

『人間行動の計量分析』（抜刷）
東京大学出版会刊，1990年5月

第2章 理想点判別分析法とその応用

高根 芳雄

1. はじめに

理想点判別分析法 (Ideal Point Discriminant Analysis ; 以下 IPDA と略称する) は本来予測変数として連続変数と離散変数が混在する場合の判別分析法として提案された (Takane & Shibayama, 1984). 判別の目的は被験者に関する情報からそれが属する集団を予測することにあるが, その背後には予測を可能ならしめる関連変数の構造分析が前提とされる. 同じ判別分析法が構造分析法としても有効に用いられることのゆえんである. 本論では紙幅の制限もあり, 主として分割表 (contingency table) の構造分析を例にとりながら IPDA の働きを見ていくことにしよう (Takane, 1987). IPDA の他の用法については Takane, Bozdogan & Shibayama (1987) を参照されたい.

例として表 1 の分割表を見てみよう (Maxwell, 1961a). このデータは 620 人の患者を分裂病 (SC), 躁うつ病 (MD), 不安症 (AX) から成る 3 つの基準群と, 4 つの症状 (不安症状 (A), 疑心症状 (S), 思考障害 (T), 罪悪妄想 (G)) の有無から成る 16 の症候パターンによって分類したものである. 表中 16 の症候パターンにおける 4 つの症状の有無は 1 (有) — 0 (無) で示されている. このようなデータで我々が興味を持つのはどの症状があるいはその組合せがどの基準群と関係が深いかを見極めることである.

このような興味に答える手法として 2 つの代表的統計手法が考えられる. 1 つは対数線形モデルで, この方法を用いると個々の症状 (症状の主効果) やその組合せ (交互作用) が基準群に及ぼしている効果の有意性が検定できる. ただこの方法では症候パターン間, 基準群間また両者の関連を示す全体像がつかみに

表1 Maxwell(1961a)のデータ

パターンの番号	予測変数のパターン				基準群における頻度		
	A	S	T	G	SC	MD	AX
1	0	0	0	0	38	69	6
2	0	0	0	1	4	36	0
3	0	0	1	0	29	0	0
4	0	0	1	1	9	0	0
5	0	1	0	0	22	8	1
6	0	1	0	1	5	9	0
7	0	1	1	0	35	0	0
8	0	1	1	1	8	2	0
9	1	0	0	0	14	80	92
10	1	0	0	1	3	45	3
11	1	0	1	0	11	1	0
12	1	0	1	1	2	2	0
13	1	1	0	0	9	10	14
14	1	1	0	1	6	16	1
15	1	1	1	0	19	0	0
16	1	1	1	1	10	1	0
				計	224	279	117

くい。他方、数量化3類(例えば駒沢, 1982; 岩坪, 1987; 双対尺度法とも呼ばれる(西里, 1982))を適用すると症候パターン及び基準群の空間配置が得られ、これより全体の関連が直ちに理解できる。ただしこの方法では逆に個別の効果の有意性を検定したり、得られた解の安定性を評価したりするのが容易でない。IPDAはこれらの2つの解析手法の欠点を補い、両者の長所を生かした解析手法である。

IPDAでは数量化3類と同様分割表の行と列が多次元空間内の点として表されるものと仮定する。次に特定の行から列に対する反応確率を両者の点間のユークリッド距離の減少関数として定義する。IPDAではこの反応確率が観測された頻度とできるだけ一致するよう点の空間配置を定める。従って定められた点間の距離が近いほど両者の関連が深いことを示している。IPDAではまた対数線形モデルにおけるように様々なモデル評価が可能である。これは分割表の行に関する情報(例えばどの症候パターンがどの症状の組合せから成っているか)を重回帰分析の予測変数のように用いて点の位置をその線形関数として定めモデルの適合度比較によってその効果の有意性を「検定」するものである。

2. モデル

本節では IPDA の基本モデルについて簡単に説明する。

分割表を $R \times C$ の行列 $F = (f_{ij})$ で表す。 f_{ij} は i 行における j 列の観測頻度を表す。分割表の行に関する情報を持った変数(以下、予測変数と呼ぶ)の行列を $X (R \times P)$ で表す。予測変数としては連続変数、離散変数、あるいは両者が混在していてもよい。ただし離散変数はあらかじめダミー変数化されているものとし、連続変数は既に中心化されているものと仮定する。また離散変数同士の交互作用、離散変数と連続変数の交互作用などもあらかじめ定義されていれば X の中に含めることができる。特に行間の構造を仮定したくない場合は $X = I$ (単位行列) と置けばよい。

行に対応する点(行点)の座標行列を $Y (R \times A)$ で表す。ここで A は点を表す空間の次元数である。 A は $R-1$, $C-1$ を越えることはない。IPDA では行に対応する点の座標行列が予測変数の一次結合によって与えられるものと仮定する。すなわち、

$$Y = XB \quad (2.1)$$

ここで $B (P \times A)$ は X にかかる重みの行列である。これより A の大きさは一次独立な予測変数の数によっても制限されることがわかる。

列に対応する点(列点)の座標行列を $M (C \times A)$ で表す。IPDA では推定するパラメータの数を最小限にとどめるため、この M が Y の重み付き平均(重心)によって与えられるものと仮定する。すなわち、

$$M = D_C^{-1} F' Y = D_C^{-1} F' X B \quad (2.2)$$

ここで D_C は F の列和を要素とする対角行列である。(2.2) をセントロイド制約と呼ぶ。この制約の経験的妥当性については第6節で触れる。

上で定義された Y と M を用いて行点と列点間のユークリッド距離を定義する。いま第 i 行と第 j 列間のユークリッド距離を d_{ij} で表すと、

$$d_{ij} = \left\{ \sum_{a=1}^A (y_{ia} - m_{ja})^2 \right\}^{1/2} \quad (2.3)$$

ここで y_{ia} , m_{ja} はそれぞれ Y , M の要素である。行 i が与えられたときの列 j

の生起(反応)確率を $p_{j|i}$ で表す. IPDA ではこの $p_{j|i}$ が

$$p_{j|i} = \frac{w_j \exp(-d_{ij}^2)}{\sum_{k=1}^C w_k \exp(-d_{ik}^2)} \quad (2.4)$$

で与えられるものと仮定する. ここで $w_j (\geq 0)$ は列 j の反応バイアスを表す. これは列 j の先験確率のようなものと考えると理解しやすい. w_j の絶対的大きさを定めるため, $\sum_j w_j = 1$ とする. モデル(2.4)は基本的には $p_{j|i}$ が w_j に比例して大きくなること, また $p_{j|i}$ が d_{ij} が大きくなるにつれ $\exp(-d_{ij}^2)$ に比例して小さくなることを示している. (2.4)は X の分布が指数分布に属するときの条件付分布として導かれる. また対数線型モデルや数量化法など他の手法との関連もつきやすい(第6節参照).

