

# ノンメトリック・アンフォールディング (非計量的展開法) へのハイブリッド・アプローチ

中 西 正 雄

## I はじめに

ノンメトリック・アンフォールディング (非計量的展開法) は1960年代から70年代にかけてポピュラーなマーケティング・リサーチ技法であった (Green and Carmone, 1969)。しかしその人気は過去20年間に下火になった。不人気の一つの理由は、おそらくこの技法の解につきまとう曖昧さによるものと思われる。この論文ではわれわれは曖昧さの一つの原因、すなわち順位 (序数) データが本質的にもっている不確定性にとくに焦点を当てる。

まずノンメトリック・アンフォールディング (以下では NMU と略称) の定義から始めよう。NMU に通常使われるデータは、 $n$  人のジャッジから得られた順位の  $n \times m$  行列である。それぞれのジャッジは、各自の選好にもとづいて、 $m$  個の対象物 (製品、ブランド、店舗など) に 1 から  $m$  までの順位をつける。(順位をつける基準は選好とはかぎらない。) 一例を示すと、ジャッジが 3 名、対象が 3 個の場合、データとなる順位行列は下記のようになる。

ジャッジ (行ポイント)	対象(列ポイント)		
	A	B	C
1	2	1	3
2	1	2	3
3	2	1	3

一般的には、ジャッジは行ポイント、評価対象物は列ポイントと呼ばれる。

NMU の目的は、このデータ行列から、行ポイントと列ポイントの座標を共通の空間中表示することである。すでにこの問題を解くために多数のアルゴリズムが開発され、汎用統計パッケージにも採用されている（たとえば SPSS の PREFSCAL）。ここでは、多くのアルゴリズムは最適解を求める際の目的関数として、下記の「ストレス」測度（もしくはその変形）を用いていることを指摘するにとどめたい（de Leeuw, 2005）。

$$(1) \quad \sigma(X, Y) = \sum_{i=1}^n \sum_{j=1}^m w_{ij} (\delta_{ij} - d_{ij}(X, Y))^2$$

この式で：

$\delta_{ij}$  = ジャッジ  $i$  の対象  $j$  に対する非選好 (disliking) の指標（現在の文脈では順位）

$d_{ij}(X, Y)$  = 共通空間におけるジャッジ  $i$  と対象  $j$  間の距離

$w_{ij}$  = ウェイト（必要ならば）

$X$  = 行ポイント（ジャッジ）の座標（ $n \times p$  行列）

$Y$  = 列ポイント（対象）の座標（ $m \times p$  行列）

および

$$(2) \quad d_{ij} = \sqrt{\sum_{s=1}^p (x_{is} - y_{js})^2}.$$

このように定義すると、NMU の目的は「観察された順位データ行列  $\{\delta_{ij}\}$  から  $X$  と  $Y$  を計算することである」と言い換えることができよう。

しかし、NMU の本質的な問題の一つは「たとえ  $Y$  があたえられても、 $X$  の座標が有限の範囲に納まるとは限らない」ことである。 $Y$  を所与とした場合、ある順位に関する解の値域は楔形をしていて、その頂点は原点にあるが、反対側は無限に広がっていることがある。この場合、 $X$  の布置になにか恣意的な制約を加えない限り、 $X$  の座標をこの値域の中に納めることは不可能である。このようなケースは稀ではない。例えば、4つの列ポイントが共通平面上に凸4辺形を形作っている場合、列ポイント間の24個（=4!）

の可能な順位のうち、すくなくとも 8 個の解の値域は無限となる。こうした有限解のない順位を含む順位データについて数値解析的な手法で  $X$  と  $Y$  の布置を求めようとすると、最終的には「退化解 (degenerate solution)」と呼ばれる無意味な解に収束してしまう。De Leeuw (2006) は「ノンメトリック・アンフォールディングの現在の定式化は、本質的に不適切に提起された問題 (ill-posed problem) であり、別のアプローチが要求されている」と述べているが、この「退化解」問題も NMU の本質的な欠陥の一つであろう。

この小論では、NMU の退化解問題の部分的解決策として、一つのハイブリッド・アプローチを提案する。このアプローチでは、二つの選好モデル—理想点モデルとベクトル・モデル—の混合を使って上記の本質的な不確定性に対応しようとする。基本的なアイデアは「NMU の標準である理想点モデルを、有限解が得られない可能性のある状況下で、部分的にベクトル・モデルで置き換えよう」というものである。この両方のモデルを計算するには、一対比較法にもとづく最尤法を用いることにする。

## II モデルと計算手続き

### モデル

この節ではハイブリッド・アプローチの基本的な論理とそれを順位データに適合させる方法について述べる。まず理想点モデルとベクトル・モデルの定義から始めよう。 $\{\delta_{ij}\}$  を順位集合 (データ行列) とし、 $X$  と  $Y$  を計算すべき座標としよう。また  $X_i$  と  $Y_j$  とをそれぞれ  $X$  の  $i$  番目の行、 $Y$  の  $j$  番目の行とする。以下では行ポイントを「個人」、列ポイントを「対象」と呼ぶが、これは説明を簡明にするためである。

ここで  $a_{ij}$  を対象  $j$  の個人  $i$  に対する「魅力」と定義する。そして「個人  $i$  から得られた  $\{\delta_{ij}\} (j=1, 2, \dots, m)$  は  $\{a_{ij}\} (j=1, 2, \dots, m)$  に関して単調である (すなわち、すべての  $k \neq j$  について、もし  $a_{ij} \geq a_{ik}$  ならば  $\delta_{ij} \geq \delta_{ik}$ )」と仮定する。この仮定のもとに、二つの選好モデルは以下のように規定される。

(3) 理想点 (Ideal Point) Model:  $a_{ij} = \exp(d_{ij}^{-2})$

(ここで  $d_{ij}$  は(2)式で規定。)

