4　尺度化と数量化

4.1　多次元尺度法

4.1.1　目　　的

対象間の類似度がデータとして与えられたとき，対象を空間内の点として表わし，互いに類似したものどうしを近くに，類似していないものどうしを遠くに配置(布置)する方法を**多次元尺度法**(multi-dimensional scaling; MDS)という(高根芳雄[1980]，斎藤堯幸[1980])．

図1の例は，トン(・)とツー(―)の組合せからなる36個のモールス信号の**混同率**(confusion rate)を示したものである．たとえば，A(・―)とK(―・―)がこの順で提示されたとき，22%の人が2つの信号を同じだと判断したことを示している．類似した信号は同じであると判断される確率が高いはずだから，これは1種の**類似性(度)データ**と考えられる．ところがこのような数値の配列をみせられただけでは，個々の刺激対について混同率の大小は云々できても一般的傾向としてどういう要因が混同率に関係しているかを読みとることは困難である．

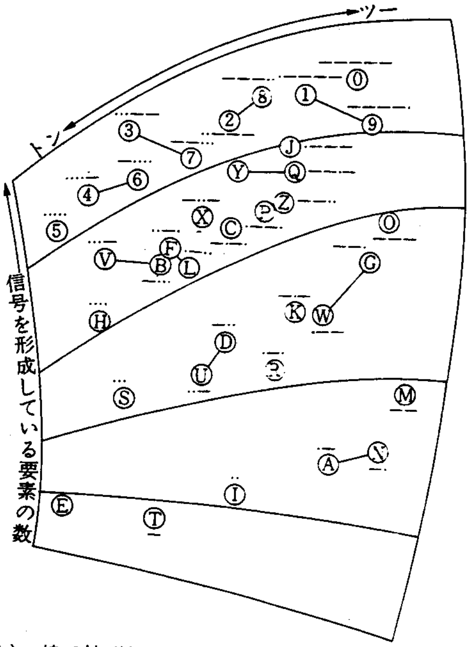
図2は図1のデータをMDSによって解析した結果を示したものである．**図2**をみると，確かに混同率が高い信号どうしは近くに，混同率の低い信号どうしは遠くに配置されていることがわかる．たとえばGとOは近くに配置されて

図1　R. N. Shepard [1963] によって分析されたモールス信号の混同行列

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7	**8**	**9**	0	
A	92	04	06	13	03	14	10	13	46	05	22	03	25	34	06	06	09	35	23	06	37	13	17	12	07	03	02	07	05	05	08	06	05	**06**	**02**	03	A
B	05	84	37	31	05	28	17	21	05	19	34	40	06	10	12	22	25	16	18	02	18	34	08	84	30	42	12	17	14	40	32	74	43	**17**	**04**	04	B
C	04	38	87	17	04	29	13	07	11	19	24	35	14	03	09	51	34	24	14	06	06	11	14	32	82	38	13	15	31	14	10	30	28	**24**	**18**	12	C
D	08	62	17	88	07	23	40	36	09	13	81	56	08	07	09	27	09	45	29	06	17	20	27	40	15	33	03	09	06	11	09	19	08	**10**	**05**	06	D
E	06	13	14	06	97	02	04	04	17	01	05	06	04	04	05	01	05	10	07	67	03	03	02	05	06	05	04	03	05	03	05	02	04	**02**	**03**	03	E
F	04	51	33	19	02	90	10	29	05	33	16	50	07	06	10	42	12	35	14	02	21	27	25	19	27	13	08	16	47	25	26	24	21	**05**	**05**	05	F
G	09	18	27	38	01	14	90	06	05	22	33	16	14	13	82	52	23	21	05	03	15	14	32	21	23	39	15	14	05	10	04	10	17	**23**	**20**	11	G
H	03	45	23	25	09	32	08	87	10	10	09	29	05	08	08	14	08	17	37	04	36	59	09	33	14	11	03	09	15	43	70	35	17	**04**	**03**	03	H
I	64	07	07	13	10	08	06	12	93	03	05	16	13	30	07	03	05	19	35	16	10	05	08	02	05	07	02	05	08	09	06	08	05	**02**	**04**	05	I
J	07	09	38	09	02	24	18	05	04	85	22	31	08	03	21	63	47	11	02	07	09	09	09	22	32	28	67	66	33	15	07	11	28	**29**	**26**	23	J
K	05	24	38	73	01	17	25	11	05	27	91	33	10	12	31	14	31	22	02	02	23	17	33	63	16	18	05	09	17	08	08	18	14	**13**	**05**	06	K
L	02	69	43	45	10	24	12	26	09	30	27	86	06	02	09	37	36	28	12	05	16	19	20	31	25	59	12	13	17	15	26	29	36	**16**	**07**	03	L
M	24	12	05	14	07	17	29	08	08	11	23	08	96	62	11	10	15	20	07	09	13	04	21	09	18	08	05	07	06	06	05	07	11	**07**	**10**	04	M
N	31	04	13	30	08	12	10	16	13	03	16	08	59	93	05	09	05	28	12	10	16	04	12	04	06	11	05	02	03	04	04	06	02	**02**	**10**	02	N
O	07	07	20	06	05	09	76	07	02	39	26	10	04	08	86	37	35	10	03	04	11	14	25	35	27	27	19	17	07	07	06	18	14	**11**	**20**	12	O
P	05	22	33	12	05	36	22	12	03	78	14	46	05	06	21	83	43	23	09	04	12	19	19	19	41	30	34	44	24	11	15	17	24	**23**	**25**	13	P
Q	08	20	38	11	04	15	10	05	02	27	23	26	07	06	22	51	91	11	02	03	06	14	12	37	50	63	34	32	17	12	09	27	40	**58**	**37**	24	Q
R	13	14	16	23	05	34	26	15	07	12	21	37	14	12	12	29	08	87	16	02	23	23	62	14	12	13	07	10	13	04	07	12	07	**09**	**01**	02	R
S	17	24	05	30	11	26	05	59	16	03	13	10	05	17	06	06	03	18	96	09	56	24	12	10	06	07	08	02	02	15	28	09	05	**05**	**05**	02	S
T	13	10	01	05	46	03	06	06	14	06	14	07	06	05	06	11	04	04	07	96	08	05	04	02	02	06	05	05	03	03	03	08	07	**06**	**14**	06	T
U	14	29	12	32	04	32	11	34	21	07	44	32	11	13	06	20	12	40	51	06	93	57	34	17	09	11	06	06	16	34	10	09	09	**07**	**04**	03	U
V	05	17	24	16	09	29	06	39	05	11	26	43	04	01	09	17	10	17	11	06	32	92	17	57	35	10	10	14	28	79	44	36	25	**10**	**01**	05	V
W	09	21	30	22	09	36	25	15	04	25	29	18	15	06	26	20	25	61	12	04	19	20	86	22	25	22	10	22	19	16	05	09	11	**06**	**03**	07	W
X	07	64	45	19	03	28	11	06	01	35	50	42	10	08	24	32	61	10	12	03	12	17	21	91	48	26	12	20	24	27	16	57	29	**16**	**17**	06	X
Y	09	23	62	15	04	26	22	09	01	30	12	14	05	06	14	30	52	05	07	04	06	13	21	44	86	23	26	44	40	15	11	26	22	**33**	**23**	16	Y
Z	03	46	45	18	02	22	17	10	07	23	21	51	11	02	15	59	72	14	04	03	09	11	12	36	42	87	16	21	27	09	10	25	66	**47**	**15**	15	Z
1	02	05	10	03	03	05	13	04	02	29	05	14	09	07	14	30	28	09	04	02	03	12	14	17	19	22	84	63	13	08	10	08	19	**32**	**57**	55	1
2	07	14	22	05	04	20	13	03	25	26	09	14	02	03	17	37	28	06	05	03	06	10	11	17	30	13	62	89	54	20	05	14	20	**21**	**16**	11	2
3	03	08	21	05	04	32	06	12	02	23	06	13	05	02	05	37	19	09	07	06	04	16	06	22	25	12	18	64	86	31	23	41	16	**17**	**08**	10	3
4	06	19	19	12	06	25	14	16	07	21	13	19	03	03	02	17	29	11	09	03	17	55	08	37	24	03	05	26	44	89	42	44	32	**10**	**03**	03	4
5	08	45	15	14	02	45	04	67	07	14	04	41	02	00	04	13	07	09	27	02	14	45	07	45	10	10	14	10	30	69	90	42	24	**10**	**06**	05	5
6	07	80	30	17	04	23	04	14	02	11	11	27	06	02	07	16	30	11	14	03	12	30	09	58	38	39	15	14	26	24	17	86	69	**14**	**05**	14	6
7	06	33	22	14	05	25	06	04	06	24	13	32	07	06	07	36	39	12	06	02	03	13	09	30	30	50	22	29	18	15	12	61	85	**70**	**20**	13	7
8	03	23	40	06	03	15	15	06	02	33	10	14	03	06	14	12	45	02	06	04	06	07	05	24	35	50	42	29	16	16	09	30	60	**89**	**61**	26	8
9	03	14	23	03	01	06	14	05	02	30	06	07	16	11	10	31	32	05	06	07	06	03	08	11	21	24	57	39	09	12	04	11	42	**56**	**91**	78	9
0	09	03	11	02	05	07	14	04	05	30	08	03	02	03	25	21	29	02	03	04	05	03	02	12	15	20	50	26	09	11	05	22	17	**52**	**81**	94	0
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	1	2	3	4	5	6	7	**8**	**9**	0	

（注）　E. Z. Rothkoph [1957] より．

図 2　モールス信号の多次元布置とその解釈



（注）　線で結ばれているのは互いに鏡映関係にある信号を表わす．R. N. Shepard [1963] より．

いるし，Ｅと０は遠くに配置されている．

図２をみると，次のようなことが容易に理解できる．i)混同率はほぼ２つの要因に規定されている．これは36個の信号がほぼ２次元空間に布置できることに対応する．ii)２つの要因は信号を構成する要素の総数と構成要素の中で特定の種類の要素（トンまたはツー）が占める割合に対応する．図２をみると同数の要素を持つ信号が横に帯状に布置され，しかも上に行くほど要素数の多い信号が見出される．また同数の要素を持つ信号の中で，左に行くほどトンの数が多くなり，逆に右に行くほどツーの数が多くなる．iii)これら２つの要因はほぼ独立に作用している．これは２つの要因の変化を示す方向がほぼ直交していることから理解できる．iv)ある信号と鏡映関係にある信号，すなわち構成要素がもとの信号と全く逆に配列されている信号は混同されやすい．図２では互いに鏡映関係を持っている信号どうしを実線で結んであるが，いずれの場合も互いに近い位置を占めている．

MDS はこのように対象間の類似性を規定している要因を探りあてるための方法である．実際には扱う類似性データの種類，距離モデルの種類，モデルとデータの一致度を定義する関数の種類などによってさまざまな MDS の手法が存在するが，いずれの場合も複雑さの中に埋没した類似性データの構造を誰の目にもわかりやすい形に提示することがその目的である．

4.1.2　概　　要

狭義の MDS はこのように対象間の類似性を点間の距離†(distance)で表現する方法である．これに対して対象を複数個の数値の組によって表現する方法一般を MDS ということもある．後者の意味では因子分析(→Ⅲ.3.4, Ⅰ.12.4)，主成分分析(→ Ⅰ.12.3, Ⅲ.3.2)などの**双線形モデル**(bilinear model)を用いた方法も１種の MDS と考えられるが，本稿では MDS を狭い意味に限定して用いる．ただし，後述するようにユークリッド距離の２乗とスカラー積(内積)の間には特別の関係があり，両者は類縁関係にあるといってさしつかえない．なお後述する数量化法(→Ⅲ.4.2)はどちらの意味でも MDS の１種と考えられる．

MDS はヤング＝ハウスホルダー(G.Young & A.S.Householder[1938])に端を発し，トーガソン(W.S.Torgerson[1952][1958])によって一応の完成をみる**計量 MDS**(metric MDS)と，1960年代に主としてシェパード(R.N.Shepard[1962])とクラスカル(J.B.Kruskal[1964a][1964b])等により発展させられた**非計量 MDS**(non-metric MDS)を区別することができる．前者が類似性の大きさを直接用いるのに対し，後者は類似性の大小関係のみにもとづいて対象の布置を決定する．非計量 MDS の出現により，MDS の適用範囲は飛躍的に増大した．

その後キャロルとチャング(J.D.Carroll & J.J. Chang[1970])により，複数個の**類似度行列**からその共通部分と独自性成分を分解してとらえる，**個人差 MDS**(individual difference MDS)が開発され，興味深い結果を生んでいる．

一方，さまざまな対象に対する選好の個人差をとらえるモデルとして**展開法**(unfolding method; **理想点モデル**ともいう)が，クームズ(C.H. Coombs[1964])によって提案されている．これは個人あるいは被験者の理想点と対象点を同時に多次元空間に配し，それらの点間の距離の減少関数によってある個人の特定の対象に対する選好の強さ(選好価)をとらえるモデルである．選好データを個人の理想と実際の対象の類似性を示すデータと解釈すれば，展開法も１種の MDS と考えられる．

MDS は対象間の類似性を最もよく表わすように点の布置を定める方法である．したがって入力データは類似性(proximity, similarity)データである．これより，点の位置を表わす座標行