(2.4)の $p_{j|i}$ を用いてデータ全体の尤度を定義する.

$$L = \prod_i \prod_j (p_{j|i})^{f_{ij}} \quad (2.5)$$

モデルのパラメータ (\mathbf{B} ; $w_j, j = 1, \dots, C$) は(2.5)が最大化するよう定められる. IPDA ではスコア法によって尤度の最大化を行う. この方法は適当な初期値から出発してパラメータの値を逐次更新して行く方法であるが, 収束が速いばかりでなく推定されたパラメータの安定性に関する情報が得られるので便利である.

最大尤度 L^* が求まったら

$$\text{AIC} = -2 \ln L^* + 2n_p \quad (2.6)$$

を求め(Akaike, 1974), モデルの適合度を比較する. ここで n_p は推定されたパラメータの数で通常(特に別の制約がなければ),

$$n_p = p^* A - \frac{A(A-1)}{2} - (C-1) \quad (2.7)$$

で与えられる. ここで p^* は一次独立な予測変数の数である. AICはその値が小さいほどモデルの適合度が良いことを示している.

3. モデル評価の実際

前節で述べた IPDA の基本的枠組の中で様々なモデルが指定でき, その中

から最小 AIC 基準に従って最適なモデルを選択する。本節ではどのようなモデル比較が可能なのか、序節でとりあげた Maxwell (1961a) のデータを通じて概観することにしよう。

このデータで我々が興味を持つのは次のような点である。

1) 分割表の行と列を空間的に表現することは果して妥当か。この間は次のように言い換えられる。IPDA は空間的表現を仮定しない次の2つのモデルより予測力が高いか。

$$p_{j|i} = f_{ij}/f_{i.} \quad (3.1)$$

$$p_{j|i} = f_{.j}/f \quad (i = 1, \dots, C) \quad (3.2)$$

ここで $f_{i.}$, $f_{.j}$ はそれぞれ i 行, j 列の周辺頻度, f は全体の頻度を表す。

(3.1) を飽和モデル, (3.2) を独立モデルと呼ぶ。

2) 空間的表現が妥当だとすると空間の次元数は何次元必要か。

3) 16 の症候パターンの中に何らかの構造があるか、例えば4つの症状の主効果だけで行と列の関係がうまく説明できるか。或いは交互作用を考慮する必要があるか。主効果はすべて有意か。——これらはすべて最も効果的な予測変数の組を選択する問題に帰着する。

4) 3つの基準群の生起パターンには有意差があるか。どの群とどの群に有意差があり、どの群とどの群に有意差がないか。——これは分散分析における多重比較の問題と同一である。

5) 推定されたパラメータはどれ位安定しているか。

以上の問題は5)を除けばいずれも特定の仮説を具現したモデルを IPDA によって当てはめ、AIC を用いた適合度比較によって解決できる。例えば次元数の決定では次元数を組織的に変えて解を求め、その中から一番適合度の良い次元数を選ぶ。予測変数の選択は様々な予測変数の組を用いて解を求め、その中から最適な解を選ぶ。これを網羅的に行うこともあれば、先験的基準から確度の高いものだけをいくつか選んで比較することもある。また消去法のような組織的アルゴリズムを用いることもできる。基準群間の有意差は例えば $m_j = m_k$ (j 番目と k 番目の基準群に差がない) という仮説のもとで解を求め、それを $m_j = m_k$ という仮説を置かない場合と比較する。これを網羅的に行うと全部で $\sum_{k=1}^C S(C, k)$ 個の仮説が存在する (Takane, 1989b)。ここで $S(C, k) =$

表2 マックスウェル(1961a)のデータの解析結果

	K ⁺⁺	次元数	AIC(パラメータの数)	AICの下限
飽和モデル	3		855(32)	
理想点判別分析				
[1]1, 2, 3				
4つの主効果	3	2	841*(9)	
変数1を除去	3	2	945(7)	
変数2を除去	3	2	871(7)	
変数3を除去	3	2	1013(7)	
変数4を除去	3	2	915(7)	
2元交互作用付加	3	2	853(21)	
3元交互作用付加	3	2	866(29)	
4元交互作用付加	3	2	869(31)	
(無制約解と同等)	3	1	918(17)	888(2) ⁺
[2](1, 2), 3	2	1	1121(17)	1091(2) ⁺
[3]1, (2, 3)	2	1	970(17)	940(2) ⁺
[4](1, 3), 2	2	1	1098(17)	1068(2) ⁺
[5](1, 2, 3)	1	0	1096(2)	1296(2) ⁺
(独立モデルと同等)				

⁺⁺K: 列間の等値制約で互いに異なると想定された列の集合数

⁺変数選択によって到達しうるAICの下限がケース[1]次元2のAICよりも大きい場合

*最小AIC

$$\sum_{g=1}^k \binom{k}{g} (-1)^g (k-g)^C / k!. \text{ 推定されたパラメータの信頼性は情報行列}$$

$$I(\theta) = E \left[\left(\frac{\partial \ln L}{\partial \theta} \right) \left(\frac{\partial \ln L}{\partial \theta} \right)' \right]$$

を収束点で評価し、その逆行列をとることによって得られる。

表2は様々な仕様のモデルをMaxwellのデータに適用した結果である。次元数の決定、予測変数の選択、基準群間の有意性の検定をすべて網羅的に行おうとすると、可能な組合せの数は膨大な数にのぼるが、この例のようにうまく方略を組むと実際に当てはめなければならないモデルの数を相当数減らすことができる。この例ではまず3つの基準群間すべてに有意差があるものと仮定し、 $X = I$ (表中無制約解と表示されている)として、1次元解、2次元解を求めた。同時に4つの症状の主効果のみを予測変数とするモデルで2次元解を求めた。後者のモデルは直観的にかなり有望なモデルと考えられたからである。AICを比べるとやはりこのモデルの適合度が一番良いことがわかる。さらにこのモデルは飽和モデル(3.1)、独立モデル(3.2)と比べても当てはまりの良いことが

わかる。これら一連の事実は IPDA による空間表現が有効であることを、しかも 16 の症候パターンの間にはかなり単純な構造が含まれていることを示している。

ただし、まだ上で述べた制約付 2 次元解が最適解と決まったわけではない。例えば同じ制約付モデルの一次解はどうであろうか。この解は表 2 には含まれていない。これは次のような理由による。先に $X = I$ として無制約の 1 次元解を求めたことを述べた。他のいかなる 1 次元解もこの解の尤度を越えることはできない。従って、この解の尤度から、どんなにうまく X を操作しても、AIC の値をそれ以上に小さくすることのできない値、いわば AIC の下限とでもいうべき値を論理的に導き出すことができる(これは予測変数をすべて取り除いても無制約一次元解の尤度がそのまま保たれるという想定のもとで簡単に計算できる)。もしこれまで求められた解の中でこの下限より小さい AIC を持つ解があればいかなる一次解もその解の適合度を越えることはできない。従って他の一次解を一切求めずして、それらのすべてを考慮外に置くことができる。実際、一次元解の下限を計算してみると 888 となり、これは上で求められた制約付 2 次元解の AIC, 841 よりも大きい。従って一次解の中には最適解は存在しないことがわかる。