(4) ベクトル (Vector) Model:  $a_{ij} = \exp(2X_i Y_j')$ ,  $X_i X_i' \leq 1$

(3)式と(4)式とに共通の記号  $X_i$  を使うことには数学的には問題はないが、この二つの式の間で  $X_i$  の意味に若干の違いがある。(3)式では  $X_i$  は個人  $i$  の共通空間における座標を示すが、(4)式では共通空間の各軸に付与されるウェイトである。(4)式のウェイトには  $X_i X_i' \leq 1$  という制約が付けられているが、われわれに必要なのは選好ベクトルの「方向」であってその「長さ」ではないから、この制約は一般性を損なうものではない。これがないとベクトル・モデルは理想点モデルと同じような退化解に悩まされることになる。

すでに上に述べたように、われわれの主なアイデアは「特定の個人について理想点モデルが  $X_i$  に関して有限の解を持たないときにベクトル・モデルで置き換えよう」というものである。この考え方が単に恣意的なものではないことは以下の理由による。

- (1) ベクトル・モデルはある個人の理想点が対象 (列ポイント) の布置から無限に遠いところにあること意味している。したがって  $X_i$  の解の値域が有限でない場合に理想点モデルをベクトル・モデルで置き換えることは、解の値域に人為的な制約を加えることに比べれば、恣意的でないと言える。
- (2) (1)式と(2)式を比較すると、ベクトル・モデルが理想点モデルに入れ子 (“nested”) になっていることが分かる。 $X_i Y_j'$  の項は  $d_{ij}^2 = (X_i - Y_j)(X_i - Y_j)'$  の一部だからである。 $a_{ij}$  を、 $\exp(d_{ij}^{-1})$  でなく、 $\exp(d_{ij}^{-2})$  と規定したり、(4)に定数 2 を加えたりしたのは、(4)が(3)に入れ子になっている点を明確にするためである。

### 最尤法による推定

次にこのハイブリッド・アプローチの解を求める方法について述べる。片平 (1990, 1991) は NMU の解を求めるのに最尤法を用いることを提案した。

ある個人がつけた対象の順位から得られる尤度関数は次のように書くことができる。

$$\ln(L_i) = \sum_{j=1}^{m-1} \ln(a_{ij} / \sum_{k=j+1}^m a_{ik})$$

上式で  $a_{ij}$  は対象  $j$  の個人  $i$  に対する魅力度である。この定式化で、片平は個人  $i$  の順位  $\{\delta_{i1}, \delta_{i2}, \dots, \delta_{im}\}$  は魅力度  $\{a_{i1}, a_{i2}, \dots, a_{im}\}$  に関して単調であると仮定し、この尤度関数を用いて  $X$  と  $Y$  とを求めることができることを示した。

われわれは片平の尤度関数と同等であるが、ヨリ単純な尤度関数を用いることにした。まずある個人の対象順位  $\{\delta_{i1}, \delta_{i2}, \dots, \delta_{im}\}$  は一連の対比較に展開できることを利用する。たとえば対象 A、B、C、D に付けられた順位  $\{1, 2, 3, 4\}$  は次の対比較行列に展開できる。

$$\begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{cccc} A & B & C & D \\ \left[ \begin{array}{cccc} 1/2 & 0 & 0 & 0 \\ 1 & 1/2 & 0 & 0 \\ 1 & 1 & 1/2 & 0 \\ 1 & 1 & 1 & 1/2 \end{array} \right] \end{array}$$

上の展開行列—“explosion” matrix と呼ばれる—は「列上位である」という。すなわち列に示された対象が行に示された対象より選好されたら 1、そうでなければ 0 が入る。また対角要素は同じ対象同志の比較なので 1/2 が入っている。

この展開行列にもとづく尤度関数は以下の式であたえられる。

$$(5) \quad \ln(L_i) = \sum_{k=1}^m \sum_{j=1}^m c_{ijk} \ln(\pi_{ijk})$$

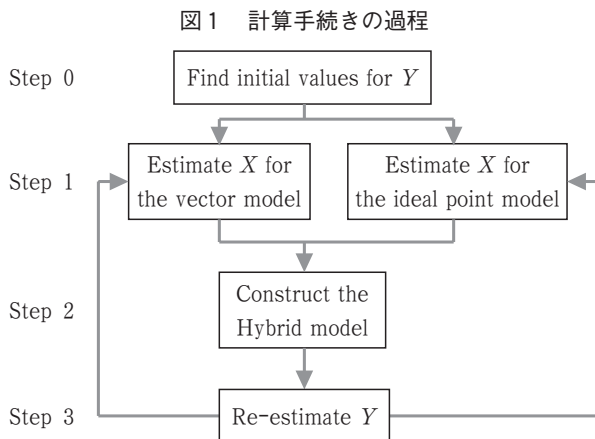
$$\text{ここで } \pi_{ijk} = \frac{a_{ij}}{a_{ij} + a_{ik}},$$

$c_{ijk}$  = 展開行列の要素 (もし個人  $i$  が対象  $j$  を対象  $k$  より選好すれば 1、さもなければ 0)

である。全体的な尤度関数は個人に理想点モデルが使われたか、ベクトル・モデルが使われたかにかかわらず、(5)式を  $n$  人の個人について合計したものである。

### 計算手続き

$X$  と  $Y$  との計算手続きは以下のようなものである。(図1参照)



ステップ0：適当な  $Y$  の初期値を定める。

ステップ1：個人  $i$  の順位  $\{\delta_{i1}, \delta_{i2}, \dots, \delta_{im}\}$  ( $i=1, 2, \dots, n$ ) について、(5)式により理想点モデルとベクトル・モデルそれぞれの最尤解  $X_i$  を求める。