列が求められる．点の位置が決まれば，一定の方式に従って点間の距離が計算できる．MDSはこの点間の距離がなんらかの意味でもとの類似性データと最もよく一致するよう点の位置＝座標を定める方法である．これを図式化すると

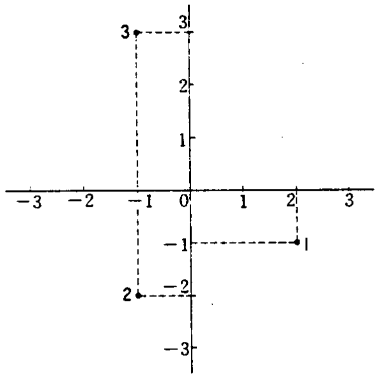
$$O \approx D(X)$$

と表わされる．ここでOはnを刺激(stimulus)数として，$n \times n$の類似性行列を表わす．Oの(i,j)要素(i行j列の要素)，o_{ij}は刺激iとjの類似性，あるいは非類似性を表わす．Oは少なくとも近似的に対称でなければならない．Xはrを空間の次元数として$n \times r$の座標行列を表わす．Xの(i,a)要素，x_{ia}は刺激iに対応した点の次元aの座標を表わす．$D(X)$はXから計算される点間の距離を要素とする$n \times n$の距離行列である．Dの(i,j)要素d_{ij}は，点iとjの距離を表わす．

点の座標を直交座標系を用いて図示すると図2に示されるような「地図」が得られる．たとえば3つの対象が2次元ユークリッド空間に位置づけられ，その座標が

$$X = \begin{bmatrix} 2 & -1 \\ -1 & -2 \\ -1 & 3 \end{bmatrix}$$

であったとすると，これら3つの対象は下図のように布置される．



こうしていったん点の位置が定められたら，後は座標軸を取り去ってもさしつかえない．ユークリッド距離は原点の移動についても座標軸の回転についても不変だからである．

点iとj間の(単純)ユークリッド距離はx_{ia}，x_{ja}を用いて

$$d_{ij} = \left[\sum_{a=1}^{r} (x_{ia} - x_{ja})^2 \right]^{1/2} \tag{1}$$

で求められる．上の例にこの式を当てはめると

$$D = \begin{bmatrix} 0 & \sqrt{10} & 5 \\ \sqrt{10} & 0 & 5 \\ 5 & 5 & 0 \end{bmatrix}$$

が得られる．$D(X)$はDがXの関数，すなわち点の位置によって変化することを示しているが，MDSはこのDがOと何らかの意味で最もよく一致するようにXを定める．OとDの一致度(agreement)をどのように評価するかは，入力データ，その他の条件によって変わってくる．

4.1.3 最も簡単な場合の解法

ここでデータとして誤差を含まないユークリッド距離が与えられたときの解法について簡単に触れておく(G.Young & A.S.Householder[1938])．この場合MDSは距離行列Dから座標行列Xを復元する操作に相当する．

d_{ij}^2を(i,j)要素とする行列を$D^{(2)}$で表わす．すると，この$D^{(2)}$はXを用いて

$$D^{(2)} = \mathrm{Diag}(XX^t)1_n1_n^t - 2XX^t + 1_n1_n^t\,\mathrm{Diag}(XX^t)$$

と表わすことができる．ここで1_nは要素がすべて1のn次元ベクトルを表わす．中心化行列を

$$J_n = I_n - 1_n1_n^t/n$$

と定義し，このJ_nを用いて$D^{(2)}$にヤング＝ハウスホルダー変換(Young-Householder transformation)すなわち

$$P = -2^{-1}J_nD^{(2)}J_n$$

を施すと，$J_n1_n = 0,\ 1_n^tJ_n = 0^t$より

$$P = J_nXX^tJ_n$$

が得られる．いま座標の原点をセントロイド(重心)にとると，$J_nX = X$．したがって

$$P = XX^t$$

が得られる．これはXが行列Pの平方根分解†(square root decomposition)によって得られることを示している．Xの列数は空間の次元数rに相当する．またPの階数に一致する．Pのスペクトル分解†を

$$P = Q_r\Lambda_rQ_r^t$$

とすると，Xの1つの解は

$$X = Q_r\Lambda_r^{1/2}$$

によって求められる．ところで$P1_n = 0$，したがってPの階数は高々$n-1$で，これはn個の点が常に$n-1$次元空間におさめられることを示している．また，Tをr次の直交行列とし，$X^* = XT$と定義すると

$$X^*X^{*t} = XTT^tX^t = XX^t.$$

したがって$D(X^*) = D(X)$が成り立つ．これはユークリッド距離が回転に関して不変であることに対応する．

4.1.4 誤差のある場合の解法

現実的には誤差なしに(非)類似性を測定する

のは不可能であるから，上記の解法はそのままでは実用価値に乏しい．ところが誤差が全くランダムならば，誤差がある場合でも上記と同様の方法が適用できる．これを一般に**トーガソン**(W.S.Torgerson[1952])**の解法**と呼んでいる．ただし，この場合入力データは非類似性データでなければならない(あるいは少なくとも既知の変換によって非類似性に変換できるものでなければならない)．

非類似性の2乗を要素とする行列を$\boldsymbol{O}^{(2)}$で表わす．この$\boldsymbol{O}^{(2)}$にヤング＝ハウスホルダー変換を施して得られる行列を$\hat{\boldsymbol{P}}$で表わす．すなわち

$$\hat{\boldsymbol{P}}=-2^{-1}\boldsymbol{J}_n\boldsymbol{O}^{(2)}\boldsymbol{J}_n$$

この$\hat{\boldsymbol{P}}$を最小2乗の意味で近似する$\boldsymbol{X}\boldsymbol{X}^t$を求める．すなわち，最小2乗基準を

$$\phi=\mathrm{tr}\{(\hat{\boldsymbol{P}}-\boldsymbol{X}\boldsymbol{X}^t)^2\}$$

と定義し，解に一意性を持たせるための制約条件，$\boldsymbol{X}^t\boldsymbol{X}=\boldsymbol{\Lambda}$ ($\boldsymbol{\Lambda}$は対角行列)のもとでϕを最小化する$\boldsymbol{X}$を求める．ϕを$\boldsymbol{X}$で偏微分して結果をゼロとおくと

$$\frac{\partial\phi}{\partial\boldsymbol{X}}=-2(\hat{\boldsymbol{P}}-\hat{\boldsymbol{X}}\hat{\boldsymbol{X}}^t)\hat{\boldsymbol{X}}=\boldsymbol{O}$$

したがって

$$\hat{\boldsymbol{P}}\hat{\boldsymbol{X}}-\hat{\boldsymbol{X}}\boldsymbol{\Lambda}=\boldsymbol{O}$$

なる固有方程式が得られる．いま$\hat{\boldsymbol{P}}$の大きいほうからr個の固有値を対角要素とする行列を$\hat{\boldsymbol{\Lambda}}_r$，それに対応する固有ベクトルの行列を$\hat{\boldsymbol{Q}}_r$とすると，$\hat{\boldsymbol{X}}$は

$$\hat{\boldsymbol{X}}=\hat{\boldsymbol{Q}}_r\hat{\boldsymbol{\Lambda}}_r^{1/2}$$

で求められる．一般に$\hat{\boldsymbol{P}}$は誤差なしの場合と違って非負定符号とは限らないが，r次元空間に布置を求めるためには少なくともr個の正の固有値を持つことが必要である．$\hat{\boldsymbol{X}}$は大きい方からr個の固有値に対応する固有ベクトルから求められる．ϕは残りの$n-r$個の固有値の2乗和に等しく，したがってϕを最小化するためには，これら残りの固有値の絶対値が小さくなければならない．また，$\hat{\boldsymbol{P}}$に負の固有値がある場合は，その絶対値はr番目の固有値を越えてはならない．

上記の解法は原則的には比尺度で測定された非類似性データにのみ適用可能であるが，加算定数を同時に推定することによって間隔尺度で測定された非類似性データにも比較的容易に拡張することができる．この項で説明した方法を**主座標分析法**(principal coordinate analysis)と呼ぶことがある(J.C.Gower[1966])．

4.1.5 類似性データの種類

MDSの入力データとなる類似性データには次のようなものがある．直接判断法として，i) **(非)類似性評定法**(rating method)，ii) **4つ組法**(method of tetrads)，iii) **3つ組法**(method of triads)，iv) **条件付き順位法**(method of conditional rank orders)，v) **分類法**(sorting method)，などが掲げられる．i)は刺激対を提示し，その(非)類似性をカテゴリー評定させる方法，ii)は2対の刺激対のうちより類似している対を選ばせる方法，iii)は標準刺激が2つの比較刺激のどちらにより類似しているかを判断させる方法，iv)は標準刺激を定め，残りの刺激を標準刺激との類似性によって順位づける方法，v)は刺激を類似性に応じていくつかの群に分類させる方法である．分類法では2つの刺激が同じ群に分類された頻度をそれらの刺激間の類似性の指標として用いることが多い．

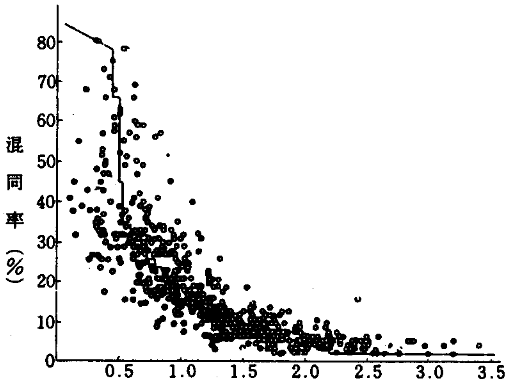
このほか間接法として，vi) **同異判断**(same-different judgment)における**混同率**，vii) 刺激同定判断(stimulus recognition)における混同率，viii) 刺激の般化率(rate of stimulus generalization)，ix) 事象の同時生起確率(probability of co-occurrence)，x) 反応時間(reaction time)，xi) 代替価(stimulus substitution)，xii) 連想価(word association)，xiii) 社会的相互作用の頻度(frequency of social interaction)，xiv) 相関係数(correlation coefficient)など多変量データから導かれた類似性指標などがMDSの入力データとして用いられる．

4.1.6 非計量 MDS

以上のように類似性データにはさまざまなものがあり，その中には距離と直線関係を持つような量に変換する方法が自明でないものも存在する．**非計量MDS**はこのようなとき，類似性の順序情報のみから対象の空間布置を定める方法である(R.N.Shepard[1962]，J.B.Kruskal[1964a][1964b])．すなわち，順序尺度で測定された類似性データに直接適用できる．しかもデータは類似性でも非類似性データでもさしつかえない．図2の結果は図1の混同率に非計量MDSを適用して得られた結果である．

非計量MDSはデータに**最適単調変換**を施し，距離モデルを単調変換されたデータに当てはめる．このとき，既知の単調変換でなく，距離モデルのパラメータ(刺激座標)を推定するのと同時に，データの最適単調変換を求めるところにその特徴がある．図3は図1のデータに非計量MDSを適用して得られた最適単調変換(線分でつないだもの)を示している．混同率は類似性データなので，距離との関係が逆単調に

図3 モールス信号の混同率の分析（図1，2から求められた距離の原データに対するプロット）



（注） R. N. Shepard [1963] より．

なっていることに注意．○で示されている実際のデータ点が最適単調変換からずれているのは誤差のためである．

いまデータの単調変換を $\delta_{ij}=F(o_{ij})$とすると，非計量 MDS の問題は，次の量の平方根

$$\phi=\sum_{j<i}(\delta_{ij}-d_{ij}(X))^2\Big/\sum_{j<i}d_{ij}^2(X)$$

をXとFに関して同時に最小化する問題に帰着する．ϕの平方根をストレスという．ϕは基本的には最小2乗基準であるが，規準化されていることに注意．これは δ_{ij}, $d_{ij}(X)$がともに0に収束するのを避けるためである．δ_{ij}はFがo_{ij}の単調関数であるという制約を受け，d_{ij}はそれがXの特定の関数であるという制約を受ける．ストレス最小化の基準はこれらの制約のもとでδ_{ij}とd_{ij}を全体として最もよく一致させることをめざしたものである．

ϕの最小化は次のようにして行う(J.B.Kruskal[1964b])．

$$\min_{X,F}\phi=\min_X\left(\min_{F|X}\phi\right)$$

(ただし，$\min_{F|X}\phi$はXを固定したときのϕのFに関する最小値.)より $\min_{F|X}\phi$を与えるFがXの関数として表わされるならば，$\min_{F|X}\phi=\phi^*$もまたXのみの関数として表わされる．すなわち

$$\min_{X,F}\phi=\min_X\phi^*$$

ϕのXとFに関する最小化はϕ^*のXに関する最小化によって置き換えられる．

では $\min_{F|X}\phi$となるようなFを求めるにはどうしたらよいか．まず固定されたXからd_{ij}を求める．d_{ij}の大きさの順序がデータから要求される単調性を満たしていれば，d_{ij}をそのままδ_{ij}とする．そうでなければ満たしていない距離の平均をとってδ_{ij}とする．こうして求められたδ_{ij}は$d_{ij}(X)$を一定としたとき，Fの単調性の制約のもとでϕを最小化することが知られている．これをクラスカルの**単調回帰法**(monotonic or isotonic regression)と呼ぶ(J.B. Kruskal[1964b])．単調回帰についてより詳しくはⅢ.7.1.3を参照．