空間の次元数が 2 次元必要であるという事実は 3 つの基準群間にすべて有意差があることを間接的に示している。もし有意差のない群があるとすると、それに対応して有意な次元の数が減少する。このことを実際に確かめるために再び $X = I$ と置いて [2] $m_1 = m_2 \neq m_3$, [3] $m_2 = m_3 \neq m_1$, [4] $m_1 = m_3 \neq m_2$ の仮説の下で解を求めてみた。するといずれの場合も AIC の値が相当大きくなり、しかも X をいかに操作しても制約付 2 次元解を越えるモデルは得られないことがわかった。これで予想どおり、3 つの基準間にはすべて有意差があることが裏付けられた。また AIC の値の大きくなり方から群 1 と他の 2 群との差は大きく、群 2 と群 3 の差は相対的に小さいことがわかる。

4 つの症状の主効果はいずれも有意であろうか。このことを確かめるために 4 つの主効果を 1 つずつ取り除いたモデルを当てはめてみた。するといずれの場合でも AIC の増加が著しいことが判明した。このことから 4 つの主効果はいずれも有意であることがわかった。AIC の増加の様子から 3 番目の症状(T)

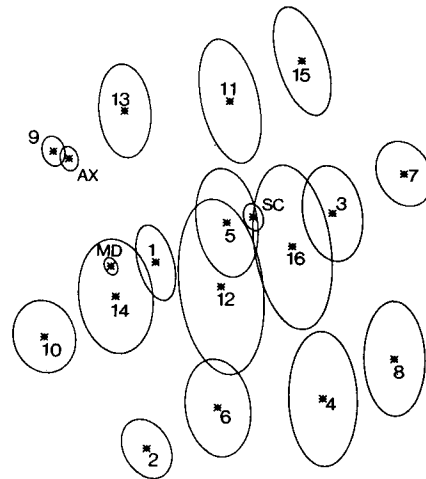


図1 Maxwell (1961a)のデータから4つの症状の主効果を予測変数として導かれた2次元解と各点の95%漸近信頼領域.

の効果が一番大きく、2番目の症状(S)の効果が一番小さいことがわかる。ここで2つ以上の主効果を同時に取り除いたらどうなるかという疑問が生じるが、主効果を1つずつ減らした場合の結果から4つの主効果をすべて含めた場合に匹敵するようなモデルは到底得られようがないことがわかる。

それでは有意な交互作用はないであろうか。この問題は交互作用の数が多い——2次の交互作用が6個、3次の交互作用が4個、4次の交互作用が1個存在する——ため100%正しい答を出すのは殆ど不可能に近い。そこで低次の交互作用から順次、しかも同次の交互作用を同時に予測変数に含め適合度に有意な改善が見られるかを調べることにした。AICを見る限り、有意な改善は見られない。このことから個々の交互作用の中で有意なものが絶対存在しないとは言いきれないが、たとえそのような交互作用があったとしてもその効果はそれほど大きくないという判断が成り立つ。従って、4つの主効果を予測変数とする2次元解がほぼ確実に最適解であると結論づけて差しつかえないであろう。

図1はこの最適解の空間表現である。16の症候パターンの布置が4つの症状の主効果の一次結合によって表されている。この図には特に示されていない

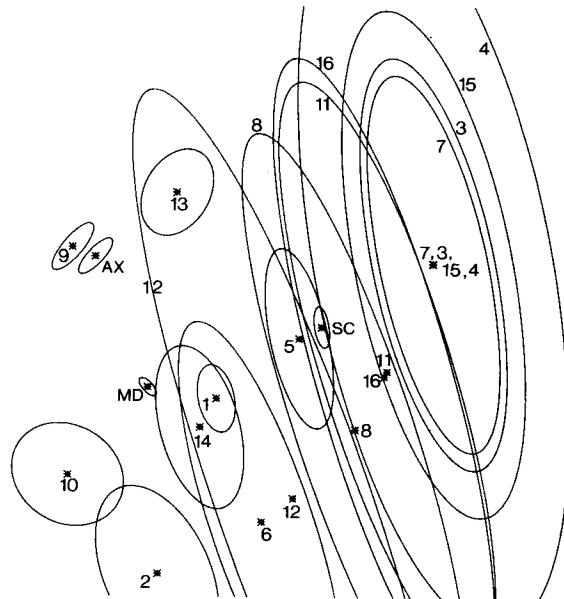


図2 Maxwell (1961a)のデータから $X = I$ として導かれた
2次元解と各点の95%漸近信頼領域。

が、T や S の存在は SC である確率を高め、T の不存在、G の存在は MD である確率を高め、A の存在は AX の確率を高める効果を持っている。図中の楕円は16の症候パターン及び3つの基準群を表す点の95%漸近信頼領域を表す。これより、この解では点の位置がかなり精密に定められていることがわかる。

参考までに図2に無制約の2次元解で得られた布置を提示する。既に述べたように、この解は最適解ではない。推定されたパラメータの数に見合って適合度が改善されなかったためである。図2の布置と図1のそれを見比べるといくつかの点はかなり移動していることがわかる。これはそれらの点はかなり動いても尤度はあまり変化しないことを示している。すなわち、これらの点の位置はあまり信頼性が高くない。これは図2の点の信頼性領域が非常に大きいことから確かめられる。このことは適当な行構造を指定することが時には安定した解を得るためにも必要であることを示している。

4. 制約条件のまとめ

本節では様々なモデルの仕様を可能にする制約条件についてまとめておく。これらの制約条件は3つのグループに大別できる。

(1) パラメータに直接課される制約

IPDAには2種類のパラメータが存在する。このうち w_j に課されるべき制約は簡単である。 w_j は制約なしで推定されるか、 $w_j = p_j = f_j/f$ と仮定されるか、すべての j につき一定 ($w_j = 1/C$) とされるかのいずれかである。面白いのは一定の w_j の組がいくつかの行に成り立ち、別の組が別の行に成り立つという仮定である(第7節参照)。一方、 B に課せられる制約は種類も多く若干説明を必要とする。

1) 予測変数の尺度水準によって異なる制約

予測変数の尺度水準を区別するため異なった水準に応じて次のような制約を置く。

(a) 名義尺度の場合： カテゴリー間の一次従属性を避けるため

$\sum_s n_{k(s)} b_{k(s)a} = 0, a = 1, \dots, A$ の制約を置く。ここで $n_{k(s)}$ は k 番目の名義尺度変数の s 番目のカテゴリーの反応頻度、 $b_{k(s)a}$ はそのカテゴリーにかかる重みを表す。これを中心化の制約と呼ぶ。

(b) 序数尺度の場合： 中心化の制約と所与の順序制約のもとに次元の重みを求め、それを多次的に重みづける。すなわち、 $b_{k(s)a} = b_{k(s)} \cdot b_{ka}$ として $b_{k(s)}$, b_{ka} を別々に求める。

