wiley.
- de Leeuw, J. (1977). Application of convex analysis to multidimensional scaling. In J.R. Barra, F. Brodeau, G. Rominer, & B. Van Cussem (Eds.), *Recent developments in statistics*. Amsterdam: North-Holland. Pp. 133-145.
- de Leeuw, J. (1988). Convergence of the majorization method for multidimensional scaling. *Journal of Classification*, **5**, 163-180.
- Dempster, A.P., Laird, N.M., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society: B*, **39**, 1-38.
- DeSarbo, W.S., & Carroll, J.D. (1985). Three-way metric unfolding via alternating weighted least squares. *Psychometrika*, **50**, 275-300.
- DeSarbo, W.S., & 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., & Rao, V.R. (1989). GENFOLD2: A general unfolding methodology for the analysis of preference/dominance data. G. De Soete, H. Feger., & K.C. Klauer (eds.). *New developments in psychological choice modeling*. Amsterdam: North-Holland, Pp. 57-75.
- De Soete, G., Carroll, J.D., & DeSarbo, W.S. (1986). The wandering ideal point model: A probabilistic multidimensional unfolding model for paired comparisons data. *Journal of Mathematical Psychology*, **30**, 28-41.
- Greenacre, M.J. & Browne, M.W. (1986). An efficient alternating least-squares algorithm to perform multidimensional unfolding. *Psychometrika*, **51**, 241-250.
- Heiser, W.J. (1987). Joint ordination of species and sites: The unfolding technique. In P. Legendre and L. Legendre (Eds.). *Developments in numerical ecology*. New York: Springer-Verlag, Pp 189-221.
- Heiser, W.J. (1991). A generalized majorization method for least squares multidimensional scaling of pseudodistances that may be negative. *Psychometrika*, **56**, 7-27.
- Hojitink, H. (1991). A latent trait model for dichotomous choice data. *Psychometrika*, **55**, 641-656.
- Hojjo, H. (1994). A new method for multidimensional unfolding. *Behaviormetrika*, **21**, 131-147.
- Hojjo, H. (1997). A marginalization model for the multidimensional unfolding analysis of ranking data. *Japanese Psychological Research*, **39**, 33-42.
- Hojjo, H. (1998). Multidimensional unfolding analyses of ranking data for groups. *Japanese*

- Psychological Research*, **40**, 166-171.
- Jedidi, K., & DeSarbo, W.S. (1991). A stochastic multidimensional scaling procedure for the spatial representation of three-mode, three-way pick any/*J* data. *Psychometrika*, **56**, 471-494.
- 池田 央 (1994). 現代テスト理論. 朝倉書店.
- Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29**, 1-29.
- Mackay, D.B., & Zinnes, J.L. (1995). Probabilistic multidimensional unfolding: An anisotropic model for preference ratio judgements. *Journal of Mathematical Psychology*, **39**, 99-111.
- McLachlan, G.J., & Krishnan, T. (1997). *The EM algorithm and extensions*. New York: Wiley.
- Neyman, J., & Scott, E.L. (1948). Consistent estimates based on partially consistent observations. *Econometrika*, **16**, 1-32.
- 岡太彬訓・今泉 忠 (1994). パソコン多次元尺度構成法. 共立出版.
- 岡本安晴 (1994). 展開法の確率モデルによる意味空間の分析. 行動計量学, **21**, 66-74.
- 岡本安晴 (1995). 数量化4類の基準による展開法. 行動計量学, **22**, 126-134.
- Ramsay, J.O. (1977). Maximum likelihood estimation in multidimensional scaling. *Psychometrika*, **42**, 241-266.
- Ramsay, J.O. (1980). The joint analysis of direct ratings, pairwise preferences, and dissimilarities. *Psychometrika*, **45**, 149-165.
- Ramsay, J.O. (1982). Some statistical approaches to multidimensional scaling data (with discussion). *Journal of the Royal Statistical Society: A*, **145**, 285-312.
- 齋藤堯幸 (1980). 多次元尺度構成法. 朝倉書店.
- Schönemann, P.H. (1966). A generalized solution of the orthogonal Procrustes problem. *Psychometrika*, **31**, 1-10.
- Schönemann, P.H. (1970). On metric multidimensional unfolding. *Psychometrika*, **35**, 349-356.
- 芝 祐順 (1991). 項目反応理論—基礎と応用—. 東京大学出版会.
- 高根芳雄 (1980). 多次元尺度法. 東京大学出版会.
- Takane, Y. (1996). An item response model for multidimensional analysis of multiple-choice data. *Behaviormetrika*, **23**, 153-167.
- Takane, Y. (1998). Choice model analysis of the "pick any/*n*" data. *Japanese Psychological Research*, **40**, 31-39.
- Takane, Y., Young, F.W., & de Leeuw, J. (1977). Nonmetric individual differences multidimensional scaling: An alternating least squares method with optimal scaling features. *Psychometrika*, **42**, 7-67.
- 柳井晴夫 (1994). 多変量データ解析法—理論と応用—. 朝倉書店.
- 柳井晴夫・繁樹算男・前川眞一・市川雅教 (1990). 因子分析—その理論と方法—. 朝倉書店.
- Zinnes, J.L., & Griggs, R.A. (1974). Probabilistic multidimensional unfolding analysis. *Psychometrika*, **39**, 327-350.

— 1998年10月14日 受付,

1999年11月9日 最終修正 —