ϕ^*をXに関して最小化するには何らかの逐次解法が必要である．クラスカルは**最急降下法**(steepest descent method)を用いている．この方法は適当な初期値から出発し，負の勾配$(-\partial\phi^*/\partial X$を$X$の現在値で評価したもの)の方向に$X$を逐次更新し，勾配が0となる点を求める方法である．$\boldsymbol{x}$をXのベクトルとし，$\boldsymbol{x}^{(q+1)}$, $\boldsymbol{x}^{(q)}$をそれぞれ$\boldsymbol{x}$の$q+1$回目，q回目の更新値のベクトルとすると，$\boldsymbol{x}$の最終値は

$$\boldsymbol{x}^{(q+1)}=\boldsymbol{x}^{(q)}-\alpha\boldsymbol{g}^{(q)}$$

を繰り返し適用することによって求められる．ここで$\boldsymbol{g}=(\partial\phi^*/\partial\boldsymbol{x})$($\boldsymbol{g}^{(q)}$はそれを$\boldsymbol{x}^{(q)}$で評価したもの)，$\alpha$はステップ・サイズを表わす．$\boldsymbol{g}$は$\phi^*=A/B$として

$$\frac{\partial\phi^*}{\partial\boldsymbol{x}}=\frac{1}{B}\left(\frac{\partial A}{\partial\boldsymbol{x}}-\phi^*\frac{\partial B}{\partial\boldsymbol{x}}\right)$$

ここで

$$\frac{\partial A}{\partial\boldsymbol{x}}=2\sum_i\sum_j(\delta_{ij}^*-d_{ij})\left(\frac{\partial\delta_{ij}^*}{\partial\boldsymbol{x}}-\frac{\partial d_{ij}}{\partial\boldsymbol{x}}\right)$$

$$\frac{\partial B}{\partial\boldsymbol{x}}=2\sum_i\sum_j d_{ij}\frac{\partial d_{ij}}{\partial\boldsymbol{x}}$$

で与えられる．ただし，δ_{ij}^*は単調回帰法によって求められたδ_{ij}の推定値を表わす．単調回帰法を用いたときは$\partial\delta_{ij}^*/\partial\boldsymbol{x}=\boldsymbol{0}$であることが知られている．$\partial d_{ij}/\partial\boldsymbol{x}$は距離関数によって変わってくるが，ユークリッド距離の場合は

$$\frac{\partial d_{ij}}{\partial x_{ka}}=\frac{(\Delta_{ik}-\Delta_{jk})(x_{ia}-x_{ja})}{d_{ij}}$$

(ただし，Δはクロネッカーのデルタ)を要素とするベクトルである．

なお非計量 MDS で用いられる逐次解法は最急降下法も含めて，勾配が0になる点を探す方法であるから，特定の初期値から出発して絶対最小(global minimum)が得られないことがある．初期値は前記のトーガソンの方法によって求めることが多いが，絶対最小への収束が疑わしい場合には，初期値をいろいろ変えて解を求めてみる必要がある．

また非計量 MDS は次元数や布置のクラスター性によっても変わってくるが，少なくとも10〜12個程度の刺激数が必要である．これは非計量 MDS では距離の大小関係のみで刺激の布置

を定めるため，相当な数の大小関係が拘束条件として働かないと解が一意に定まらなかったり，退化した解が得られたりするからである．

4.1.7 適用の要件

MDS を実際に適用する場合いくつかの考慮が必要である．まずデータの尺度水準の問題がある．尺度水準は類似性データが距離に関して持っている情報の多寡に関係しており，**比尺度**(ratio scale)，**間隔尺度**(interval scale)，**順序尺度**(ordinal scale)を区別する．比尺度の場合，データは距離の比に関する情報を持つと仮定される($o_{ij} \approx ad_{ij}$)．間隔尺度の場合，データは距離の差に関する情報を持つと仮定される($o_{ij} \approx ad_{ij}+b$)．順序尺度の場合は距離の大小関係に関する情報のみが有効とされる．データの尺度水準によって適用される MDS の手法も当然異なってくるから，MDS を適用するさいは注意しなければならない．一般に**尺度水準**(scale level)が低いほど(水準の高い方から，比尺度，間隔尺度，順序尺度である)適用範囲が広いが，尺度水準が高いほど推定の効率は高い．

次に距離関数の問題がある．数学的に距離と呼ばれるのは距離の3公理を満たす量であり，MDS で通常用いられるユークリッド距離を含む広い概念である．MDS でもユークリッド距離以外の距離関数(たとえばミンコフスキーのパワー距離)をあてはめる試みがなされているが，現在のところ解法の安定性の問題があり，実用性に乏しい．またユークリッド距離は，それが正しくない場合でもかなりよい近似を与えることが知られている．したがって，特に探索的目的で MDS を使う場合は，特別な理由がない限りユークリッド距離を用いるのが妥当であろう．

解析的に解が求められる場合以外は MDS を適用するのにさいして，空間の次元数を指定しなければならない．ただし次元数の指定は多分に仮説的な意味あいが強く，次元数の異なった解を求めて，適合度の変化と次元の解釈可能性を考慮しながら，最適な次元数を決定する．

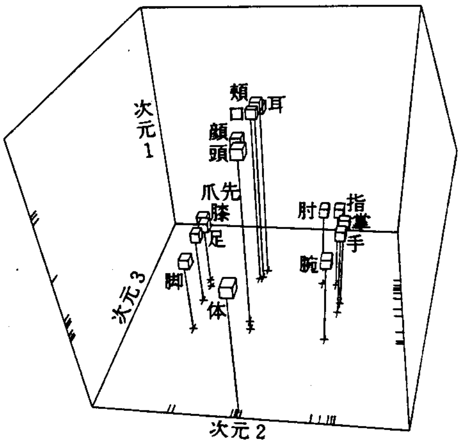
最後に適合度の指標(goodness of fit index)について簡単に触れる．MDS ではモデルをデータにあてはめる基準として最小2乗基準が用いられることが多い．これまでに説明した方法はいずれも最小2乗基準を用いている．パラメータを推定するときの便利さから基準を定義する要素としてスカラー積(内積)を用いたり，距離を用いたり，あるいは距離の2乗を用いたりする．最小化された基準の値は，データに対するモデルの適合度の指標として，たとえば次元数を決定するのに用いられる．最小2乗基準は分布の仮定を必要としない便利さから広範囲に用いられているが，一方その分布は未知でそれを適合度の指標として用いるときおのずと限界がある．最近ではデータに特定の分布の仮定をおく最尤推定法†が開発され(J.O.Ramsay[1977][1982], Y.Takane[1978][1981], Y.Takane & J.D.Carroll[1981])，次第に広く使われるようになってきた．最尤法の利点は推定値の信頼性その他のモデル評価が既製の統計理論にのっとって比較的容易に行えることである．

4.1.8 3元(個人差)MDS

これまでに述べた方法はいずれも単一の類似性行列を扱う方法であった．あるいは複数のデータ行列が存在する場合でも，それらの間に組織的な違いがなく，平均をとって解析してもかまわないという前提があった．ところが実際の場面ではこのような前提が成り立たないこともあり，逆にデータ行列間の組織的な違いがかえって興味深い結果をもたらすこともある．**3元 MDS**(3-way MDS; individual difference MDS)はこのようなときに適用される方法である．データが3方向に変化することからこの名前で呼ばれる．また第3の方向が個人を表わすことが多いため個人差 MDS†とも呼ばれる．ただし，第3の方向は測定時点や異なった測定条件のようなものであってもかまわない．ただ歴史的にこの方法が類似性判断における個人差をとらえるモデルとして発展したため個人差 MDS と呼ばれることが多い．3元 MDS の発展はキャロルとチャング(J.D.Carroll & J.J.Chang[1970])の貢献によるところが大きい．

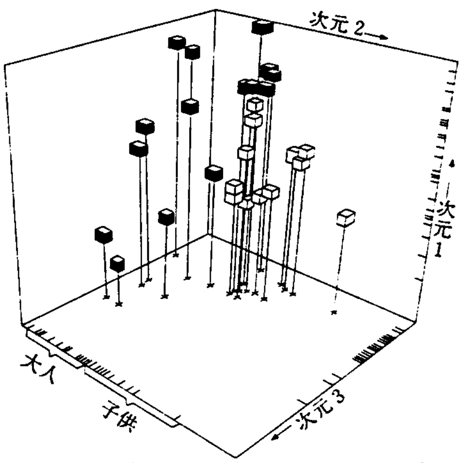
3元 MDS は複数個の類似性行列の共通部分と独自成分を分離表現する方法である．方法を理解するために図4および図5をみてみよう．これは大人と子供(7歳)のグループにさまざまな体の部位に関する言葉の類似性を判断させ，結果を3元 MDS を用いて分析したものである．図4は2つのグループの被験者(個人)に共通した体の部位間の類似性を表わしている．この図から体の部位間の類似性を表現するには3次元空間が必要なことがわかる．3つの次元はそれぞれ顔の部位対四肢(次元1)，上肢対下肢(次元2)，全体と末梢部位(次元3)を対照する次元と解釈することができる．図5は各被験者が類似性判断をするときにこれらの3つの次元にかけた重みを表わしたものである．この重みは個人がどの次元をどの程度重要とみなしてい

図 4　体の部位の名称の 3 次元プロット



（注）このプロットから次元 1 が頭部-四肢，次元 2 が上肢-下肢，次元 3 が中心-細部を対照する次元であることが明らかになる．

図 5　刺激布置に対応する個人差ウェイトのプロット



（注）子供が次元 2 に比重を置いて体の部位の非類似性を判断しているのに対し，大人の場合は次元 1 に重点を置く人と次元 3 に重点を置く人に分かれる点が注目をひく．

るかを反映している．図中黒で塗りつぶしたボックスが大人の重み，そうでないものが 7 歳児の重みを示している．これによると，2 つのグループは次元 2 できわだった対照をなしている．子供は大人と比べ相対的に次元 2，すなわち上肢と下肢の違いを重視している．大人は子供に比べるとグループ内での差が大きい．次元 1 を最も重視するものもあり，次元 3 を重視するものもいるが，次元 2 を重視する大人は 1 人もいない．この例は 3 元 MDS によって概念の認知構造の発達差をとらえたものである．3 元 MDS はこのように類似性判断における年代差，性差などの集団差を表現する方法としても利用できる．

3 元 MDS は，次式で定義される重み付きユークリッド距離モデルをあてはめる．d_{ijk} を個人 k（k 番目のデータ行列に対応する）における刺激 i, j 間の距離として

$$d_{ijk}=\left[\sum_{a=1}^{r} w_{ka}(x_{ia}-x_{ja})^2\right]^{1/2} \tag{2}$$

ここで $w_{ka}(\geq 0)$ は個人 k が次元 a にかける重みを表わす．図 5 はこの w_{ka} をプロットしたものである．解に一意性を持たせるため通常，空間の原点を重心におき，x_{ia} の分散を次元ごとに 1 と定める．これは w_{ka} と x_{ia} の間に単位の互換性が存在するためである．(2)式で x_{ia} の分散を一定とせず，すべての k と a につき，$w_{ka}=1$ とおくと，単純ユークリッド距離(1)が得られる．重み付きユークリッド距離の大きな特徴は単純ユークリッド距離と違って軸の回転の自由度が存在しないことである．したがって，求められた軸方向をそのまま解釈できる．単純ユークリッド・モデルでは距離モデルとは別の基準から軸方向を定めなければならない．解釈にあたっては多くの場合主軸方向が選ばれるが，主軸方向が常に経験的意味づけが容易とは限らず，MDS の結果を解釈するさい意味のある方向を探さなければならない．

誤差のない場合，3 元 MDS は単純ユークリッド・モデルの場合のヤング＝ハウスホルダーと同様の手続きで解くことができる．これはシェーネマン(P.H.Schönemann[1972])による．$D_k^{(2)}$ を d_{ijk}^2 を要素とする行列とすると

$$D_k^{(2)}=\mathrm{Diag}(XW_kX^t)1_n1_n^t-2XW_kX^t+1_n1_n^t\,\mathrm{Diag}(XW_kX^t)$$