(c) 間隔尺度の場合： カテゴリーにかかる重みは所与と考え、変数にかかる次元の重みだけを求める。

2) 次元数にかかわる制約

これは B の列数に関する。IPDAでは異なった次元数で解を求め、その中から最適な解を選ぶことによって次元数を決定する。

3) 定値制約、等値制約 B の特定の要素を特定の値に固定したり、2つ以上の要素が同じ値をとるように制約をかける。特に0値制約は変数の寄与の有無を検定するのに便利である。より一般的には $GBH' = Q$ の形の制約を課す

ことも可能である。ここで G , H , Q は定数行列である。

(2) 行構造にかかわる制約

予測変数を適当に操作することによって様々な行構造が表現できる。既に述べたように、特にとり立てて当てはめるべき構造がない場合には $X = I$ とする。これは行の間に特別な関係を想定しないことを意味する。ただ一般的には特定の行構造を具現した予測変数を定義し、その適合度を $X = I$ の場合と比較できるほうが望ましい。 X としては例えば離散変数の主効果のみを用いたり、いくつか交互作用を含めたりすることができる。また既に述べたように、離散変数と連続変数の交互作用を含めることも可能である (Takane ら, 1987)。

この他、行が実験のデザインに関係している場合、そのデザインを反映させたり (例えば対比較)、行が特定の刺激に対応している場合にはその刺激のデザインを予測変数として用いることもできる。また分散分析におけるコントラストやトレンドなども指定することができる。なお、特定の変数の影響を取り除く方法に関しては Takane (1987; p. 503) を参照されたい。

特定の変数の効果は、その変数を含めたモデルとそれを除いたモデルの適合度を比較することによって検定できる。これを組織的に行うと変数選択ができる。また特定の行どうしが等しいというような仮説の検定も特殊な変数(ないし変数群)の有意性検定の問題としてとらえられる。

(3) 列構造にかかわる制約

(2.2)式のセントロイド制約は殆どの場合に課される。列制約の中で最も重要なのは列間の等値制約である。例えば $m_1 = m_2$ は列1と列2に実質的な差がないことを表す。等値制約を組織的に用いて列間の多重比較を行う。またこの制約は列間に順序性がある順序制約を明示的に課しないと列点の表現がその順序性を満たさない場合にも有効性を発揮する(順序を乱す列には等値制約をかければよい)。

列にもセントロイド制約を置くかわりに行と同様の構造を指定することも原理的には可能である。分割表が正方形で行と列に対応関係が存在する場合は $M = Y$ とすることも考えられる (Takane & Shibayama, 1986)。

5. いくつかの例

前節に引き続き IPDA を用いて解析した分割表の例を2つあげる。

5.1 死刑判決と人種差別 (Takane, 1989a)

表3のデータは米国ジョージア州で過去殺人犯が死刑の宣告を受けた頻度と受けなかった頻度を被告の人種(黒人か白人か)と被害者の人種によって分類した結果である。一見して被告や被害者の人種が死刑判決に関係しているらしいことがわかる。ではいったいどのように関係しているのか。

表3 被告の人種と被害者の人種別死刑判決の頻度

	被告の人種	被害者の人種	死刑判決		死刑の割合
			可	否	
(1)	黒人	黒人	18	1420	1.2%
(2)	黒人	白人	50	178	21.9%
(3)	白人	黒人	2	62	3.1%
(4)	白人	白人	58	678	7.8%

(注) 「チャンス」という雑誌の第1号(1988年刊行)の7ページに引用されている Baldus と Woodworth の研究から転載。

この問に答えるため死刑か否かを基準変数として IPDA を適用した。この場合、可能な予測変数としては、1. 被告の人種、2. 被害者の人種、3. 1と2の交互作用(被告の人種と被害者の人種が一致するかしないか)の3つが考えられる。これら3つの予測変数の組合せから全部で8つのモデルができる。これらのモデルをすべて当てはめてみた。結果は表4に報告されている。表中3つの予測変数が特定のモデルに含まれるかどうかは Yes と No のパターンで示されている。8つのモデルのうち3つの予測変数を全て含むモデルは行に何ら特殊な関係を想定しない無制約モデルに一致する。さらにこの場合のように基準群が2つしか存在しない場合は無制約モデルと飽和モデルが一致する。また予測変数を全く含まないモデルは独立モデルと同等である。最小 AIC 基準から被害者の人種と、被告と被害者の人種の交互作用のみを含むモデルが一番適合

表4 表3のデータの解析結果

	被告人種	被害者の人種	2つの変数の の交互作用 ⁺⁺⁺	AIC	パラメータの数
1 ⁺	Yes	Yes	Yes	866.6	4
2	Yes	Yes	No	869.3	3
3	Yes	No	Yes	919.3	3
4	No	Yes	Yes	864.8*	3
5	Yes	No	No	999.8	2
6	No	Yes	No	895.1	2
7	No	No	Yes	937.5	2
8 ⁺⁺	No	No	No	1009.5	1

* 最小 AIC 解 + 飽和モデルと同等 ++ 独立モデルと同等
⁺⁺⁺ 被告人種と被害者の人種が一致するかどうか

表5 重みの推定値と標準誤差

変数	カテゴリー	重み	標準誤差
被害者	黒人	.473	.041
	白人	-.730	.063
被告と被害者の人種	一致	.087	.041
	不一致	-.653	.105
4つの組合せの座標			
(1)被告 被害者	黒人 黒人	.560	.031
(2)被告 被害者	黒人 白人	-1.383	.072
(3)被告 被害者	白人 黒人	-.180	.139
(4)被告 被害者	白人 白人	-.643	.074
理想点の座標			
死刑	可	-.755	.030
	否	.041	.002

度がよいことがわかる。

表5はこの2つの予測変数のカテゴリーにかかる重み、被告人種と被害者の人種の4つの組合せに対応する点の座標、死刑か否かの2つの基準群に対応する点の座標を示している。「死刑」の座標が負なので小さい座標値を持つ点ほど、死刑判決の確率が高いことを示している。また予測変数のカテゴリーにかかる重みは負の値ほど死刑の確率を高める要因となっている。すなわち、被

害者が白人のほうが黒人のときよりも死刑の確率が相当高い。しかも被告と被害者の人種が一致しないときのほうが死刑の確率が高い。従って、被害者が白人で加害者が黒人のとき(2)が死刑の確率が一番高くなり、被害者、加害者が共に黒人のとき(1)が死刑の確率が一番低い。

1つ注意しなければならないのは、座標の上では(4)のほうが(2)よりも「死刑」のカテゴリーに近いのに、(2)のほうが死刑の確率が高くなるのは何故かという点である。これは確率が距離そのものではなく、2つの基準群に対応する点までの距離の2乗の差によって決まるためである。実際、(2)から「無死刑」までの距離の2乗と「死刑」までの距離の2乗の差は(4)からの同じ差よりも大きい。従って、(2)の無死刑の確率は(4)のそれよりも低く、逆に死刑の確率は高くなる。