ステップ2：すべての個人について理想点モデルの解が得られたら、その中で  $X_i$  の値域が無限となりそうなケースを探し、ベクトル・モデルと置き換えてハイブリッド・モデルを構築する。

ステップ3：上で得られたハイブリッド・モデルの  $\{X_i\}$  ( $i=1, 2, \dots, n$ ) を所与として、(5)式により  $Y$  を再計算する。

解の値が収束するまでステップ1からステップ3を繰り返す。この計算手続

きの結果として得られる  $X$  の最終解は理想点モデルの布置とベクトル・モデルのウェイトとの混合になっている。

この最尤化手続きは「交互最尤化」推定法と呼ぶことができる。まず  $Y$  の初期値を定め、それを所与として  $X$  を推定し、またその推定結果を使って  $Y$  を推定するという手続きを繰り返すからである。この繰り返しは(5)式の対数尤度関数が収束するまで行われる。これまでの経験では、収束は必ずしも早くないが、一旦理想点モデルとベクトル・モデルの適切な混合が見つかると、確実に収束に向かう。

数値例に進む前に、いくつかの主要な課題を取り上げておく必要がある。第一は、「 $Y$  の初期値をどう選ぶか」という課題である。われわれは、順位データの行列  $\{\delta_{ij}\}$  をあたかも数値データであるように扱って、これを特異値分解 ( $UAV^{-1}$ ) し、 $VA$  を  $Y$  の初期値とした。この方法は順位データに関わるタブーをすべて無視していると非難されるかもしれないが、対象 (すなわち列ポイント) 間の類似性についてわれわれが利用できるデータは、順位行列  $\{\delta_{ij}\}$  の列間距離 (もしくは列間相関) しかないことを考えれば、この方法はあながち誤りとは言えないであろう。 $X$  の初期値には理想点モデルでは原点を、ベクトル・モデルでは特異値分解の左ベクトル ( $U$ ) を用いた。

第二の問題は、「 $X$  と  $Y$  とにどんな人為的な制約が必要か」についてである。すでにベクトル・モデルの  $X'_i$  について  $X_i X'_i \leq 1$  という制約が必要であることは述べた。われわれはもう一つ  $Y$  について  $\sum Y_j Y'_j = m$  ( $m$  は対象の数) という制約を加える。これは列ポイントの原点からの自乗距離の平均値を 1 とする制約に等しい。この制約は対象 (列ポイント) の布置が無限に拡大する (または原点に収束する) ことを防ぐためのもので、とくに厳しい制約ではない。なお理想点モデルの  $X$  についてなんらかの制約を加えることは、不必要であるだけでなく、ハイブリッド・アプローチ開発の趣旨に反する。

第 3 の問題は、「計算手続きのステップ 2 でどの個人の理想点モデルをベ

クトル・モデルと置き換えるか」の判断に関するものである。われわれの計算手続きを繰り返す中で、理想点が対象（列ポイント）布置の中心から外へ外へと移動する個人は、最終的に退化解となる可能性が強い。しかし、「中心から外へ移動する理想点を見つけよ」という指示を計算手続きに加えることは難しい。そこで以下では「理想点の原点からの距離が2以上になったとき、その個人にはベクトル・モデルを適用する」ことにした。すなわち  $\sqrt{X_i X_i}$  が2以上になったとき、「その個人の理想点は退化解となる」と判断するのである。もちろんこの判断基準は恣意的であるが、経験的には、もしベクトル・モデルに置き換えなければ、このような個人の理想点はほとんどの場合に無限遠となることが分っている。

なおこの判断基準を適用した場合、次のステップ3で  $Y$  の値が変わると、またベクトル・モデルから理想点モデルに切り替わる個人が出てくる。このため理想点モデルとベクトル・モデルの  $X$  の値はすべての個人について計算を続けねばならない。すなわちハイブリッド・モデルは一度選べばそれで終わりなのではなく、ステップ1から3を繰り返すごとに再規定される。この手続きがもたらす問題については後に述べる。

### III 数 値 例

#### 数値例1

ハイブリッド・モデルの論理と計算手続きをチェックするために、個人（行ポイント）10人が5個の対象（列ポイント）に順位をあたえたデータを用いた。この順位は無作為に生成したものである。表1(A)は原データで、表1(B)は左表のそれぞれのセルから行平均（すなわち3）を引いたものである。計算手順は以下のとおりである。

ステップ0：表1(B)の行列を特異値分解して、 $Y$ の初期値を得た。元来は4次元の特異値解が得られるが、ここでは2次元のみを採用した。（表2参照）

ステップ1：対象（列ポイント）の初期布置（調整済み）を所与として、



表1 小さな数値例

1(A) Rankings

Individuals	Products				
	A	B	C	D	E
1	3	1	5	2	4
2	3	2	5	4	1
3	5	1	3	2	4
4	2	3	1	5	4
5	5	1	2	3	4
6	5	1	4	3	2
7	4	1	2	5	3
8	1	5	2	3	4
9	5	1	2	4	3
10	4	5	1	3	2

1(B) Row Centered Rankings

Individuals	Products				
	A	B	C	D	E
1	0	-2	2	-1	1
2	0	-1	2	1	-2
3	2	-2	0	-1	1
4	-1	0	-2	2	1
5	2	-2	-1	0	1
6	2	-2	1	0	-1
7	1	-2	-1	2	0
8	-2	2	-1	0	1
9	2	-2	-1	1	0
10	1	2	-2	0	-1

表2 Yの初期値

特異値 ( $\Lambda$ )

次元	特異値	貢献(%)	累積(%)
1	7.06	36.91	36.91
2	5.07	26.53	63.44
3	3.65	19.08	82.52
4	3.34	17.48	100

列ポイント (U $\Lambda$ )

初期値 (調整済み)