と表わされる．ここで W_k は $w_{ka}(a=1,\cdots,r)$ を対角要素とする行列である．$D_k^{(2)}$ のそれぞれにヤング＝ハウスホルダー変換を施すと

$$P_k=-2^{-1}J_nD_k^{(2)}J_n=XW_kX^t$$

が得られる（$J_nX=X$ であることに注意）．いましばらく X と W_k を $N^{-1}\sum_{k=1}^{N}W_k=I$ となるよう単位調整するものとすると（N は個人数）

$$\overline{P}\equiv\frac{1}{N}\sum_{k=1}^{N}P_k=XX^t$$

となる．$\overline{P}$ の 1 つの平方根分解を YY^t とすると，T を適当な直交行列として $X=YT$ が成り立つ．よって任意の k につき $P_k=YTW_kT^tY^t$ でなければならない．これより

$$A_k\equiv(Y^tY)^{-1}Y^tP_kY(Y^tY)^{-1}=TW_kT^t$$

が得られる．これは T および W_k が A_k のスペ

クトル分解によって得られることを示している．T が求められたら $X=YT$．これより $W_{k'}$ $(k'\neq k)$ は

$$W_{k'}=(X^tX)^{-1}X^tP_{k'}X(X^tX)^{-1}$$

によって求められる．

誤差のある場合は若干解法が異なる．この場合，$O_k^{(2)}$ を個人 k の観測された非類似性の2乗の行列として，観測されたスカラー積行列を

$$\hat{P}_k=-2^{-1}J_nO_k^{(2)}J_n$$

と定義し，この $\hat{P}_k$ に XW_kX^t をあてはめる．最小2乗基準

$$\phi=\sum_{k=1}^{N}\mathrm{tr}\{(\hat{P}_k-XW_kX^t)^2\}$$

と定義し，ϕ を最小化する X および $W_k(k=1,\cdots,N)$ を求める．これは解析的には解けないが，キャロルとチャング(J.D.Carroll & J.J. Chang[1970])は ϕ を X と W_k につき交互に最小化し全体の最小を得るINDSCALという解法を提案している．

式の展開法を容易にするために，ϕ を別の行列を用いて表わす．ϕ の定義式の中にある2つの X を区別して $X^{(L)},X^{(R)}$ とする．

$$W=\begin{bmatrix}w_{11}\cdots w_{1r}\\ \vdots \quad \vdots\\ w_{N1}\cdots w_{Nr}\end{bmatrix}$$

$$Z=\begin{bmatrix}x_{11}^{(L)}x_{11}^{(R)}\cdots x_{1r}^{(L)}x_{1r}^{(R)}\\ x_{11}^{(L)}x_{21}^{(R)}\cdots x_{1r}^{(L)}x_{2r}^{(R)}\\ \vdots \quad \vdots\\ x_{n1}^{(L)}x_{n1}^{(R)}\cdots x_{nr}^{(L)}x_{nr}^{(R)}\end{bmatrix}$$

$$Q=\begin{bmatrix}\hat{p}_{111}\cdots \hat{p}_{11N}\\ \hat{p}_{121}\cdots \hat{p}_{12N}\\ \vdots \quad \vdots\\ \hat{p}_{nn1}\cdots \hat{p}_{nnN}\end{bmatrix}$$

と定義すると

$$\phi=\mathrm{tr}\{(Q-ZW^t)^t(Q-ZW^t)\}$$

これより $X^{(L)}$ と $X^{(R)}$ を一定としたときの W の最小2乗推定値は

$$\hat{W}^t=(Z^tZ)^{-1}Z^tQ$$

によって求められる．次に $X^{(R)}$ と W を一定としたときの $X^{(L)}$ の推定値は($X^{(L)}$ と W を一定としたときの $X^{(R)}$ の推定値は $X^{(L)}$ と $X^{(R)}$ の役割を交換すればよい)

$$V^{(R)t}=\begin{bmatrix}w_{11}x_{11}^{(R)}\cdots w_{1r}x_{1r}^{(R)}\\ \vdots \quad \vdots\\ w_{11}x_{n1}^{(R)}\cdots w_{1r}x_{nr}^{(R)}\\ \vdots \quad \vdots\\ w_{N1}x_{11}^{(R)}\cdots w_{Nr}x_{1r}^{(R)}\\ \vdots \quad \vdots\\ w_{N1}x_{n1}^{(R)}\cdots w_{Nr}x_{nr}^{(R)}\end{bmatrix},\quad R=\begin{bmatrix}\hat{P}_1\\ \vdots\\ \hat{P}_N\end{bmatrix}$$

と定義すると

$$\phi=\mathrm{tr}\{(R^t-X^{(L)}V^{(R)})^t(R^t-X^{(L)}V^{(R)})\}$$

と書けるから

$$\hat{X}^{(L)}=R^tV^{(R)}(V^{(R)t}V^{(R)})^{-1}$$

によって求められる．このようにして $X^{(L)}$, $X^{(R)}$, $W_k(k=1,\cdots,N)$ を順次収束が得られるまで繰り返し求める．$X^{(L)}$ と $X^{(R)}$ は同じ行列に収束することが知られている．

上記のINDSCALは比尺度で測定された非類似性データを前提とするが，序数データを扱う非計量3元MDSの方法としてALSCAL (Y. Takane & F.W.Young[1977])がある．

4.1.9　多次元展開法

これまで述べてきたMDSは，いずれも n 個の刺激間に明らかに類似性を表わすと考えられる量が観測されている場合に適用される方法であった．図6，図7の例は，選好データもその目的によって1種の類似性データと解釈することができて，MDSが適用可能であることを示している．

図6は82人のベルギー人の学生に16個の刺激に対する選好順位を求めた結果である．刺激は異なった家族構成で，それぞれ0人から3人に到る息子の数と娘の数の組合せからなる．16個の刺激がどの組合せに対応しているかは図の下に示してある．括弧でくくられた2つの数字は(息子の数，娘の数)を表わす．たとえば，(2, 1)は2人息子と1人娘の組合せを指す．このデータからたとえば，被験者1は(1,1)の組合せを1番望ましいと考え，(2,1)の組合せをその次に望ましいと考えていることがわかる．ところが，このようなデータから全体的傾向としてどのような傾向があるかを看取することはきわめて困難である．

図7はMDSの変法である(多次元)展開法†(multidimensional unfolding, C.H.Coombs[1964])を図6のデータに適用して得られた結果を示している．展開法は次のような仮定から導かれる．

i)個々の被験者は「理想(ideal)」を持っている．実際の刺激はこの理想に近いほどその被験者に好まれる．

ii)理想および実際の刺激は多次元空間内の点として表わされる．ある被験者の選好はその被験者の理想点と刺激点間の距離の減少関数として与えられる．このような考えから展開法は理想点と刺激点の布置を同時に多次元空間内に定める．

図7では個人の理想点は・で，刺激点は数字の対(息子の数と娘の数の組合せ)で示されてい

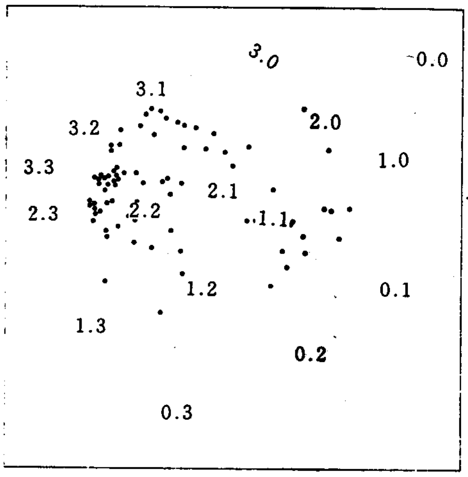
図6　ベルギー人の学生の家族構成に対する選好データ

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	16	6	8	4	7	1	2	12	5	3	10	13	9	11	14	15
2	16	7	4	5	15	1	3	8	10	9	2	6	14	13	11	12
3	16	14	9	7	15	6	4	2	10	5	1	11	8	3	12	13
4	16	15	13	12	14	8	3	5	11	2	1	7	10	4	6	9
5	16	12	10	13	14	8	2	7	11	6	5	1	15	9	3	4
6	16	7	5	6	15	11	1	2	13	10	4	3	14	9	12	8
7	16	12	11	6	15	5	2	3	14	4	1	8	13	7	9	10
8	16	8	6	5	14	9	4	2	15	11	3	1	13	12	7	10
9	14	15	10	12	16	4	2	5	11	3	1	7	13	6	8	9
10	15	16	13	9	14	10	7	8	11	5	4	2	12	6	1	3
11	16	13	4	3	15	1	2	9	11	8	5	7	12	14	10	6
12	16	14	6	3	15	5	2	4	12	8	1	7	13	9	11	10
13	16	14	9	8	15	4	1	11	10	2	5	3	13	12	6	7
14	16	14	10	11	15	9	5	6	13	8	4	2	12	7	3	1
15	15	14	12	16	13	11	9	10	8	7	5	6	2	3	1	4
16	16	6	2	7	8	1	4	10	3	5	9	11	13	14	12	15
17	16	15	13	10	14	9	8	7	12	5	3	6	11	4	1	2
18	16	8	5	2	15	3	1	10	9	4	6	7	14	11	12	13
19	16	14	12	13	15	10	7	6	11	8	1	3	9	5	4	2
20	15	12	10	14	11	6	2	5	9	1	3	8	13	4	7	16
21	16	15	10	11	13	9	8	5	12	7	1	3	14	6	2	4
22	16	14	11	10	15	9	5	6	13	7	1	3	12	8	4	2
23	16	11	6	12	13	4	1	5	14	2	3	7	15	10	9	8
24	16	15	13	12	14	7	8	10	4	1	5	9	2	3	6	11
25	16	14	10	9	15	8	6	5	12	7	1	2	13	11	4	3
26	16	12	11	10	15	5	6	8	14	7	4	2	13	9	3	1
27	14	15	11	9	16	8	4	5	13	6	2	3	12	7	10	4
28	16	13	12	14	10	1	4	15	5	3	2	8	11	6	7	9
29	16	14	8	7	15	10	6	4	13	9	5	1	12	11	3	2
30	1	2	5	7	4	3	6	10	14	12	8	9	16	15	13	11
31	14	15	10	11	16	6	4	7	12	8	1	3	13	9	5	2
32	16	14	11	9	15	10	2	4	13	3	1	6	12	5	7	8
33	16	4	3	8	6	1	2	10	7	5	9	12	13	11	14	15
34	15	7	6	9	4	1	3	12	5	2	10	14	8	11	13	16
35	16	7	6	5	15	8	3	1	14	9	2	4	13	10	11	12
36	16	2	4	5	11	3	1	6	13	7	8	9	15	14	12	10
37	16	7	8	9	6	3	2	12	4	1	10	14	5	11	13	15
38	16	4	9	10	7	3	2	5	11	6	1	13	12	8	14	15
39	16	12	11	10	15	7	5	8	14	6	1	3	13	9	4	2
40	16	2	3	12	4	1	6	10	5	7	8	9	13	11	14	15
41	8	2	3	10	5	1	4	11	6	7	9	12	13	14	15	16
42	16	9	4	7	12	1	2	8	5	3	6	10	11	13	14	15
43	16	15	13	12	14	8	5	7	10	4	1	3	11	6	2	9
44	16	14	6	5	15	9	7	1	13	10	2	3	12	11	8	4
45	16	10	8	9	12	5	4	2	14	11	1	3	15	13	7	6
46	16	13	1	2	15	3	4	6	11	5	7	8	14	12	9	10
47	1	2	5	9	3	4	7	12	6	8	11	14	10	13	16	15
48	10	3	1	5	4	2	7	11	6	8	12	14	9	13	15	16
49	16	13	4	1	15	6	2	3	14	7	5	8	12	10	9	11
50	16	2	7	9	3	1	5	11	4	6	8	13	12	10	14	15
51	16	15	10	13	14	5	7	8	12	6	2	4	11	9	3	1
52	16	15	12	11	14	10	3	4	13	2	1	6	9	8	5	7
53	16	12	11	9	13	10	4	5	14	6	1	2	15	8	3	7
54	16	12	11	9	15	7	5	6	14	10	1	2	13	8	3	4
55	16	15	13	11	14	4	2	6	12	1	3	8	10	5	7	9
56	16	14	13	11	15	6	2	8	12	4	1	3	10	7	5	9
57	16	15	13	10	14	11	6	8	12	5	3	4	9	7	2	1
58	16	15	13	9	14	11	8	5	12	7	3	2	10	4	1	6
59	16	14	8	7	15	12	6	2	13	10	3	1	11	9	4	5
60	12	16	14	10	15	11	8	6	13	7	4	2	9	5	1	3
61	16	15	13	12	14	9	8	6	11	7	2	4	10	5	3	1
62	16	14	10	13	15	9	7	2	11	8	1	4	12	3	5	6
63	16	14	11	13	15	8	6	5	12	7	4	1	10	9	3	2
64	16	12	5	2	15	11	6	1	14	10	7	3	13	9	8	4
65	16	10	11	13	12	9	7	5	14	8	4	2	15	6	3	1
66	16	10	8	3	15	9	5	2	14	11	6	1	13	12	7	4
67	16	10	12	14	11	9	7	3	13	8	1	2	15	6	4	5
68	15	5	3	6	10	2	1	9	8	7	4	11	12	13	14	16
69	16	15	13	10	14	6	8	9	12	5	1	7	11	3	2	4
70	16	12	8	9	15	10	6	4	14	7	5	2	13	11	3	1
71	6	3	5	9	2	1	8	13	4	7	11	15	10	12	14	16
72	16	12	11	10	15	9	5	7	14	6	4	1	13	8	2	3
73	14	13	8	11	12	2	4	5	15	3	1	9	16	7	6	10
74	16	10	11	14	12	9	1	3	13	5	2	4	15	7	6	8
75	16	14	13	15	12	6	10	8	11	5	1	4	9	7	3	2
76	16	12	10	11	14	4	1	3	13	5	2	6	15	7	8	9
77	16	12	4	3	15	9	5	2	14	10	6	1	13	11	8	7
78	16	10	11	13	15	4	2	6	14	3	1	5	12	7	8	9
79	1	2	3	4	9	5	6	7	14	11	10	8	16	17	13	12
80	16	14	11	10	15	9	7	5	13	8	4	2	12	6	3	1
81	16	15	13	11	14	9	8	6	12	7	4	3	10	5	2	1
82	16	12	8	4	15	11	7	3	14	10	6	2	13	9	5	1