なお、ここで報告した解析では死刑の確率が被害者、加害者の人種の組合せに関係しているということはいえるが、人種が判決を左右しているとはまではいえない。死刑の確率と人種の間を説明する別の要因が存在するかも知れないからである。例えば、黒人が白人を殺すときはより暴力的な殺し方をすることが多いというような事実があれば、どうしてこの場合死刑の確率が高くなるかを説明できる。BaldusとWoodworthの原研究では、他に280もの説明変数を用意して死刑の確率と人種の間を説明しようと試みたがうまく行かなかった。このことから人種そのものが死刑か否かの判決に直接影響を及ぼしていることがほぼ確実であるといえる。

5.2 子宮癌手術後10年後の生存率(Madsen, 1976)

表6は299人の子宮癌の患者を次の5つの2値変数によって分類した結果である。

1. 腫瘍の段階(S)——低い(1), 高い(2).
2. 手術のタイプ(O)——全面切除(1), 部分切除(2).
3. 放射線治療(R)——なし(1), あり(2).
4. 腫瘍の病理(P)——局在的(1), 既に広がってしまっている(2).
5. 生存期間——手術後10年以内に死亡(1), 10年以上生存(2).

このデータでの我々の興味は生存期間が1~4の変数によっていかに予測で

表6 Madsen (1976)の例2で用いられたデータ

パターン番号	予測変数のパターン				生存期間	
					10年未満	10年以上
	S	O	R	P		
1	1	1	1	1	1	21
2	1	1	1	2	9	20
3	1	1	2	1	0	23
4	1	1	2	2	17	41
5	1	2	1	1	0	4
6	1	2	1	2	1	9
7	1	2	2	1	1	2
8	1	2	2	2	2	7
9	2	1	1	1	1	3
10	2	1	1	2	37	3
11	2	1	2	1	1	4
12	2	1	2	2	63	7
13	2	2	1	1	0	0
14	2	2	1	2	3	1
15	2	2	2	1	0	1
16	2	2	2	2	13	4

きるかという点である。

まず、すべての主効果、交互作用を含むモデルを当てはめ(基準群の数が2つなので、このモデルの適合度は飽和モデルのそれに一致する)、それによって得られた重みの大きさから有力と思われるモデルをいくつか選んで適合度を比較することにした(表7参照)。モデル1と3の比較から変数3の主効果は殆

表7 Madsen (1976)のデータでの変数選択

モデルの番号	予測変数	AIC(パラメータ数)
1	1,2,3,4	252.9 (5)
2	1,3,4	254.3 (4)
3	1,2,4	251.0 (4)
4	1,4	252.3 (3)
5	1,4,1×4	252.6 (4)
6	1,2,4,2×4	250.9 (5)
7	1,4,2×4	249.0*(4)
8	4,2×4	349.7 (3)
9	1,2×4	260.8 (3)
10	飽和モデル	263.5(15)
11	独立モデル	416.8 (1)

表8 Madsen (1976) のデータの Kategorii の重みの推定値

変数	カテゴリー	重み(標準誤差)
1	1	.58(.04)
	2	-.65(.05)
4	1	.66(.18)
	2	-.17(.05)
2×4	1	-.13(.05)
	2	.28(.12)
予測変数のパターン		変数パターンの座標
1	1111, 1121	1.53
2	1112, 1122	.28
3	1211, 1221	1.11
4	1212, 1222	.69
5	2111, 2121	.29
6	2112, 2122	-.95
7	2211, 2221	-.12
8	2212, 2222	-.54
生存期間		理想点の座標
	10年未満	-.62
	10年以上	.61

どないものと考えられる。変数2の主効果はモデル1と2, 3と4を比較する限りはあるように見えるが、変数2と4の交互作用を含めると変数2の主効果は無視してもかまわないことがわかる(モデル6と7の比較)。モデル4と7の比較から変数2と4の交互作用がわずかながら有意であることがわかる。またモデル7と8, 7と9の比較からそれぞれ変数1の主効果, 4の主効果が有意であることがわかる。特に変数1の効果が大きい。結局、変数1, 4の主効果, 変数2と4の交互作用を含むモデル7が最適モデルとして選ばれる。

表8で3つの予測変数のカテゴリーにかかる重みを見ると、腫瘍の発達段階の低いもの、腫瘍が局在的なもの、また腫瘍が局在的な場合は部分切除、広がってしまっている場合は全面切除をした場合(変数2と4の交互作用)が生存率が高くなることがわかる。これは生存率が高いほうの基準群がプラス側に位置し、正の重みを持つカテゴリーが生存に有利に働くと考えられるからである。結局もとの4つの2値変数で1111, 1121, 1211, 1221のパターンを持つ群が生存率が高く、2112, 2122のパターンを持つ群の死亡率が一番高い。

6. 連関モデル, 正準分析との比較

IPDA の基本モデル ((2.4)式) は d_{ij}^2 が

$$d_{ij}^2 = \sum_{a=1}^A y_{ia}^2 - 2 \sum_{a=1}^A y_{ia} m_{ja} + \sum_{a=1}^A m_{ja}^2 \quad (6.1)$$

と展開できることを考慮すると,

$$p_{ji} = \frac{w_j^* \exp(y_{ia}^* m_{ja}^*)}{\sum_{k=1}^C w_k^* \exp(y_{ia}^* m_{ja}^*)} \quad (6.2)$$

と書き直すことができる. (6.1) の $\sum y_{ia}^2$ は分子, 分母に共通なので消去され, $\sum m_{ja}^2$ は w_j^* に吸収されてしまう. (6.2) の y_{ia}^* , m_{ja}^* は y_{ia} , m_{ja} を単位調整したものである. (6.2) は m_{ja}^* が y_{ia}^* の関数となっていることを除けば, 連関モデル (例えば Goodman, 1981) から導かれる条件付確率に一致する. 従って, IPDA と連関モデルの基本的な違いは (2.2) のセントロイド制約を置くか置かないかにかかっていることがわかる. 2つのモデルの適合度を比較すると, セントロイド制約の妥当性が検証できる.

(2.4)式の p_{ji} を (2.5)式に代入して対数尤度を展開すると,

$$\ln L = - \sum_{i=1}^R f_i \ln \left(\sum_k w_k \exp(-d_{ik}^2) \right) + \sum_{j=1}^C f_j \ln w_j - \sum_{i,j} f_{ij} d_{ij}^2 \quad (6.3)$$

明らかに距離モデルのパラメータの推定に最も重要な役割を果すのは右辺の第3項である. しかるに $\sum_{i,j} f_{ij} d_{ij}^2$ を一定の規準化条件のもとで最小化すると数量化3類が導かれることが知られている (Heiser, 1981). 従って, IPDA は点の空間表現に関する限り, 数量化3類と類似した結果をもたらすことが予想される. Y が (2.1)式のように予測変数の一次結合として表される場合, 数量化3類は2類に帰着する. 従って, この場合, IPDA は数量化2類と類似した結果を得ると考えられる. 数量化法は本来記述的な方法であるが, 最近 Gilula と Haberman (1986, 1988) が「正準分析」と呼ばれる, 数量化2類・3類の最尤法とでもいうべき方法を提案している.