列ポイント	Dim 1	Dim 2	Dim 3	Dim 4	SUMSQ	Dim 1	Dim 2	SUMSQ
A	4.11	1.30	-0.32	-2.08	23.00	1.13	0.19	1.31
B	-5.57	-0.06	-0.44	-1.35	33.00	-1.40	0.08	1.98
C	1.38	-4.32	-0.31	0.55	21.00	0.29	-1.03	1.15
D	0.19	2.07	-1.95	1.97	12.00	0.09	0.74	0.55
E	-0.11	1.02	3.02	0.91	11.00	-0.10	0.03	0.01

10人の個人について理想点モデルとベクトル・モデルの  $X_i$  を推定した。結果は表3に示す。

ステップ2：表3で、すでに個人2、3、5、10（黄色でハイライト）以外の理想点は原点からの距離が2以上になっている。そこでこの6人については以降ベクトル・モデルを適用した。

ステップ3：ステップ2の結果を踏まえて、Yの布置を再推定した。そ

表3 繰り返し1の計算結果

Ideal Point Model				Vector Model				Hybrid Model (Step 3)			
	Dim 1	Dim 2	Row SSQ		Dim 1	Dim 2	Row SSQ		Dim 1	Dim 2	Row SSQ
A	1.130	0.186	1.15	A	1.130	0.186	1.15	A	0.952	0.314	1.00
B	-1.405	0.077	1.41	B	-1.405	0.077	1.41	B	-1.473	0.047	1.47
C	0.286	-1.031	1.07	C	0.286	-1.031	1.07	C	0.095	-1.196	1.20
D	0.088	0.738	0.74	D	0.088	0.738	0.74	D	-0.035	0.556	0.56
E	-0.099	0.031	0.10	E	-0.099	0.031	0.10	E	0.122	0.242	0.27
1	-1.094	1.760	2.07	1	-0.479	0.878	1.00	1	-0.479	0.878	1.00
2	-0.483	0.306	0.57	2	-0.298	0.391	0.49	2	-0.483	0.306	0.57
3	-1.394	-0.049	1.40	3	-0.996	0.089	1.00	3	-1.394	-0.049	1.40
4	0.267	-2.694	2.71	4	0.207	-0.978	1.00	4	0.207	-0.978	1.00
5	-1.537	-0.962	1.81	5	-0.892	-0.451	1.00	5	-1.537	-0.962	1.81
6	-25.600	1.679	25.66	6	-0.985	0.175	1.00	6	-0.985	0.175	1.00
7	-1.868	-2.474	3.10	7	-0.646	-0.763	1.00	7	-0.646	-0.763	1.00
8	39.219	2.082	39.27	8	0.985	-0.175	1.00	8	0.985	-0.175	1.00
9	-25.600	-10.570	27.70	9	-0.821	-0.572	1.00	9	-0.821	-0.572	1.00
10	0.220	-1.003	1.03	10	0.308	-0.794	0.85	10	0.220	-1.003	1.03

初期値のみ log-likelihood = -49.8360    Step3 log-likelihood = -28.0942

の値は上の図の左端にある。またこのときの対数尤度(5)の値は -28.09 であった。

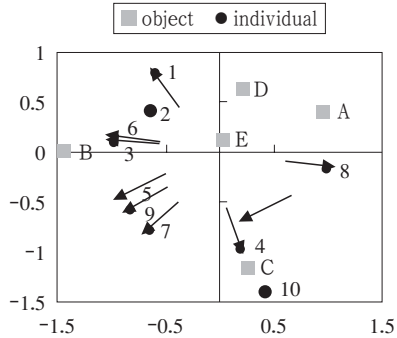
上のステップ1から3を(5)式が収束するまで繰り返した結果が表4ならび

表4 ハイブリッド・アプローチの最終解

	Dim 1	Dim 2	Row SSQ
A	0.950	0.399	1.062
B	-1.439	0.010	2.072
C	0.261	-1.160	1.415
D	0.206	0.627	0.435
E	0.025	0.125	0.016
1	-0.597	0.803	0.679
2	-0.653	0.429	0.610
3	-0.996	0.089	1.000
4	0.207	-0.978	1.000
5	-0.892	-0.451	1.000
6	-0.985	0.175	1.000
7	-0.646	-0.763	1.000
8	0.985	-0.175	1.000
9	-0.821	-0.572	1.000
10	0.413	-1.397	2.122

log-likelihood = -27.388

図2 小さな数値例の最終解



に図2である。(5)式の値が収束するまでに、11回の繰り返しが必要であった。最終的には個人2と10(黄色で表示)を除いて他の個人にはすべてベクトル・モデルが適用された。図2では各個人のベクトル( $X_i$ )は矢印で示されている。この最終解の対数尤度は-27.39であり、初期値を入れただけの段階における-49.84から大幅に改善されている。またこの結果から再構築された順位と原データ(表1(A))の順位との間の順位相関を計算したところ、10人の平均で0.87であった。

比較のために同じデータ(表1(A))をSPSSのPREFSCALによって分析した結果が表5と図3である。(Yについてハイブリッド・アプローチと同じ制約を加えているので、Xについても調整が施されている。)

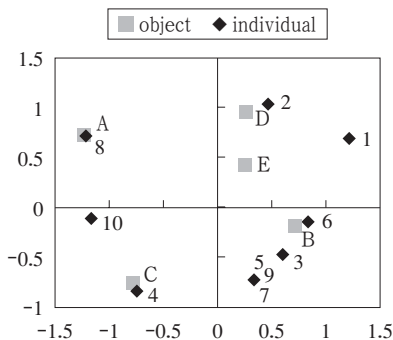
PREFSCALから得られた布置をわれわれのモデルに代入して対数尤度-41.66を得た。しかしPREFSCALはストレス((1)式)を最小化の基準に用いているので、われわれの結果と直接には比較できない。そこでPREFSCALの結果から各個人の順位行列を再構築して、原データとの順位相関を計算したところ、その平均値は0.80であった。これから見ると、われわれのハイブリッド・アプローチの方が適合度がよいと言えるだろう。対象