（注）　表中1行目の1は(0,0)，2は(1,0)，3は(2,0)，4は(3,0)，5は(0,1)，6は(1,1)，7は(2,1)，8は(3,1)，9は(0,2)，10は(1,2)，11は(2,2)，12は(3,2)，13は(0,3)，14は(1,3)，15は(2,3)，16は(3,3)を表わす．

る．刺激の布置から(0,0)を除いて横軸がほぼ子供の総数(息子の数＋娘の数，縦軸が性差(息子の数－娘の数)に対応する次元と考えられる．0,0の組合せはほとんどの被験者が最も好ましくないと判断したため，位置を定める情報に乏しい．図7における位置もそれほど信頼性が高くないと考えられ，解釈のさいは無視してもさしつかえない．また(1,0)と(0,1)，(3,2)と(2,3)では息子と娘が1人ずつ入れかわっているが，前者のほうが距離がへだたっている．これは，息子と娘の入れかわりの影響が前者のほう

図7　家族構成の選好の展開法モデルによる分析



（注）　W. J. Heiser[1981]より．

が大きいことを示している．理想点の布置から，ベルギーの学生は大家族を好む傾向が強いこと，わずかながら娘よりも息子を好む傾向があることがわかる．

展開法はこのように類似性データと解釈された選好データを距離モデルを用いて表現する方法である．このとき選好の個人差は理想点の位置によって表わされる．展開法では通常，データが長方形をしている(図6参照)ことが普通のMDSと違っている．

n 個の刺激と N 人の個人の理想点の距離が誤差なしで与えられている場合，刺激座標 $\boldsymbol{X}$ および理想点の座標 $\boldsymbol{Y}$ を解析的に求めることができる(P.H.Schönemann[1970])．距離の2乗の行列($N\times n$)を $\boldsymbol{D}^{(2)}$ で表わす

$$\boldsymbol{D}^{(2)}=\mathrm{Diag}(\boldsymbol{YY}^t)\boldsymbol{1}_N\boldsymbol{1}_n^t-2\boldsymbol{YX}^t+\boldsymbol{1}_N\boldsymbol{1}_n^t\,\mathrm{Diag}(\boldsymbol{XX}^t)$$

と表わされるから，$\boldsymbol{D}^{(2)}$ にヤング＝ハウスホルダー変換を施すと

$$\boldsymbol{P}=-2^{-1}\boldsymbol{J}_N\boldsymbol{D}^{(2)}\boldsymbol{J}_n=\boldsymbol{J}_N\boldsymbol{YX}^t\boldsymbol{J}_n$$

が得られる．ここで $\boldsymbol{J}_N$ は N 次の中心化行列

$$\boldsymbol{I}_N-\boldsymbol{1}_N\boldsymbol{1}_N^t/N$$

を表わす．いま $\boldsymbol{P}$ の特異値分解†に基づいて $\boldsymbol{P}=\boldsymbol{HG}^t$ とと表わすと，適当な正則行列を $\boldsymbol{T}$ として

$$\boldsymbol{J}_n\boldsymbol{X}=\boldsymbol{GT},\quad \boldsymbol{J}_N\boldsymbol{Y}=\boldsymbol{H}(\boldsymbol{T}^{-1})^t$$

が成り立たなければならない．いま $\boldsymbol{X},\boldsymbol{Y}$ の共通の原点を $\boldsymbol{X}$ の重心におくと

$$\boldsymbol{X}=\boldsymbol{GT},\quad \boldsymbol{Y}=\boldsymbol{H}(\boldsymbol{T}^{-1})^t+\boldsymbol{1}_N\boldsymbol{y}_0^t$$

が成り立つ．

$\boldsymbol{y}_0$ は理想点の座標の原点を刺激座標の重心に合せるために必要な移動(translation)の大きさを表わす．

$$\boldsymbol{G}=\begin{bmatrix}\boldsymbol{g}_1^t\\ \vdots\\ \boldsymbol{g}_n^t\end{bmatrix},\quad \boldsymbol{H}=\begin{bmatrix}\boldsymbol{h}_1^t\\ \vdots\\ \boldsymbol{h}_N^t\end{bmatrix}$$

とすると

$$d_{ik}^2=\boldsymbol{g}_i^t\boldsymbol{T}\boldsymbol{T}^t\boldsymbol{g}_i-2(\boldsymbol{g}_i^t\boldsymbol{h}_k+\boldsymbol{g}_i^t\boldsymbol{T}\boldsymbol{y}_0)+\boldsymbol{h}_k^t(\boldsymbol{T}\boldsymbol{T}^t)^{-1}\boldsymbol{h}_k+2\boldsymbol{h}_k^t(\boldsymbol{T}^{-1})^t\boldsymbol{y}_0+\boldsymbol{y}_0^t\boldsymbol{y}_0$$

したがって

$$f_{ink}\equiv d_{ik}^2-d_{nk}^2+2(\boldsymbol{g}_i-\boldsymbol{g}_n)^t\boldsymbol{h}_k$$

と定義すると

$$f_{ink}=\boldsymbol{g}_i^t\boldsymbol{M}\boldsymbol{g}_i-\boldsymbol{g}_n^t\boldsymbol{M}\boldsymbol{g}_n-2(\boldsymbol{g}_i-\boldsymbol{g}_n)^t\boldsymbol{T}\boldsymbol{y}_0$$

と表わされる．ここで $\boldsymbol{M}=\boldsymbol{T}\boldsymbol{T}^t$ である．

$$\boldsymbol{m}=(m_{11},\cdots,m_{rr},m_{12},\cdots,m_{(r-1)r})^t$$

$$\boldsymbol{s}_i=(g_{i1}^2-g_{n1}^2,\cdots,g_{ir}^2-g_{nr}^2,2(g_{i1}g_{i2}-g_{n1}g_{n2}),\cdots,2(g_{i(r-1)}g_{ir}-g_{n(r-1)}g_{nr}))^t$$

$$\boldsymbol{z}=\boldsymbol{T}\boldsymbol{y}_0$$

$$\boldsymbol{u}_i=(2(g_{n1}-g_{i1}),\cdots,2(g_{nr}-g_{ir}))^t$$

$$\boldsymbol{K}_0=\begin{bmatrix}\boldsymbol{s}_1^t & \boldsymbol{u}_1^t\\ \vdots & \vdots\\ \boldsymbol{s}_{n-1}^t & \boldsymbol{u}_{n-1}^t\end{bmatrix}$$

と定義し，$\boldsymbol{K}_0$ を N 個重ねた行列を $\boldsymbol{K}$ とすると，f_{ink} を要素とする $N(n-1)$ 次元のベクトル $\boldsymbol{f}$ は

$$\boldsymbol{f}=\boldsymbol{K}\begin{bmatrix}\boldsymbol{m}\\ \boldsymbol{z}\end{bmatrix}$$

と表わされる．これより

$$\begin{bmatrix}\boldsymbol{m}\\ \boldsymbol{z}\end{bmatrix}=(\boldsymbol{K}^t\boldsymbol{K})^{-1}\boldsymbol{K}^t\boldsymbol{f}$$

が求められる．$\boldsymbol{m}$ を適当に並べ換えると $\boldsymbol{M}$ が得られ，$\boldsymbol{M}$ を平方根分解することによって $\boldsymbol{T}$ が求められる．最後に $\boldsymbol{T}$ と $\boldsymbol{z}$ より，$\boldsymbol{y}_0=\boldsymbol{T}^{-1}\boldsymbol{z}$ が得られる．

上記の解法は厳密には誤差のない $\boldsymbol{D}^{(2)}$ が与えられた場合にのみ通用する．誤差がわずかであれば $\boldsymbol{O}^{(2)}$ を $\boldsymbol{D}^{(2)}$ の代りに用いて解を求めることができるが，誤差が少しでも大きくなると，安定した解が得られないことが知られている．したがって，得られた解が本当の解であるかどうか十分な吟味が必要である．図7は非計量多次元展開法を用いて得た結果であるが，非計量データの場合はさらに問題が多く，退化した解が得られやすい．これはデータ数に比して，パラメータ数が多いためと考えられる．

展開法はマーケティング・リサーチで，よく売れる商品の開発や，特定の消費者層をねらった広告の戦略を開発する手段としてよく用いられる (W.S.De Sarbo & V.R.Rao[1984])．

あらかじめ通常の類似性データから MDS によって求められた刺激布置に，それとは別に与えられた選好データを最もよく表わすように個人の理想点を埋め込む方法を**外部展開法**(external unfolding)という(J.D.Carroll[1972])．この場合，個人 k の選好ベクトルを $\boldsymbol{s}_k$，k の理想点と刺激間の距離の2乗を要素とするベクトルを $\boldsymbol{d}_k^{(2)}$ とし

$$\boldsymbol{s}_k\approx a_k\boldsymbol{d}_k^{(2)}+b_k\boldsymbol{1}_n\quad(a_k\text{は通常負の値をとる})$$

なるモデルをあてはめる．$\boldsymbol{X}$ を刺激座標の行列，$\boldsymbol{y}_k$ を個人 k の理想点の座標ベクトルとすると

$$\boldsymbol{d}_k^{(2)}=\mathrm{Diag}(\boldsymbol{X}\boldsymbol{X}^t)\boldsymbol{1}_n-2\boldsymbol{X}^t\boldsymbol{y}_k+(\boldsymbol{y}_k^t\boldsymbol{y}_k)\boldsymbol{1}_n$$

だから

$$\boldsymbol{X}^*=[\mathrm{Diag}(\boldsymbol{X}\boldsymbol{X}^t)\boldsymbol{1}_n,-2\boldsymbol{X}^t,\boldsymbol{1}_n]$$

$$\boldsymbol{c}_k=\begin{bmatrix}a_k\\ a_k\boldsymbol{y}_k\\ b_k^*\end{bmatrix}\quad\text{ただし}\quad b_k^*=b_k+a_k\boldsymbol{y}_k^t\boldsymbol{y}_k$$

とすると

$$\hat{\boldsymbol{c}}_k=(\boldsymbol{X}^{*t}\boldsymbol{X}^*)^{-1}\boldsymbol{X}^{*t}\boldsymbol{s}_k$$

が得られる．これより

$$\hat{a}_k=\hat{c}_{1k},\ \hat{\boldsymbol{y}}_k=(\hat{c}_{2k},\cdots,\hat{c}_{(r+1)k})^t/\hat{a}_k$$

が求められる．このように外部展開法は単純な重回帰分析†に還元されるところから適用が簡単で実用価値が高い．

外部展開法で $\hat{a}_k$ が0となるとき，$\hat{\boldsymbol{y}}_k$ は無限に発散してしまうが，この場合その発散する方向だけが意味を持つと考えられる．この特別の場合を理想点モデルに対し**ベクトル選好モデル**という．このとき個人の選好はその個人の理想方向を表わすベクトルに個々の刺激を射影することによって得られる．

ベクトル選好モデルはお金のように多ければ多いほど望ましいといった対象の場合によくあてはまる．これに対し温度のようにあまり寒すぎず，暑すぎないところに適温があるような場合は理想点モデルのほうがよくあてはまる．ベクトル選好モデルでは等選好曲(直)線が個人のベクトルに直交する平行線になるのに対し，理想点モデルではそれが理想点を中心とした同心円をなす．なおベクトル選好モデルは双線形モデルの1種(主成分分析†の特別の場合)と考えられる．

4.1.10 プログラム，参考文献

MDS の適用は計算機プログラムの使用をぬきにしては考えられない．幸い汎用プログラム・パッケージである SAS†には ALSCAL(Y. Takane & F.W.Young[1977])，最尤 MDS 法の MLSCALE(J.O.Ramsay[1977][1982])などが収められている．ベル研究所で開発されたプログラム INDSCAL(J.D.Carroll & J.J.Chang[1970])や MDSCAL(J.B.Kruskal[1964a][1964b])などは直接ベル研究所に請求すればよい．プログラムの使い方に関する一般的解説書として，

S.S.Schiffman & M.L.Reynolds[1981]を掲げておく．この本には本稿で直接触れなかったプログラムの紹介もなされている．

MDS の方法論的側面をもっと重点的に解説した邦書に高根芳雄[1980]，斎藤堯幸[1980]，林知己夫・飽戸弘編[1976]がある．英語の文献では M.L.Davison[1983]があるのみである．MDS を応用するさいの助言を豊富に含んだ実践的手引書として，J.B.Kruskal & M.Wish[1978]を掲げることができる．

4.2 数量化法

数量化法(quantification method)は林知己夫によって開発されたカテゴリカル・データの解析手法の総称である(C.Hayashi[1952]，林知己夫・樋口伊佐夫[1970]，駒沢勉[1978])．開発者の名前をとって**林の数量化理論**(quantification theory)と呼ばれることもある．主な手法は1類から4類までの4種類で，このうち特に重要なのは3類である．