IPDA の適合度を実際のデータを用いて連関モデル及び「正準分析」に比較

してみた。用いられたデータは次の通りである。

1. Srole, Michael, Opler と Rennie (1962) : 1660 人のマンハッタンの住民を親の社会経済的地位を表す 6 つのカテゴリと 4 つの精神衛生上のカテゴリで分類したデータ。
2. Fisher (1940) : 5387 人の生徒を 4 色の目の色と 5 色の髪の毛の色で分類したデータ。
3. Snee (1974) : 592 人の被験者を 4 色の目の色と 4 色の髪の毛の色で分類したデータ。
4. Cramèr (1946) : 25263 のスウェーデンの家族を収入 (4 水準) と子供の数 (5 水準) で分類したデータ。
5. Rasch-Andersen (Goodman, 1986) : 告訴が取り下げられたデンマークの男性 7055 人を年度別 (4 年度) と年齢別 (5 水準) に分類したデータ。
6. Caussinus (1986) : 2730 人の患者を 5 つの年齢層と癌の種類によって分類したデータ。

表 9 は飽和モデルとそれぞれの方法で最適と判断されたモデルの尤度比から導かれたカイ 2 乗の値 (G^2) とその自由度を示している。AIC の値は比較の便宜上

$$AIC = G^2 - 2df \quad (6.4)$$

によって計算した。この式によると飽和モデルの AIC が 0 にセットされる。従って、AIC が負のモデルはいずれも飽和モデルより当てはまりが良いこと

表 9 最尤正準分析法と RC 連関モデルとの比較

データ	理想点判別分析		正準分析		連関モデル	
	$G^2(df)$	AIC	$G^2(df)$	AIC	$G^2(df)$	AIC
Srole, 他	3.6(12)	-20.4*	2.8(8)	-13.2	3.6(8)	-12.4
Fisher	9.4 (6)	-2.6*	4.7(2)	0.7	8.7(5)	-1.3
Snee	2.0 (4)	-6.0*	1.4(1)	-0.6	0.8(3)	-5.2
Cramèr	2.8 (6)	-9.2*	2.6(2)	-1.4	3.7(4)	-4.3
Rasch-Andersen	3.5 (9)	-14.5*	3.6(6)	-8.4	3.5(6)	-8.5
Caussinus	0.8 (5)	-9.2*	0.3(2)	-3.7	0.2(2)	-3.8

G^2 は飽和モデルとの差を表す対数尤度比から導かれたカイ 2 乗の値を示す。
* 最小 AIC

を示している。この基準によると、表9のモデルの殆どが一応満足のいくモデルであるが、6つのデータすべてについてIPDAの当てはまりが一番よい。このことは M に対するセントロイド制約が単に妥当であるばかりでなく、適合度を犠牲にすることなくパラメータ数を減らす有効な手段になっていることを示している。

7. 系列カテゴリー法(SCM)との比較(Takane, 1989b)

分割表の行に順序がある場合の扱いについては既に述べた(第4節(1)パラメータに直接かかる制約, 1)-(b)を参照のこと)。列に順序がある場合はどうか。IPDAでは順序を無視して多次元解を求めるか、列点が一次元空間内に順序通りに並ぶよう制約された解を求めるかのどちらかである。後者の場合、特に順序制約を課さなくとも、順序通りに並んだ列点が得られることもある。順序制約を課さないと順序が乱される場合は、列に関する等値制約を用いて順序制約をかける。どの列を等値しなければならないかは若干判断力を要する。例えば列が1,2,3,4の順で並ばなければならないのに1,3,2,4の順で並んだものとする。この場合、可能性としては $m_2 = m_3, m_1 = m_2, m_1 = m_2 = m_3, m_3 = m_4, m_2 = m_3 = m_4$ の5通りの等値制約が考えられる。5通りの解を求め、その中で順序制約を満たし、かつ適合度の一番良いものを選ぶ必要がある($m_2 = m_3, m_1 = m_2 = m_3, m_2 = m_3 = m_4$ の制約はほぼ確実に順序制約を満たす解を得るが、 $m_1 = m_2, m_3 = m_4$ については定かでない)。もっと複雑な場合はすべての可能な等値制約について解を求め、その中から一番適合度の高い解を選ぶことも考えられる。可能な等値制約の数は $2^{c-1}-2$ (-2 は等値制約を全くかけない場合とすべての列を等値する場合を除いたためである)で、これは列に順序がない場合の可能な等値制約の数に比べるとずっと少ない。列に順序がある場合、等値制約がかけられるのは隣接する列に限られるからである。

一方、計量心理学の分野では順序づけられたカテゴリーを一次元空間上の区間として表現する方法がよく用いられる(Torgerson, 1958)。この方法では行を一定の分布を持つ確率変数と考え、この確率変数が (t_{j-1}, t_j) で定義される j

番目の区間にはいる確率をもって p_{ji} とするものである。 i 番目の行に対応する確率変数の分布を平均 $x'_i \mathbf{b}$ のロジスティック分布とすると (McCullagh, 1980), この変数が t_j を越えない確率は

$$b_{ij} = 1/[1 + \exp(-(t_j - x'_i \mathbf{b}))] \quad (7.1)$$

で与えられる。従って

$$p_{ji} = b_{ij} - b_{i(j-1)} \quad (7.2)$$

が導かれる。これより IPDA と同様 (2.5) 式でデータ全体の尤度を定義し、それを最大化するようモデルのパラメータ (\mathbf{b} 及び $t_j, j = 1, \dots, C$) を定める。この方法を計量心理学の伝統に従って、系列カテゴリー法 (Successive Categories Method) 略して SCM 法と呼ぶ。モデルからすべての j につき $t_{j-1} \leq t_j$ が成り立つ (さもないと p_{ji} が負になってしまう)。従って、SCM にはもともと順序制約が組み込まれていると考えられる。(2.5) 式は

$$L = \left\{ \prod_{i \in G} \prod_{g \in G} (p_{g|i})^{f_{ig}} \right\} \left\{ \prod_{i \in g} (p_{j|i}/p_{g|i})^{f_{ij}} \right\} \quad (7.3)$$

と書き換えられる。ここで g は隣接する列の集合で、 G は g の集合、 $f_{ig}, p_{g|i}$ はそれぞれ

$$f_{ig} = \sum_{j \in g} f_{ij}, \quad p_{g|i} = \sum_{j \in g} p_{ji}$$

と定義される。(7.3) を用いると、SCM の順序制約を意識的にはずすことができる。まず順序を問わない列を合併して新しい分割表を作る。それを SCM で分析すると (7.3) の最初の部分が最大化される。従って、(7.3) の後半部分を別個に計算して前半の部分にかけてやれば、 g に含まれる列間に順序制約を課さない場合の尤度が計算できる。