表5 PREFSCALの解

	Dim 1	Dim 2	Row SSQ
A	-1.243	0.719	2.063
B	0.703	-0.179	0.526
C	-0.799	-0.741	1.188
D	0.247	0.956	0.975
E	0.241	0.435	0.247
1	1.212	0.685	1.939
2	0.455	1.028	1.264
3	0.606	-0.469	0.587
4	-0.766	-0.804	1.234
5	0.322	-0.702	0.596
6	0.824	-0.125	0.695
7	0.322	-0.702	0.596
8	-1.257	0.712	2.087
9	0.322	-0.702	0.596
10	-1.190	-0.112	1.428

log-likelihood=-41.655

図3 PREFSCALの解



布置に関しては、PREFSCAL の対象布置を左右反転すれば、ほぼハイブリッド・アプローチの布置に合致する。個人の布置では、個人2と10以外の個人ベクトルはほぼ PREFSCAL 解と（中心から）同じ方向を向いている。個人2と10の理想点の位置が PREFSCAL とハイブリッド・アプローチとの間で大きく異なったのはなぜか分かっていない。

## 数値例 2

次の例は SPSS の Sample Files から取ったもので、15種類のベーカリー製品について、42名のジャッジ（男性21名、女性21名）が朝食の際の選好順序を示したものである（breakfast\_overall.sav. 付録A）。

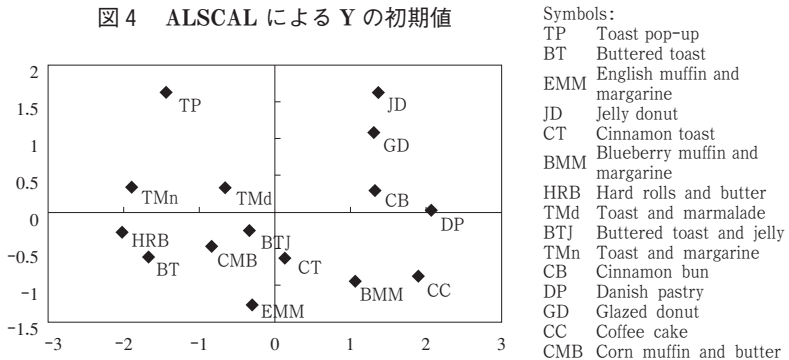
ステップ0：この順位データをあたかも数値データのように仮定し、ALSCAL を用いて  $Y$  の初期値を決定した。（図4）

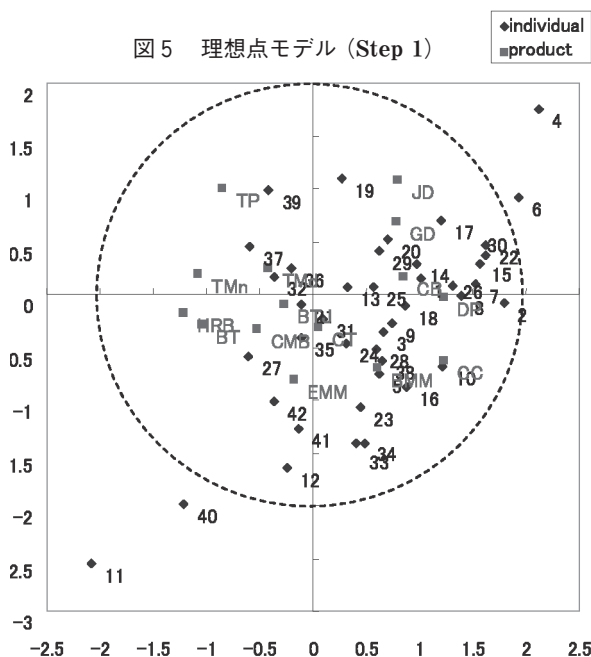
ステップ1： $Y$  の初期値を所与とし、42人の個人（＝ジャッジ）の布置を理想点モデルにより計算した結果が図5である。点線の円は  $Y_i Y_j'$  の平均値の2倍の距離を示している。

ステップ2：すでに個人4、6、11、40は円の範囲を超えて、退化解になる兆しを見せているので、以降の計算ではこれらの個人にはベクトル・モデルを適用した。

ステップ3：ステップ2で選んだハイブリッド・モデルを受けて、 $Y$  の

図4 ALSCAL による  $Y$  の初期値





布置を再推定した。

ステップ1から3を対数尤度が最大化するまで繰り返し、得られた結果を図6 (A)に示した (各対象の座標は付録Bに)。この最終解では、個人4、6、7、10、11、15、22、および30にベクトル・モデルが適用された (図では矢印でベクトルを示している)。その他の個人は理想点で表されている。対数尤度の最大値は $-1911.73$ で、各個人ごとに復元した順位と元の順位との順位相関係数の平均値は $0.753$ であった。

図6 (B)は同じデータを PREFSCAL で分析した結果である。二つの結果を比較可能にするために、 $Y$ にはハイブリッド・アプローチと同じ制約を課し、それにともない $X$ も調整している。この布置をわれわれの理想点モデルに入れた場合の対数尤度の最大値は $-2039.68$ で、個人ごとの順位相関係数の平均値は $0.745$ であった。ハイブリッド・アプローチの方が僅かに適