4.2.1 数量化法1類

数量化法1類(quantification method I)は説明変数(アイテム，項目)がカテゴリカル・データ，基準変量(目的変数)が計量データ†で与えられたとき適用される方法で，説明変数のカテゴリーに基準変数を最もよく予測するような数値を付与することを目的とする．これは説明変数をダミー変数で表示すると，1種の重回帰分析に還元する．したがって同じくダミー変数への重回帰分析に相当する分散分析†と形式的に一致する．ただし数量化法1類ではカテゴリー効果の有意性を検定することよりも，効果そのものの解釈に重点が置かれる．なお数量化法1類で説明変数が2つ以上ある場合，カテゴリー間に1次従属性が生じ数量化の一意性が失われる．これは効果の絶対的大きさでなく，その差のみが解釈可能であることを意味する．数量化法1類の適用のさい，カテゴリー間の1次従属性をなくすため各説明変数の最後のカテゴリーを除くことがよく行われるが，これは最後のカテゴリーに数量0を与えることに相当する．こうすることにより数量化に一意性が得られる．

i 番目の説明変数をダミー変数表示した行列を $\boldsymbol{H}_i$ とし，この $\boldsymbol{H}_i$ から任意の列(通常は最後の列)を取り除いた行列を $\tilde{\boldsymbol{H}}_i$ とする．この $\tilde{\boldsymbol{H}}_i$ を用いて

$$\boldsymbol{H}=[\boldsymbol{1}_N,\tilde{\boldsymbol{H}}_1,\cdots,\tilde{\boldsymbol{H}}_n]$$

と定義する．ここで $\boldsymbol{1}_N$ はその要素がすべて1よりなる N 次元ベクトル(N はサンプル数)，n は説明変数の数を表わす．i 番目の変数のカテゴリー数を n_i とすると $\boldsymbol{H}$ の階数はその列数

$$\sum_{i=1}^{n}(n_i-1)+1=\sum_{i=1}^{n}n_i-n+1 \quad (<N)$$

に一致する．

$\boldsymbol{y}$ を基準変数のベクトルとすると，数量化法1類は $\boldsymbol{y}\approx\boldsymbol{Hb}$ で，$\boldsymbol{b}$ を求める問題に帰着する．最小2乗基準を

$$\phi=(\boldsymbol{y}-\boldsymbol{Hb})^t(\boldsymbol{y}-\boldsymbol{Hb}) \tag{3}$$

と定義すると，$\boldsymbol{b}$ の最小2乗推定値は

$$\hat{\boldsymbol{b}}=(\boldsymbol{H}^t\boldsymbol{H})^{-1}\boldsymbol{H}^t\boldsymbol{y}$$

で求められる．説明変数の効果の大きさを表わす指標としては，特定の変数にかかわる $\hat{\boldsymbol{b}}$ の要素のレンジ(range)を用いることが多い．このさい各変数で除かれたカテゴリーには暗黙裡に0が付与されていることに注意しなければならない．もちろん冗長なカテゴリーを取り除かず $\boldsymbol{H}$ に含め，そのかわりに $\boldsymbol{H}^t\boldsymbol{H}$ の一般逆行列(→Ⅷ2.4.2)を使うことも可能である．

(→Ⅷ2.4.2)

$\boldsymbol{y}$ を観測個体(多くの場合，個人)の1次元ユークリッド空間における座標を表わすものとすると，ϕ を距離モデルによって解釈できる．$\boldsymbol{y}^*=\boldsymbol{Hb}$ を $\boldsymbol{H}$ を通して予測できる観測個体の座標ベクトルとすると，(3)式は

$$\phi=\sum_{k=1}^{N}d^2_{y_k y_k^*}$$

と表わされる．ϕ の最小化の基準は y_k^* が全体として最も y_k に近く位置づけられるよう $\boldsymbol{b}$ を定める問題に帰着する．

数量化法1類の適用例として柳井晴夫・高根

図8 分譲マンションの立面図

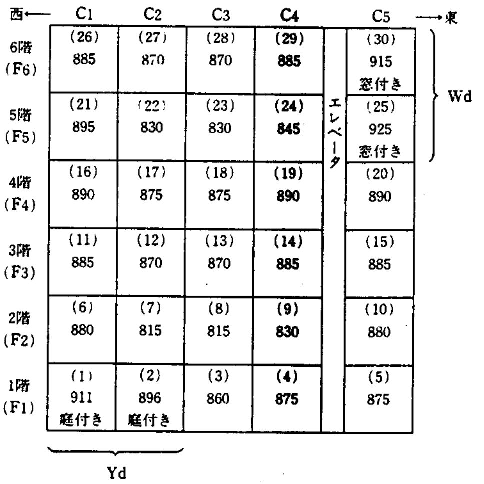


表1　分譲マンションの価格決定要因データのダミー変数表示

	基準変数	定数	階		数			平面的位置				庭	窓	交互作用
部屋番号	Y	d	F1	F2	F3	F4	F5	C1	C2	C3	C4	Yd	Wd	Int
1	911万	1	1	0	0	0	0	1	0	0	0	1	0	0
2	896	1	1	0	0	0	0	0	1	0	0	1	0	0
3	860	1	1	0	0	0	0	0	0	1	0	0	0	0
4	875	1	1	0	0	0	0	0	0	0	1	0	0	0
5	875	1	1	0	0	0	0	0	0	0	0	0	0	0
6	880	1	0	1	0	0	0	1	0	0	0	0	0	1
7	815	1	0	1	0	0	0	0	1	0	0	0	0	0
8	815	1	0	1	0	0	0	0	0	1	0	0	0	0
9	830	1	0	1	0	0	0	0	0	0	1	0	0	0
10	880	1	0	1	0	0	0	0	0	0	0	0	0	1
11	885	1	0	0	1	0	0	1	0	0	0	0	0	0
12	870	1	0	0	1	0	0	0	1	0	0	0	0	0
13	870	1	0	0	1	0	0	0	0	1	0	0	0	0
14	885	1	0	0	1	0	0	0	0	0	1	0	0	0
15	885	1	0	0	1	0	0	0	0	0	0	0	0	0
16	890	1	0	0	0	1	0	1	0	0	0	0	0	0
17	875	1	0	0	0	1	0	0	1	0	0	0	0	0
18	875	1	0	0	0	1	0	0	0	1	0	0	0	0
19	890	1	0	0	0	1	0	0	0	0	1	0	0	0
20	890	1	0	0	0	1	0	0	0	0	0	0	0	0
21	895	1	0	0	0	0	1	1	0	0	0	0	0	1
22	830	1	0	0	0	0	1	0	1	0	0	0	0	0
23	830	1	0	0	0	0	1	0	0	1	0	0	0	0
24	845	1	0	0	0	0	1	0	0	0	1	0	0	0
25	925	1	0	0	0	0	1	0	0	0	0	0	1	1
26	885	1	0	0	0	0	0	1	0	0	0	0	0	0
27	870	1	0	0	0	0	0	0	1	0	0	0	0	0
28	870	1	0	0	0	0	0	0	0	1	0	0	0	0
29	885	1	0	0	0	0	0	0	0	0	1	0	0	0
30	915	1	0	0	0	0	0	0	0	0	0	0	1	0

芳雄[1977]から引用する．図8はあるマンションの間取りと分譲価格を示したものである．カッコの中の数字が部屋番号，その下の数値が価格を表わす．価格の規定要因を探るために，価格を基準変数，部屋のある階(F)，平面位置(C)，庭(Yd)と余分の窓(Wd)のあるなし，階と平面位置の交互作用(Int)を説明変数として数量化法1類を適用した．表1は要因をダミー変数化したものである．たとえば1号室は1階の西端に位置しているので F1＝1，F2＝F3＝F4＝F5＝0，C1＝1，C2＝C3＝C4＝0，庭付きなので Yd＝1，余分の窓はないので Wd＝0，また2階と5階のカド部屋ではないので Int＝0 となる．

表2は数量化法1類の結果を示したものである．定数項dに与えられた885万はdのみが1であとはすべて0となるような部屋の予測価格を示している．実際の部屋の価格は，その部屋が持っている特徴に応じて，価格が上下され，全体の価格が決まる．たとえば，部屋がエレベータの止まらない2階，5階にあれば－50万，余分の窓があれば，30万が加算される．このようにしてたとえば18号室の予測価格は885万＋5万－15万＝875万と計算される．これはたまたま実際の価格と一致する．数量化法1類はこのように基準変数†の値を要因の貢献度に応じて分解，再構成する方法である．これよりどの要因がどのくらいプラスにあるいはマイナスの要因として働いているかが理解できる．

表2 カテゴリーに与えられた数量

要因	カテゴリー	与えられる数値（単位万円）
階数	(F1) 1階	−10.0
	(F2) 2階	−55.0
	(F3) 3階	0.0
	(F4) 4階	5.0
	(F5) 5階	−40.0
	(F6) 6階	0.0
平面的位値	(C1) 1番目	0.0
	(C2) 2番目	−15.0
	(C3) 3番目	−15.0
	(C4) 4番目	0.0
	(C5) 5番目	0.0
余分の窓	(Wd)あり	30.0
	なし	0 0
庭	(Yd)あり	36.0
	なし	0
交互作用	(Int)	50.0
価格水準	(d)	885.0

4.2.2 数量化法2類

数量化法2類(quantification method II)は説明変数ばかりでなく，基準変量もカテゴリー変量†で与えられる場合の手法で，2組のカテゴリー変数が互いに他を最もよく予測するように同時に数量化される．したがって説明変数，基準変量がともにカテゴリー変量であるような正準相関分析†あるいは説明変数がカテゴリー変量で与えられる正準判別分析†の1種とも考えられる．一般に基準変量のカテゴリー数が2より大きい場合，2類で求められるカテゴリーの数量化は1組だけとは限らないが，1類は2類で基準変量の数量化がすでに1組だけ与えられている場合に相当するとも考えられる．

いま，基準変量に対応するダミー変数を$\boldsymbol{G}$，説明変数に対応するダミー変数を$\boldsymbol{H}$で表わし，$\boldsymbol{G}, \boldsymbol{H}$にかかる重みをそれぞれ$\boldsymbol{A}, \boldsymbol{B}$とすると，判別問題は$\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{G}\boldsymbol{A}=\boldsymbol{I}$, $\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}=\boldsymbol{I}$の制約条件のもとで$\mathrm{tr}(\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B})$を$\boldsymbol{A}$と$\boldsymbol{B}$について最大化する問題に帰着する．これより，次の一般化固有方程式が導かれる．

$$\boldsymbol{G}^t\boldsymbol{H}(\boldsymbol{H}^t\boldsymbol{H})^-\boldsymbol{H}^t\boldsymbol{G}\boldsymbol{A}-\boldsymbol{G}^t\boldsymbol{G}\boldsymbol{A}\boldsymbol{\Lambda}=\boldsymbol{O} \qquad (4)$$
$$\boldsymbol{H}^t\boldsymbol{G}(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}-\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}\boldsymbol{\Lambda}=\boldsymbol{O}$$

ここでは$\boldsymbol{\Lambda}$は固有値を対角要素とする対角行列を表わし，$(\boldsymbol{H}^t\boldsymbol{H})^-$は$\boldsymbol{H}^t\boldsymbol{H}$の任意の一般逆行列†を表わす．なおこれは説明変数が2変数以上のとき$\boldsymbol{H}^t\boldsymbol{H}$の逆行列がとれないためである．1類の場合と同様，各変数から最後のカテゴリーを取り除いてその代り定数項を含めて$\boldsymbol{H}$を定義すれば一般の逆行列が定義でき，それを$(\boldsymbol{H}^t\boldsymbol{H})^-$の代りに用いてもよい．最大固有値1に対応する固有ベクトルは定数ベクトルとなるので解から除外する必要がある．また$\boldsymbol{A}$と$\boldsymbol{B}$は次の関係を満たす．

$$\boldsymbol{A}=(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}\boldsymbol{\Lambda}^{-1} \qquad (5)$$

いまこれを多次元尺度法(特に多次元展開法)的に解釈すると次のようになる．判別群，観測個体(被験者)がともに多次元ユークリッド空間の点として表わされるものとする．判別群に対応する点の位置を座標$\boldsymbol{A}$で表わし，被験者に対応する点の座標を$\boldsymbol{Y}=\boldsymbol{H}\boldsymbol{B}$で表わす．被験者と判別群間の距離の2乗を要素とする行列(N人$\times g$グループ)を$\boldsymbol{D}^{(2)}$で表わすと

$$\boldsymbol{D}^{(2)}=\mathrm{Diag}(\boldsymbol{H}\boldsymbol{B}\boldsymbol{B}^t\boldsymbol{H}^t)\boldsymbol{1}_N\boldsymbol{1}_g^t-2\boldsymbol{H}\boldsymbol{B}\boldsymbol{A}^t+\boldsymbol{1}_N\boldsymbol{1}_g^t\mathrm{Diag}(\boldsymbol{A}\boldsymbol{A}^t)$$