IPDA と SCM は $C = 2$ の場合は完全に一致する (Takane, 1989b) が、 $C \geq 3$ の場合は一致しない。従って後者の場合、2つの方法の適合度を実際のデータを用いて比較してみると面白い。表 10 は Agresti, Choang & Kezouh (1987) で用いられた Maxwell (1961b) のデータで、5歳から15歳までの少年 223 人を5つの年齢層と夜驚症の重症度 (4水準) に応じて分類した結果である。列には明らかに自然の順序が存在する。表 11 は表 10 のデータに様々なモデルを当てはめた結果を示している。まず、もとの4つの列に何の制約も加えない場合の結

表 10 少年の夜驚症の重症度を年齢別集計した Maxwell(1961b)のデータ

年齢層	軽症			重症	計
	1	2	3	4	
5-7	7	4	3	7	21
8-9	10	15	11	13	49
10-11	23	9	11	7	50
12-13	28	9	12	10	59
14-15	32	5	4	3	44
計	100	42	41	40	223

表 11 Maxwell(1961b)データの解析結果

飽和モデル		574.5(15)		
独立モデル		583.0 (3)		
		列間の等値制約	すべて異なる 1, 2, 3, 4	列 2 と列 3 等しい 1, (2, 3), 4
IPDA				
行=名義尺度, dim =2		565.5(10)		
行=名義尺度, dim =1		561.7 (7)		562.1 (7)
行=間隔尺度, dim =1		561.6 (4)		562.0 (4)
SCM				
行=間隔尺度, dim =1		565.7 (7)		563.3 (7)
行=間隔尺度, dim =1		564.4 (4)		562.0 (4)
連関モデル				
行=間隔尺度, dim =1		566.2 (6)*		566.6 (6)*

* Agresti ら(1987)によって求められた。

果から見て行くことにしよう。行に特別の構造を仮定しないで(行=名義尺度), IPDA を適用し, 1次元解と2次元解を得た。両者の比較から空間の次元数は一次元で十分らしいことがわかる。このことはさらに行が年齢の線形関数で表現できる(行=間隔尺度)ことから裏付けられる(この場合, 1次元解しか求められない。しかもこの1次元解の当てはまりが一番よい)。同様のことを SCM でも行った。SCM の場合も, 行は年齢の線形関数として表した方がパラメータの数を軽減できて当てはまりもよい。ところが IPDA の最適解と SCM の最適解を比較すると, わずかながら IPDA の方が当てはまりが良い。また, IPDA の最適解は Agresti ら(1987)によって当てはめられた連関モデルよりもよい適合度を示している(これは前節の結果とも一致する)。

表 12 Maxwell (1961b) のデータから求められた座標の推定値

	順序制約なし	順序制約あり
列の座標		
1	.25	.25
2	-.20	-.15
3	-.10	-.15
4	-.31	-.31
行の座標		
1	-1.40	-1.40
2	-.69	-.69
3	-.13	-.13
4	.44	.44
5	1.00	1.00

表 12 の第 1 列は IPDA の最適モデルで得られた行及び列点の座標を示している。行点の座標は年齢層の中間値の線形関数になっているので 1 行目を除くと等間隔に配置される。列点の位置は列 2 と 3 の順序が入れ替わっているのがわかる。そこで順序制約を入れるため、 $m_2 = m_3$, $m_1 = m_2$, $m_1 = m_2 = m_3$, $m_3 = m_4$, $m_2 = m_3 = m_4$ の仮定のもとで IPDA を当てはめた。この結果、 $m_2 = m_3$ の仮定が順序制約を満たし、最もよい適合度を示すことがわかった。 $m_2 = m_3$ の仮定のもとでも行はやはり年齢の線形関数として表した方が当てはまりがよい。このモデルが列に対する順序制約のもとで IPDA の最適解を与える。しかも、このモデルはもともと順序制約を含む SCM の最適モデルよりも当てはまりがよい。また、このモデルは順序制約を課さない場合と同様、連関モデルの順序制約付解よりも当てはまりがよい。

SCM で列 2 と 3 の順序制約を既に述べた方法に従って意識的にはずした場合はどうであろうか。この場合、当てはまりはもとのモデルよりも若干よくなるが、IPDA で順序制約を置かない場合の適合度には及ばない。

表 12 の第 2 列は列に順序制約を置いた場合の最適モデルで得られたパラメータの値を示している。行の座標はほとんど変わらない。また列の座標も等置された m_2 と m_3 の値がおおよそ順序制約を置かない場合の値の平均に近い値をとる他はほとんど変化しない。

本節の結果はわずか一例にすぎないが、ほぼ同様のことが他の多数の例でも確かめられている (Takane, to appear).

8. 信号検出理論との関係 (Takane, 1989a)

心理学の実験で一定の刺激の組の中からランダムに刺激を提示し、被験者にその刺激が何であるかを当てさせる実験がある。これを刺激認定 (stimulus recognition または identification) 実験と呼ぶ。実験では様々な条件が操作され、それによって反応の確率がどのように変わってくるかを調べる。

用いられる実験操作には2種類のもので考えられる。1つは刺激の提示時間や反応の期限設定のように反応の質を一様に高めたり、低下させたりする効果を持つ。もう1つは刺激の提示確率や報酬表の操作のように一定の反応にバイアスをかける (特定の反応の確率を上げたり下げたりする) ものである。本節で問題にするのは後者である。反応のバイアス効果はその名の通り、バイアス・パラメータ w_j の変化としてとらえることができる。

刺激認定実験は刺激も反応も2つに限らないが、本節では一番簡単な 2×2 の場合のみを取りあげる。より一般的な刺激認定データの解析については Takane & Shibayama (1986) を参照されたい。 2×2 の場合は信号検出理論との関係で詳細な研究がなされている (例えば Green と Swets, 1966)。また、刺激が次元の場合は、SCM 型の定式化も可能である。

表 13 は2刺激認定実験から得られた 2×2 表を5つたてに並べて提示したものである (Laming, 1968)。各試行ごとに刺激 a か b のどちらかが提示され、被験者はそれを言い当てる。可能な2つの反応を刺激 a, b に対応して A, B と表す。1つの刺激提示条件のもとで 2×2 表が1つ得られる。この例では刺激 a の提示確率を .25 から .125 のきざみで .75 まで5段階に変化させ、それに応じた反応の変化をみている (1条件で24人の被験者から200反応ずつ、5条件で計24,000の反応が集められている)。

既に述べたように、バイアス・パラメータは刺激の提示確率によって組織的に変化することが予想される。それでは刺激間の距離はどうであろうか。信号検出理論によると、被験者の感受性の測度である $d = |y_a - y_b|$ は不変でなけ