図 6 (A) ハイブリッド・アプローチの最終解

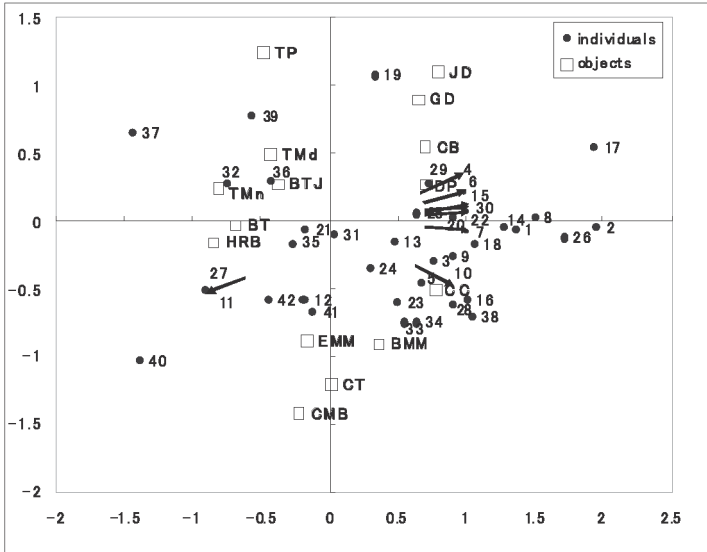
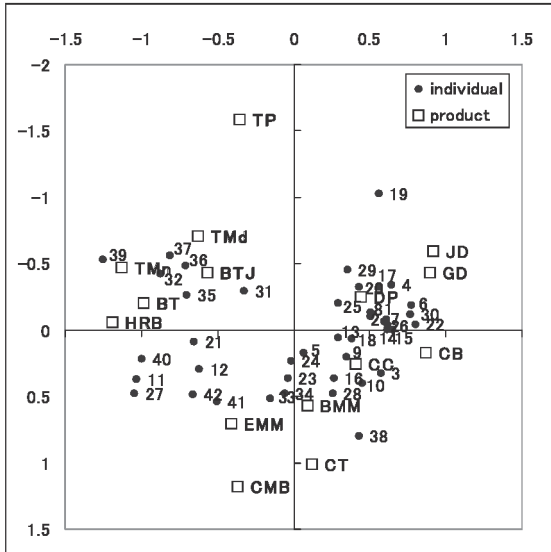


図 6 (B) PREFSCAL の解 (縦軸は上下反転)



合度が高かった。

ハイブリッド・アプローチの最終解と PREFSCAL 解との座標を比較すると、明らかに対象（列ポイント）の座標は両方の解で似たようなパターンを示している。（この点は図 6 (B) の布置を上下逆にしてみるとよく分かる。この結果は両方の方法で似たような Y の初期値を使っているからかもしれない。）対象（ベーカリー製品）はおおよそ次の三つのグループに分けることができる

グループ 1 : JD, GD, CB, and DP

グループ 2 : CMB, CT, EMM, BMM and CC

グループ 3 : HRB, BT, BTJ, TMn, TMd, and TP

これらのグループは、それぞれ「ドーナツとペーストリー」、「マフィンと丸パン」、「トーストとロール」と名付けることができよう。この対象のグルーピングは納得できるものである。

われわれはさらにこれらの布置が各個人の最も好む対象を予測するのにどれだけ役立つかを調べた。このためにまず各個人を元データの 1 位の対象がどれかによって、上の 3 グループの一つに割り当てた。（すなわち 1 位が JD であればグループ 1 に割り当てた。）次にハイブリッド・アプローチの最終解から再構築された 1 位の対象がどのグループに属するかを調べた。同じように PREFSCAL 解からも 1 位の対象がどれであるかを調べた。

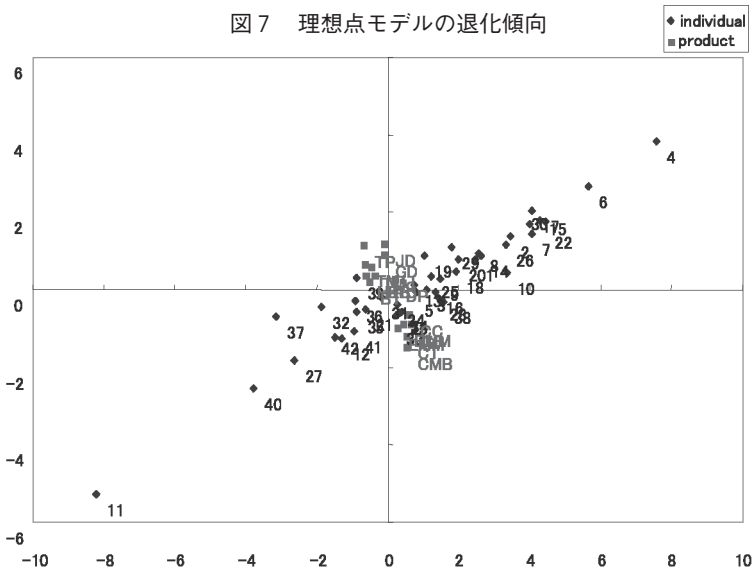
表 6 製品グループへの分類比較

6 (A) ハイブリッド・アプローチ 6 (B) PREFSCAL

元データによる分類 <sup>a)</sup>	再構築順位による分類			Total	再構築順位による分類			Total
	1	2	3		1	2	3	
グループ 1	10	8	1	19	14	4	1	19
グループ 2	2	9	1	12	2	9	1	12
グループ 3		4	7	11		3	8	11
Total	10	21	9	42	16	16	10	42

<sup>a)</sup>各個人は第 1 位に選んだ対象により分類されている。

上の表は元データとハイブリッド・アプローチおよび PREFSCAL の解と



の合致度をクロス表分析したものである。正解のパーセンテージ（対角セルの和を個人数で割ったもの）はハイブリッド・アプローチが64.3%、PREFS CAL が73.8%であった。これで見ると、PREFS CAL の方が予測精度が高いと言える。