が成り立つ．したがって，被験者とそれが属する判別群間の距離の2乗和は

$$\begin{aligned}\phi&=\mathrm{tr}(\boldsymbol{G}^t\boldsymbol{D}^{(2)})\\&=\mathrm{tr}(\boldsymbol{G}^t\,\mathrm{Diag}(\boldsymbol{H}\boldsymbol{B}\boldsymbol{B}^t\boldsymbol{H}^t)\boldsymbol{1}_N\boldsymbol{1}_g^t-2\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}\boldsymbol{A}^t\\&\quad+\boldsymbol{G}^t\boldsymbol{1}_N\boldsymbol{1}_g^t\,\mathrm{Diag}(\boldsymbol{A}\boldsymbol{A}^t))\\&=\mathrm{tr}(\boldsymbol{1}_g^t\boldsymbol{G}^t\,\mathrm{Diag}(\boldsymbol{H}\boldsymbol{B}\boldsymbol{B}^t\boldsymbol{H}^t)\boldsymbol{1}_N-2\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}\\&\quad+\boldsymbol{1}_g^t\,\mathrm{Diag}(\boldsymbol{A}\boldsymbol{A}^t)\boldsymbol{G}^t\boldsymbol{G}\boldsymbol{1}_g)\\&=\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B})-2\mathrm{tr}(\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B})\\&\quad+\mathrm{tr}(\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{G}\boldsymbol{A})\\&=\mathrm{tr}\{(\boldsymbol{H}\boldsymbol{B}-\boldsymbol{G}\boldsymbol{A})^t(\boldsymbol{H}\boldsymbol{B}-\boldsymbol{G}\boldsymbol{A})\}\end{aligned}$$

と表わされる．ϕを$\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}=\boldsymbol{I}$のもとで最小化する．まず，$\phi$を$\boldsymbol{A}$について最小化する．

$$-\frac{1}{2}\frac{\partial\phi}{\partial\boldsymbol{A}}=\boldsymbol{G}^t(\boldsymbol{H}\boldsymbol{B}-\boldsymbol{G}\boldsymbol{A})\equiv\boldsymbol{O}$$
$$\rightarrow\boldsymbol{A}=(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B} \qquad (6)$$

したがって

$$\begin{aligned}\phi^*&=\min_A\phi\\&=\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B})-\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{G}(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B})\\&=\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t(\boldsymbol{I}-\boldsymbol{G}(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t)\boldsymbol{H}\boldsymbol{B})\end{aligned}$$

が成り立つ．

ϕ^*を

$$\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}=\boldsymbol{I}$$

のもとで最小化することは，同じ制約条件のもとで

$$\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{G}(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B})$$

を最大化するのと同等である．これにより(4)式の固有値問題に還元される．なお(5)式と(6)式の違いは，(6)式では$\boldsymbol{A}^t\boldsymbol{G}^t\boldsymbol{G}\boldsymbol{A}=\boldsymbol{I}$の制約条件を置かないことによる．

$\boldsymbol{H}\boldsymbol{B}$の重心が原点となるよう$\boldsymbol{B}$を求めると，$\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}=\boldsymbol{I}$のもとでは$\mathrm{tr}(\boldsymbol{B}^t\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B})=r$ (rは

判別空間†の次元数)は被験者間の距離の総2乗和に相当するから，数量化法2類はそれを一定としたときに，被験者と判別群間の距離の2乗和が最小となるよう，被験者および判別群の布置を定める方法であるということができる．なお，以上の議論は$\boldsymbol{H}$がカテゴリカルであるという性質は用いていないから，一般の線形(正準)判別分析の場合にも成り立つものと考えられる．なお，数量化法2類の源泉は，フィッシャー(R.A.Fisher[1938])にあるといわれる．

4.2.3 数量化法3類

数量化法3類(quantification method III)は多変量カテゴリカル・データの構造分析のための手法である．1類や2類と異なり，外的基準を持たない．連続変量の場合でいえば，1類，2類が回帰分析，判別分析に相当するのに対し，3類は主成分分析に相当する．3類は外的基準によらず，「内的一貫性」の原理からカテゴリーの数量化を行う．

数量化法3類の歴史は古く，リチャードソンとクーダー(M.Richardson & G.F.Kuder[1933])の**交互平均法**(method of reciprocal averages)に由来する．その後，ガットマン(L.Guttman[1941])，マウング(K.Maung[1941])，C.Hayashi[1952]等により別の基準から同等の方法が導かれ，ベンゼクリ(J.P.Benzecri[1973])，S.Nishisato[1980]等によって洗練され現在に到っている．数量化法3類というのは日本での呼び名で，同じ方法が，外国では**双対尺度法**(dual scaling; S.Nishisato[1980]，西里静彦[1982])，**最適尺度法**(optimal scaling)，**コレスポンデンス・アナリシス**(Analyse des correspondences; J.P.Benzecri[1973])，**等質性分析**(homogeneity analysis; A.Gifi[1981])などと呼ばれている．

数量化法3類を導く内的一貫性の基準にはさまざまなものがあり，これまで次のような基準が用いられてきた．

i)分散比(個人間分散／全分散→最大)
(L.Guttman[1941], S.Nishisato[1980])
ii)交互平均法
(M.Richardson & G.F.Kuder[1933])
iii)2変量相関
(L.Guttman[1941], C.Hayashi[1952])
iv)正準相関(竹内啓・柳井晴夫[1972])
v)同時線形回帰(J.C.Lingoes[1964])
vi)一般化分散(斎藤堯幸・小川定暉[1972])
vii)等質性(A.Gifi[1981])
viii)展開法
(Y.Takane[1980a], W.J.Heiser[1981])

幸いなことにこれらの基準はすべて同一の解析法に帰着することが知られている．

これらの基準のうち分散比による方法が最も標準的であるが，ここではMDSとの関係から，展開法からの導出を試みる．3類は分割表†の解析にも用いられるが，まず多肢選択データの場合について説明する．多肢選択データを

$$f_{ki(j)}=1\ (\text{被験者 } k \text{ が項目 } i \text{ でカテゴリー } j \text{ を選択したとき})$$
$$=0\ (\text{その他のとき})$$

と定義する．展開法モデルに従い，被験者およびカテゴリーを多次元空間内の点として表わし，それぞれの座標を$\boldsymbol{Y},\boldsymbol{X}$で表わす．2類と違うのは$\boldsymbol{X},\boldsymbol{Y}$が$\boldsymbol{GA},\boldsymbol{HB}$のように基準変数，説明変数の1次結合でなければならないというような制約が存在しないことである．

$d_{ki(j)}$を被験者kと項目カテゴリー間のユークリッド距離とし，ϕを次のように定める

$$\phi=\sum_k\sum_{i,j}f_{ki(j)}d^2_{ki(j)}=\mathrm{tr}(\boldsymbol{F}^t\boldsymbol{D}^{(2)})$$

ここで$\boldsymbol{F},\boldsymbol{D}^{(2)}$はそれぞれ$f_{ki(j)}$, $d^2_{ki(j)}$を要素とする行列である．$\boldsymbol{F},\boldsymbol{D}^{(2)}$の要素はともに行が$N$人の被験者に対応し，列は変数のカテゴリーに対応するよう配列されているものとする．$\boldsymbol{D}_C$を$\boldsymbol{F}$の列和を対角要素とする行列とする．数量化法3類は$\boldsymbol{X}^t\boldsymbol{D}_C\boldsymbol{X}=\boldsymbol{I}$という条件のもとで$\phi$，すなわち被験者とその被験者に選ばれたカテゴリーの距離の2乗和を最小にする方法として導かれる．

$$\boldsymbol{D}^{(2)}=\mathrm{Diag}(\boldsymbol{YY}^t)\boldsymbol{1}_N\boldsymbol{1}_n^t-2\boldsymbol{YX}^t+\boldsymbol{1}_N\boldsymbol{1}_n^t\mathrm{Diag}(\boldsymbol{XX}^t)$$

と表わされるから

$$\phi=\mathrm{tr}(\boldsymbol{Y}^t\boldsymbol{D}_R\boldsymbol{Y})-2\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{F}^t\boldsymbol{Y})+\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{D}_C\boldsymbol{X})$$

ここで$\boldsymbol{D}_R$は$\boldsymbol{F}$の行和を対角要素とする行列である．ϕをまず$\boldsymbol{Y}$について最小化すると

$$-\frac{1}{2}\frac{\partial\phi}{\partial\boldsymbol{Y}}=\boldsymbol{FX}-\boldsymbol{D}_R\boldsymbol{Y}\equiv\boldsymbol{O}$$

よって$\boldsymbol{Y}=\boldsymbol{D}_R^{-1}\boldsymbol{FX}$．このとき

$$\phi^*=\min_Y\phi=-\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{F}^t\boldsymbol{D}_R^{-1}\boldsymbol{FX})+\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{D}_C\boldsymbol{X})$$

したがって，ϕ^*を$\boldsymbol{X}^t\boldsymbol{D}_C\boldsymbol{X}=\boldsymbol{I}$の条件のもとでさらに$\boldsymbol{X}$について最小化することは同じ条件のもとで$\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{F}^t\boldsymbol{D}_R\boldsymbol{FX})$を最大化することと同等である．この問題は次のような一般化固有値問題に帰着する．

$$\boldsymbol{F}^t\boldsymbol{D}_R^{-1}\boldsymbol{FX}-\boldsymbol{D}_C\boldsymbol{X\Lambda}=\boldsymbol{O} \tag{7}$$

以上の導出過程は数量化法2類の場合と酷似しているが，3類の場合ϕを$\mathrm{tr}\{(\boldsymbol{FX}-\boldsymbol{Y})^t(\boldsymbol{FX}-\boldsymbol{Y})\}$の形に書くことはできない．それは一般に$\boldsymbol{F}$が2つ以上の項目から成るため$\boldsymbol{F1}_n=\boldsymbol{1}_N$

を満たさないことによる．2類で $\mathrm{tr}(\boldsymbol{G}^t\boldsymbol{D}^{(2)})$を

$$\mathrm{tr}\{(\boldsymbol{GA}-\boldsymbol{HB})^t(\boldsymbol{GA}-\boldsymbol{HB})\}$$

と書けるのは $\boldsymbol{G}\boldsymbol{1}_g=\boldsymbol{1}_N$ を満たすからである．

なお ϕ を $\boldsymbol{Y}^t\boldsymbol{D}_C\boldsymbol{Y}=\boldsymbol{I}$ の条件のもとで最小化すると，$\boldsymbol{X}=\boldsymbol{D}_C^{-1}\boldsymbol{F}^t\boldsymbol{Y}$．よって

$$\phi^{**}=\min_x\phi=\mathrm{tr}(\boldsymbol{Y}^t\boldsymbol{D}_R\boldsymbol{Y})-\mathrm{tr}(\boldsymbol{Y}^t\boldsymbol{F}\boldsymbol{D}_C^{-1}\boldsymbol{F}^t\boldsymbol{Y})$$

これより

$$\boldsymbol{F}\boldsymbol{D}_C^{-1}\boldsymbol{F}^t\boldsymbol{Y}-\boldsymbol{D}_R\boldsymbol{Y}\boldsymbol{\Lambda}=\boldsymbol{O}$$

なる一般化固有方程式が導かれる．これは(7)の前から $\boldsymbol{F}\boldsymbol{D}_C^{-1}$ をかけ $\boldsymbol{D}_R^{-1}\boldsymbol{F}\boldsymbol{X}$ を $\boldsymbol{Y}$ で置き換えた式に一致する．したがって，どちらを解いても解が定数倍されるのを除いては，本質的な違いはない．

数量化法3類は多肢選択データで被験者を判別群としたとき，数量化法2類に一致することが知られている(竹内啓・柳井晴夫[1972])．たとえば $\boldsymbol{F}$ がそれぞれ2カテゴリー，3カテゴリーから成る2つの項目から成り，その最初の部分が次のようであったとする．

$$\boldsymbol{F}=\begin{bmatrix}1&0&0&1&0\\0&1&0&0&1\\0&1&1&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\end{bmatrix}$$

これより $\boldsymbol{G}$ および $\boldsymbol{H}$ を次のように定義する．

$$\boldsymbol{G}=\begin{bmatrix}1&0&0&\cdots\\1&0&0&\cdots\\0&1&0&\cdots\\0&1&0&\cdots\\0&0&1&\cdots\\0&0&1&\cdots\\\vdots&\vdots&\vdots&\cdots\end{bmatrix}\qquad\boldsymbol{H}=\begin{bmatrix}1&0&0&0&0\\0&0&0&1&0\\0&1&0&0&0\\0&0&0&0&1\\0&1&0&0&0\\0&0&1&0&0\\\vdots&\vdots&\vdots&\vdots&\vdots\end{bmatrix}$$

$\boldsymbol{G}$ は行に対応する反応がどの被験者によるものなのかを示している．第1列に1があるのはその反応が被験者1によるものであることを示している．上の例では最初の2つの反応が被験者1によるもの，次の2つが被験者2によるものであることがわかる．一方，$\boldsymbol{H}$ は所定の反応がどの項目のどのカテゴリーに属するかを示す．上の例では，1番目の反応は項目1のカテゴリー1に属し，2番目の反応は項目2のカテゴリー2に属することがわかる．