表 13 2つの刺激の提示確率を操作して得られた
Laming (1968) の刺激認定データの解析

条件	$P(a)$	刺激	反応		(1)		(2)	
			A	B	d^+	w_1	d^{++}	w_1
1.	.250	a	1121	79	1.965	.261	1.950	.271
		b	33	3567				
2.	.375	a	1726	74	1.955	.381	1.950	.384
		b	49	2951				
3.	.500	a	2330	70	1.948	.482	1.950	.478
		b	59	2341				
4.	.625	a	2942	58	1.923	.615	1.950	.616
		b	83	1717				
5.	.750	a	3572	28	1.977	.756	1.950	.745
		b	8	1117				
AIC					5458.9	5454.6*		
パラメータ数					10	6		

* 最小 AIC

+ $P(a)$ が異なると d が変化するという仮定のもとで得られた。飽和モデルと同等。

++ $P(a)$ が変わっても d は一定という仮定のもとで得られた。

ればならない。しかし、実際にはどうか。この問題は d が一定という制約のもとで解を求め、その適合度を d を一定と仮定しないで求めた解の適合度と比較することによって解決される。 d を一定としない場合の解は IPDA を個々の 2×2 表に個別に適用して得られる(ただし 2×2 表に IPDA を適用したものは飽和モデルと同等である)。全体の尤度は個々の分析で得られた尤度を掛け合わせるによって得られる。表 13 の (1) と名付けられた列に、このモデルのもとで推定されたパラメータの推定値及び適合度が与えられている。見てわかる通り、 d の値に余り変化はない。 w_1 の値が刺激提示確率によって変化するのは予想通りである。表 13 の (2) に d が一定という制約条件のもとで求められたパラメータの推定値及び適合度が与えられている。最小 AIC 基準によるとこのモデルの方が当てはまりがよい。従って、このデータに関する限り、 d は刺激の提示確率によらず一定と考えてさしつかえない。26,000 個の観測値がわずか 6 つのパラメータで説明できるというのは実に驚異的である。

9. 結 語

本章では IPDA による分割表の解析を様々な角度から見て来た。応用例の解析に当たっては、分析法の特徴がなるべく前面に出るよう配慮したつもりである。序節でも述べたように、IPDA の利点は分割表の空間的表現と詳細なモデル評価にある。このことから IPDA はデータ全体のゲシュタルト的理解と解析的理解を 1つの方法に統合したものであることができる。

IPDA は今後も様々な方向に発展しうる手法である。例えば、距離モデル (2.3) を一般化したり、欠測値を扱えるようにすると手法の守備範囲は一層広がる。また、モデルを探索する一助として、誤差解析を含む診断機能を持たせることも大切である。また将来的にはこれまで実際のデータ解析を通じて蓄積されてきた IPDA を最も有効に使うためのノウハウを組み入れたエキスパート・システムを開発する計画も存在する。

なお IPDA の計算プログラム入手条件については下記の住所に連絡されたい。

Professor Yoshio Takane
Department of Psychology
McGill University
1205 Dr. Penfield Ave.
Montreal Quebec H3A 1B1
Canada

e-mail address : PS81 @ musica. mcgill. ca

参 考 文 献

- Agresti, A., Chuang, C., & Kezouh, A. (1987) Order-restricted score parameters in association models for contingency tables. *Journal of the American Statistical*

- Association*, 82, 619-623.
- Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19, 716-723.
- Caussinus, H. (1986) Discussion of paper by L. A. Goodman. *International Statistical Review*, 54, 274-278.
- Cramer, H. (1946) *Mathematical methods of statistics*. Princeton, NJ : Princeton University Press.
- Fisher, R. A. (1940) The precision of discriminant functions. *Annals of Eugenics*, 10, 422-429.
- Gilula, Z., & Haberman, S. J. (1986) Canonical analysis of contingency tables by maximum likelihood. *Journal of the American Statistical Association*, 81, 780-788.
- Gilula, Z., & Haberman, S. J. (1988) The analysis of contingency tables by restricted canonical and restricted association models. *Journal of the American Statistical Association*, 83, 760-771.
- Goodman, L. A. (1981) Association models and canonical correlation in the analysis of cross-classifications having ordered categories. *Journal of the American Statistical Association*, 76, 320-334.
- Goodman, L. A. (1986) Some useful extensions of the usual correspondence analysis and the usual log-linear models approach in the analysis of contingency tables. *International Statistical Review*, 54, 243-309.
- Green, D. M., & Swets, J. A. (1966) *Signal detection theory and psychophysics*. Huntington, NY : Krieger, (reprinted 1973).
- Heiser, W. J. (1981) *Unfolding analysis of proximity data*. Doctoral Dissertation, University of Leiden.
- 岩坪秀一(1987) 『数量化法の基礎』朝倉書店.
- 駒沢勉(1982) 『数量化理論とデータ処理』朝倉書店.
- Laming, D. (1968) *Information theory of choice-reaction time*. London : Academic Press.
- Madsen, M. (1976) Statistical analysis of multiple contingency tables : Two examples. *Scandinavian Journal of Statistics*, 3, 97-106.
- Maxwell, A. E. (1961a) Canonical variate analysis when the variables are dichotomous. *Educational and psychological Measurement*, 21, 259-271.
- Maxwell, A. E. (1961b) *Analyzing qualitative data*. London : Methuen.
- McCullagh, P. (1980) Regression models for ordinal data (with discussion) . *Journal of Royal Statistical Society, Series B*, 42, 109-142.
- 西里静彦(1982) 『質的データの数量化』朝倉書店.
- Snee, R. (1974) Graphical display of two-way contingency tables. *American Statistician*, 38, 9-12.
- Srole, L., Michael, T. S., Opler, S. T., & Rennie, T. A. C. (1962) *Mental health in metropolis : The midtown Manhattan study*. New York : McGraw Hill.

- Takane, Y. (1987) Analysis of contingency tables by ideal point discriminant analysis. *Psychometrika*, 52, 493-513.
- Takane, Y. (1989a) Ideal point discriminant analysis : Implications for multiway data analysis. In R. Coppi, et al., (Eds.), *Analysis of multiway data matrices*. Amsterdam : North Holland, (in press).
- Takane, Y. (1989b) Ideal point discriminant analysis and ordered response categories. *Behaviormetrika*, 26, 31-47.
- Takane, Y., Bozdogan, H., & Shibayama, T. (1987) Ideal point discriminant analysis, *Psychometrika*, 52, 371-392.
- Takane, Y., & Shibayama, T. (1984) Multiple discriminant analysis for predictor variables measured at various scale levels. Proceedings of the 12th annual meeting of the Behaviormetric Society of Japan, 99-100.
- Takane, Y., & Shibayama, T. (1986) Comparisons of models for stimulus recognition data. In de Leeuw, J. et al., (Eds.), *Multidimensional data analysis*. Leiden : DSWO Press, 119-148 (with discussion).
- Torgerson, W. S. (1958), *Theory and methods of scaling*. New York : Wiley.