数値例2の結果を要約すると、ハイブリッド・アプローチの最終解とPREFS CAL 解とを比較して、それほど大きな差はないことが判明した。対象のグルーピングや消費者の選好について、どちらからもほぼ同じような結論を導くことができる。とすればなぜわざわざハイブリッド・アプローチを使う必要があるのか、という疑問が出てくるかもしれない。しかしその理由は「同じデータに理想点モデルのみを適用するとどんな解が得られるか」を見ることで明らかになる。図7は理想点モデルの繰り返し計算を50回行った時点での結果を示す。

50回の繰り返しでもまだ収束はしていないが、この点ですでに退化解の兆候が顕著に出ている。行ポイント（個人）はすべて一直線上に並び、列ポイ



ントはそれと直角の直線上に並びつつある。これはよくある円周上に並ぶ退化解とは異なる形をしているが、この繰り返し計算を続けると退化解となることは疑いない。すなわち PREFSCAL 解は、理想点モデルの布置になんらかの制約を加えることで得られたものであることが分る。ハイブリッド・アプローチの制約 ( $\sum Y_j Y'_j \leq m$ 、およびベクトル・モデルについて  $X_i X'_i \leq 1$ ) のような最低限度の制約の元に、図 6 (A) の解が得られたことはこのアプローチの一つの利点であると言える。さらにハイブリッド・アプローチは布置の計算に対数尤度を最大化する方法を取っている。この小論ではその効用を十分に示すことはできなかったが、ハイブリッド・アプローチの様々な解に関して統計的な推論を行うことが考えられる。これは今後の研究課題としたい。

#### IV 結論およびディスカッション

この小論ではノンメトリック・アンフォールディングに関するハイブリッド・アプローチとその解法（最尤法）を提案した。このアプローチの論理は「理想点モデルが適用できない個人についてはベクトル・モデルを適用する」という簡単なものである。われわれは、この方法は理想点モデルまたはベクトル・モデルのどちらかだけを使ってノンメトリック・アンフォールディングを行おうとするより、合理的であると考え。計算手続きとしての最尤法は簡明であり、二つのモデルを無理なく包括することができる。

ハイブリッド・アプローチの適合度は、数値例 1 では PREFSCAL より良く、数値例 2 ではほとんど変わらなかった。一般論として、共通空間の次元が小さい（2 または 3）のに列ポイントの数が多い場合、ベクトル・モデルはうまく機能しないようである。これは解の値域が無限になる順位が相対的に少なくなるからである。こうした場合には理想点モデルの方がうまく当てはまる順位が相対的に増え、ベクトル・モデルが適用される個人が相対的に少なくなる。そうなれば適合度の見地からハイブリッド・アプローチの相対的優位性も低くなるであろう。しかしすでに述べたように、われわれのハイブリッド・アプローチは、データとの適合手続きとしてストレス ((1)式) を

最小化する方法でなく、最尤法を用いている。これにより、統計的な推論の可能性があることは前述した。たとえば共通空間の次元を2から3に増やしたことによる適合度の差を統計的に検定できるかもしれない。今後もこのような可能性を追求する必要がある。

最後にハイブリッド・アプローチにもいくつかの困難な問題が残っていることも指摘しておきたい。まず最も困難な問題は、 $Y$ の初期値の設定である。今回の二つの数値例では、最終解の $Y$ の布置は初期値から大きく異なっていない。つまり、ハイブリッド・アプローチの解は $Y$ の初期値に大きく依存している可能性がある。この問題に対処するには $Y$ の初期値をランダムに選ぶような工夫が必要であるかもしれない。その際、ランダムな初期値を複数選んで複数の解を比較することも考えられよう。こうした初期値の設定についての検討は今後の研究課題の一つである。

またこの小論で用いた交互最尤法の欠点は、収束が遅いことである。前述したように、いったんベクトル・モデルを適用すべき個人が識別されたら、計算手続きは遅くとも単調に収束する。しかし一つの繰り返しにおいてある個人の $X_i X'_i$ が2を超えたとしても、次の（またはその後の）繰り返しにおいてこの個人の $X_i X'_i$ が2以下になることがある。さらに悪いことには、今回の数値例において特定の個人が理想点モデルとベクトル・モデルとの間を行ったり、来たりする現象が見られた。つまりどの個人にベクトル・モデルを適用すべきか、なかなか明確にならないために収束が遅くなるのである。この問題に対し、 $X_i X'_i$ の閾値を2から大幅に（たとえば3に）上げて、その代わりに一旦この閾値を超えた個人にはすべてベクトル・モデルを適用することが考えられる。この方策の是非についても今後の検討課題としたい。

(筆者は関西学院大学名誉教授)

#### <参考文献>

- de Leeuw, Jan (2005), "Multidimensional Unfolding," an entry in Everitt, Brian S., and David C. Howell (eds.), *The Encyclopedia of Statistics in Behavioral Science*, Wiley.
- de Leeuw, Jan (2006), "On Degenerate Nonmetric Unfolding Solutions," Department of Statis-

tics Papers, Department of Statistics, UCLA, UC Los Angeles.

Green, Paul E., and Frank J. Carmone (1969), "Multidimensional Scaling: An Introduction and Comparison of Nonmetric Unfolding Techniques," *Journal of Marketing Research*, 6 (August 1969), 330-41.

Katahira, Hotaka (1990), "Perceptual Mapping Using Ordered Logit Analysis," *Marketing Science*, 9, No. 1 (Winter 1990), 1-17.

片平秀貴 (1991)、『新しい消費者分析』、東京大学出版会。

## 付録A 数値例2のデータ

id	gende	TP	BT	EMM	JD	CT	BMM	HRB	TMd	BTJ	TMn	CB	DP	GD	CC	CMB
1	1	13	12	7	3	5	4	8	11	10	15	2	1	6	9	14
2	2	15	11	6	3	10	5	14	8	9	12	7	1	4	2	13
3	1	15	10	12	14	3	2	9	8	7	11	1	6	4	5	13
4	2	6	14	11	3	7	8	12	10	9	15	4	1	2	5	13
5	1	15	9	6	14	13	2	12	8	7	10	11	1	4	3	5
6	2	9	11	14	4	7	6	15	10	8	12	5	2	3	1	13
7	1	9	14	5	6	8	4	13	11	12	15	7	2	1	3	10
8	2	15	10	12	6	9	2	13	8	7	11	3	1	5	4	14
9	1	15	12	2	4	5	8	10	11	3	13	7	9	6	1	14
10	2	15	13	10	7	6	4	9	12	11	14	5	2	8	1	3
11	1	9	2	4	15	8	5	1	10	6	7	11	13	14	12	3
12	2	11	1	2	15	12	3	4	8	7	14	10	9	13	5	6
13	1	12	1	14	4	5	6	11	13	2	15	10	3	9	8	7
14	2	13	11	14	5	4	12	10	8	7	15	3	2	6	1	9
15	1	12	11	8	1	4	7	14	10	9	13	5	2	6	3	15
16	2	15	12	4	14	5	3	11	9	7	13	6	8	1	2	10
17	1	7	10	8	3	13	6	15	12	11	9	5	1	4	2	14
18	2	7	12	6	4	10	1	15	9	8	13	5	3	14	2	11
19	1	2	9	8	5	15	12	7	10	6	11	1	3	4	13	14
20	2	10	11	15	6	9	4	14	2	13	12	8	1	3	7	5
21	1	12	1	2	10	3	15	5	6	4	13	7	11	8	9	14
22	2	14	12	10	1	11	5	15	8	7	13	2	6	4	3	9
23	1	14	6	1	13	2	5	15	8	4	12	7	10	9	3	11
24	2	10	11	9	15	5	6	12	1	3	13	8	2	14	4	7
25	1	15	8	7	5	9	10	13	3	11	6	2	1	12	4	14
26	2	15	13	8	5	10	7	14	12	11	6	4	1	3	2	9
27	1	11	3	6	14	1	7	9	4	2	5	10	15	13	12	8
28	2	6	15	3	11	8	2	13	9	10	14	5	7	12	1	4
29	1	15	7	10	2	12	9	13	8	5	6	11	1	3	4	14
30	2	15	10	7	2	9	6	14	12	8	11	5	3	1	4	13
31	1	11	4	9	10	15	8	6	5	1	13	14	2	12	3	7
32	2	9	3	10	13	14	11	1	2	4	5	15	6	7	8	12
33	1	15	8	1	11	10	2	4	13	14	9	6	5	12	3	7
34	2	15	8	3	11	10	2	4	13	14	9	6	5	12	1	7
35	1	15	6	10	14	12	8	2	4	3	5	11	1	13	7	9
36	2	12	2	13	11	9	15	3	1	4	5	6	8	10	7	14
37	1	5	1	6	11	12	10	7	4	3	2	13	9	8	14	15
38	2	15	11	7	13	4	6	9	14	8	12	1	10	3	2	5
39	1	6	1	12	5	15	9	2	7	11	3	8	10	4	14	13
40	2	14	1	5	15	4	6	3	8	9	2	12	11	13	10	7
41	1	10	3	2	14	9	1	8	12	13	4	11	5	15	6	7
42	2	13	3	1	14	4	10	5	15	6	2	11	7	12	8	9

## 付録3 ハイブリッド・アプローチの最終解

X-座標				Y-座標			
Individuals	$X_1$	$X_2$	SSQ	Objects	$Y_1$	$Y_2$	SSQ
1	1.362	-0.062	1.859	TP	-0.476	1.244	1.774
2	1.956	-0.034	3.827	BT	-0.687	-0.027	0.473
3	0.753	-0.283	0.646	EMM	-0.169	-0.888	0.818
4	0.941	0.338	1.000	JD	0.796	1.114	1.875
5	0.669	-0.449	0.650	CT	0.008	-1.206	1.454
6	0.975	0.221	1.000	BMM	0.360	-0.909	0.956
7	0.997	-0.073	1.000	HRB	-0.845	-0.160	0.739
8	1.500	0.029	2.250	TMd	-0.424	0.503	0.433
9	0.897	-0.250	0.867	BTJ	-0.37	0.276	0.214
10	0.883	-0.470	1.000	TMn	-0.810	0.244	0.716
11	-0.860	-0.511	1.000	CB	0.699	0.556	0.797
12	-0.191	-0.578	0.370	DP	.707	0.276	0.576
13	0.466	-0.149	0.240	GD	0.652	0.900	1.234
14	1.266	-0.041	1.606	CC	0.782	-0.502	0.863
15	0.997	0.074	1.000	CMB	-0.227	-1.424	2.078
16	0.999	-0.575	1.329	Note: ベクトル・モデルを採用した個人は黄色で表示			
17	1.933	0.544	4.034				
18	1.062	-0.162	1.154				
19	0.328	1.086	1.286				
20	0.906	0.028	0.821				
21	-0.172	-0.060	0.033				
22	0.995	0.100	1.000				
23	0.497	-0.597	0.604				
24	0.288	-0.351	0.206				
25	0.639	0.061	0.412				
26	1.716	-0.121	2.960				
27	-0.906	-0.504	1.076				
28	.897	-0.605	1.170				
29	0.720	0.284	0.599				
30	0.996	0.092	1.000				
31	0.025	-0.088	0.008				
32	-0.749	0.274	0.636				
3	0.548	-0.749	0.862				
34	0.628	-0.748	0.954				
35	-0.271	-0.154	0.097				
36	-0.432	0.296	0.274				
37	-1.442	0.659	2.513				
38	1.039	0.703	1.573				
39	-0.567	0.792	0.950				
40	-1.384	-1.029	2.973				
41	-0.122	-0.660	0.451				
42	-0.448	-0.576	0.533				