明らかに $\boldsymbol{F}=\boldsymbol{G}^t\boldsymbol{H}$ が成り立つ．いま $\boldsymbol{G}$ を基準変数，$\boldsymbol{H}$ を説明変数として数量化2類を適用すると(5)より

$$\boldsymbol{A}=(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}=\boldsymbol{D}_R^{-1}\boldsymbol{F}\boldsymbol{B}$$

$$\boldsymbol{H}^t\boldsymbol{G}(\boldsymbol{G}^t\boldsymbol{G})^{-1}\boldsymbol{G}^t\boldsymbol{H}\boldsymbol{B}-\boldsymbol{H}^t\boldsymbol{H}\boldsymbol{B}\boldsymbol{\Lambda}$$
$$=\boldsymbol{F}^t\boldsymbol{D}_R^{-1}\boldsymbol{F}\boldsymbol{B}-\boldsymbol{D}_R\boldsymbol{B}\boldsymbol{\Lambda}=\boldsymbol{O}$$

が得られる．これは $\boldsymbol{B}$ を $\boldsymbol{X}$，$\boldsymbol{A}$ を $\boldsymbol{Y}$ としたときの3類の式に一致する．なお以上の展開は正準相関分析にもとづいた3類の導出になっていることに注意．

すでに述べたように数量化法3類は多肢選択データだけでなく，2次元分割表(2 dimensional contingency table)の分析にも用いられる．分割表は $\boldsymbol{F}$ が $\boldsymbol{F}_1,\boldsymbol{F}_2$ の2項目から成るとして，$\boldsymbol{F}^*=\boldsymbol{F}_1^t\boldsymbol{F}_2$ により求められる．この場合，$\boldsymbol{F}^*$ を $\boldsymbol{F}$ のかわりに用いれば，これまでの議論がそのまま成り立つ．これは $\boldsymbol{F}_1$ と $\boldsymbol{F}_2$ で正準相関分析，また判別分析を行うことに相当する．なお多肢選択データとしての $\boldsymbol{F}$ とそれより導かれた $\boldsymbol{F}^*$ の3類による分析結果には特殊な関係が成り立つことが知られている(西里静彦[1982])．

最後に3類の適用例を1つ掲げておく(Y. Takane[1980a])．これはやや特殊な例であるが，分類法(sorting method)によるデータが1種の多肢選択データであること，したがって3類によって分析できることを示す点で興味深

図9　表3のデータを被験者1についてのみダミー変数表示したもの

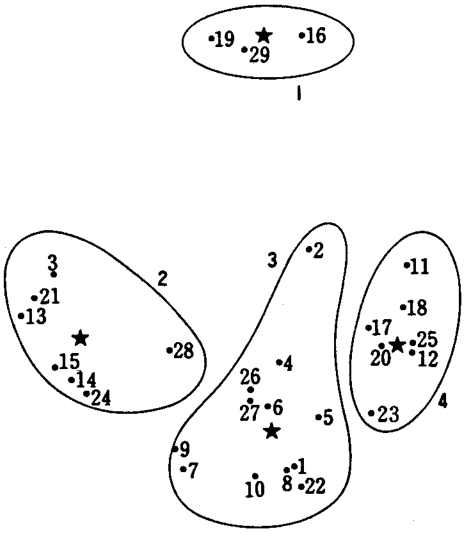
分類カテゴリー

$$\boldsymbol{F}_1=\begin{array}{c}\begin{array}{ccccccc}1&2&3&4&5&6&7\end{array}\\\left[\begin{array}{ccccccc}
1&0&0&0&0&0&0\\
0&1&0&0&0&0&0\\
0&0&1&0&0&0&0\\
0&1&0&0&0&0&0\\
0&0&0&0&1&0&0\\
1&0&0&0&0&0&0\\
0&0&0&0&1&0&0\\
0&0&0&0&1&0&0\\
0&0&0&0&1&0&0\\
0&0&0&0&1&0&0\\
0&0&0&0&0&1&0\\
0&0&0&0&0&1&0\\
0&0&1&0&0&0&0\\
0&0&1&0&0&0&0\\
0&0&1&0&0&0&0\\
0&0&0&0&0&1&0\\
0&0&0&0&0&1&0\\
0&0&0&0&0&1&0\\
0&1&0&0&0&0&0\\
0&0&0&0&0&1&0\\
0&0&1&0&0&0&0\\
1&0&0&0&0&0&0\\
0&0&0&0&0&0&1\\
0&0&1&0&0&0&0\\
0&0&0&0&0&1&0\\
0&1&0&0&0&0&0\\
1&0&0&0&0&0&0\\
0&0&0&0&0&1&0\\
0&1&0&0&0&0&0
\end{array}\right]\end{array}
\begin{array}{l}\\1\ \text{刺}\\2\\3\ \text{激}\\4\\5\\6\\7\\8\\9\\10\\11\\12\\13\\14\\15\\16\\17\\18\\19\\20\\21\\22\\23\\24\\25\\26\\27\\28\\29\end{array}$$

表3　分類されたグループ・ナンバーによって表示された分類データ

		刺	激																										
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
被験者	1	1	2	3	2	4	1	5	5	5	5	6	6	3	3	3	6	6	6	2	6	3	1	7	3	6	2	**1**	6	2
	2	1	2	1	2	3	3	4	2	4	5	5	3	1	5	5	1	3	2	1	3	5	5	5	4	3	2	**2**	1	1
	3	1	2	3	2	4	2	2	1	2	1	5	6	7	7	7	5	2	2	3	6	3	1	1	2	6	6	**6**	7	3
	4	1	2	2	2	3	4	4	3	4	4	5	5	4	3	4	6	2	5	6	2	4	1	1	4	5	4	**4**	4	6
	5	1	2	3	4	5	6	6	7	6	7	8	8	3	3	3	9	9	8	2	9	3	1	8	3	8	4	**2**	1	2
	6	1	1	2	1	1	1	1	1	1	1	3	3	2	2	2	4	3	3	4	3	2	1	3	2	3	1	**1**	2	4
	7	1	2	3	2	3	2	4	4	5	1	6	1	5	3	5	6	2	6	5	1	5	1	2	4	2	4	**4**	3	6
	8	1	2	3	4	5	4	4	1	4	4	6	5	3	4	4	2	5	6	2	1	3	1	5	4	6	4	**4**	5	2
	9	1	2	3	4	2	5	6	7	6	1	2	8	9	9	9	10	2	7	10	8	3	1	7	9	2	11	**11**	4	5
	10	1	2	3	4	5	6	7	6	7	6	8	4	3	9	9	10	4	8	10	2	3	1	5	9	6	11	**11**	9	10

図10　図8のデータの3類による分析結果



1	Accept	16	Lack
2	Beg	17	Lend
3	Belong	18	Lose
4	Borrow	19	Need
5	Bring	20	Offer
6	Buy	21	Own
7	Earn	21	Receive
8	Find	23	Return
9	Gain	24	Save
10	Get	25	Sell
11	Get rid of	26	Steal
12	Give	27	Take
13	Have	28	Use
14	Hold	29	Want
15	Keep		

い．データは図10の下に掲げる29の HAVE に関係した言葉を10人の被験者に類似性に応じていくつかのグループに分類してもらった結果である．表3の行は10人の被験者に対応し，列は29の言葉に対応する．表中の数字は各被験者がそれぞれの言葉をどのグループに分類したかを示している．ただし，どのグループを何番目と呼ぶかは任意である．図9は表3のデータをダミー変数を用いて多肢選択データの形に書き換えたものである．紙幅の都合上，被験者1のデータのみダミー変数化した結果を示す．行は29の刺激に対応し，列は被験者によって作られた分類グループに対応する．これは特定の行に対応した刺激がどの列に対応したグループに分類されたかを 1-0 のパターンで示したものである．これを $\boldsymbol{F}_1$ とし，$\boldsymbol{F}=[\boldsymbol{F}_1,\cdots,\boldsymbol{F}_N]$ と定義する（N 人の被験者のダミー変数化したデータを横に並べた行列）．この $\boldsymbol{F}$ を数量化法3類にかけると，図10の結果が得られる．

図10では主として刺激に対応した点が示されている．各被験者の分類グループに対応する点は，各被験者によって特定のグループに分類された刺激の重心で与えられる．図10では10人のうち1人の被験者（被験者6）によって形成された4つの分類グループが丸で囲まれていると同時にその重心が星印で示されている．図10を見ると，左のほうに Have, Own, Belong などが位置づけられ，逆に左のほうには Give, Sell, Lose などが布置されている．したがって水平方向を安定した所有状態と近い未来において失われる可能性の強い所有状態を対照する軸と考えることができる．また上のほうには Lack, Need, Want が位置し，下のほうには，Receive, Get, Gain などが位置している．したがって縦方向を安定した非所有状態と近い将来変化する可能性の強い非所有状態を対照する軸と考えることができる．なおこの場合，右上から左下にかけて近い将来における非所有と所有，左上から右下にかけて現在における非所有と所有の状態が区別されているとも解釈できる．

なお，分類データを処理する場合，通常のMDSでは2つの刺激が同じグループに分類された頻度を数え，その頻度をそれらの刺激間の類似性の指標として用いることが多いが，これは $\sum_{k=1}^{N} F_k F_k^t$ を類似性行列とみなすことに相当する．これに対し数量化法3類は，類似性行列として

$$\sum_{k=1}^{N} F_k (F_k^t F_k)^{-1} F_k^t$$

をとる．これは $F_k F_k^t$ を被験者についてたし合わせる前に分類グループの大きさで割ることを意味する．すなわち，2つの刺激が同じグループに分類された場合，グループの大きさによって類似度への貢献の度合が変わってくる．グループが小さいほど，そのグループ内の刺激の類似性は大きいとみなされる．

4.2.4 数量化法4類

数量化法4類(quantification method IV)は n 個の対象間の類似性データが与えられたとき，ある基準から対象の多次元布置を求める手法で，MDS†の基本的考えに最も似ている．e_{ij} を対象 i と j の類似性を表わす指標とし，e_{ij} を要素とする行列を $\boldsymbol{E}$ で表わすと，その基準は

$$\phi = \sum_{i<j} e_{ij} d_{ij}^2 = \frac{1}{2}\mathrm{tr}(\boldsymbol{E}\boldsymbol{D}^{(2)}) \tag{8}$$

と表わされる．4類はこの ϕ を対象間の距離の総2乗和が一定($\sum_{i<j} d_{ij}^2 = c$)という制約のもとで最小化する $\boldsymbol{X}$ を求める．ϕ を最小化するということは大きい e_{ij} には(すなわち i と j の類似度が大きいときは)小さい d_{ij}^2 が，小さい e_{ij} には大きな d_{ij}^2 が対応するよう対象の布置を定めることを意味する．(8)式の ϕ は形式的に3類の最小化基準と一致する．ただし，3類の場合の $\boldsymbol{D}^{(2)}$ が通常被験者とカテゴリー間の距離の2乗を要素とする $N\times n$ の行列であるのに対し，4類の場合は n 個の対象間の距離の2乗を要素とする正方行列であるところが違っている．

$\boldsymbol{S}$ を対比較の計画(デザイン)行列†とする．たとえば $n=4$ の場合，$\boldsymbol{S}$ は次のような行列を表わす．

$$S=\begin{bmatrix} 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

一般に $\boldsymbol{S}^t\boldsymbol{S} = n\boldsymbol{J}_n$ が成り立つ($\boldsymbol{J}_n$ は中心化行列を表わす)．いま $\boldsymbol{D}_{-e}$ を $-e_{ij}$ を対角要素とする $n(n-1)/2$ 次の行列とすると

$$\phi^* \equiv -\phi = \mathrm{tr}(\boldsymbol{D}_{-e}\,\mathrm{Diag}(\boldsymbol{S}\boldsymbol{X}\boldsymbol{X}^t\boldsymbol{S}^t)) = \mathrm{tr}(\boldsymbol{X}^t\boldsymbol{S}^t\boldsymbol{D}_{-e}\boldsymbol{S}\boldsymbol{X})$$

また

$$\sum_{i<j} d_{ij}^2 = \mathrm{tr}(\mathrm{Diag}(\boldsymbol{S}\boldsymbol{X}\boldsymbol{X}^t\boldsymbol{S}^t)) = n\,\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{J}_n\boldsymbol{X}) = n\,\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{X})$$

したがって ϕ^* を $\sum_{i<j} d_{ij}^2=$ 一定のもとで最小化することは ϕ を $\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{X})=$ 一定のもとで最大化することと同等である．制約条件を少し強くとり，$\boldsymbol{X}^t\boldsymbol{X}=\boldsymbol{I}$ とすると(このとき $\mathrm{tr}(\boldsymbol{X}^t\boldsymbol{X}) = r$; r は空間の次元数)，この問題は次の固有値問題に帰着する．

$$\boldsymbol{S}^t\boldsymbol{D}_{-e}\boldsymbol{S}\boldsymbol{X} - \boldsymbol{X}\boldsymbol{\Lambda} = \boldsymbol{O}$$

したがって $\boldsymbol{S}^t\boldsymbol{D}_{-e}\boldsymbol{S}$ の大きいほうから r 個の固有ベクトルに対決する固有ベクトルを $\boldsymbol{X}$ とすればよい．$\phi^* = \mathrm{tr}(\boldsymbol{\Lambda})$ となるからである．ただし r 個の固有値の中に0が入る場合はそれを避けるため $-e_{ij}$ に適当な正定数を加え $\boldsymbol{D}_{-e}^*$ を定義し，$\boldsymbol{S}^t\boldsymbol{D}_{-e}^*\boldsymbol{S}$ のスペクトル分解を求めるとよい．$\boldsymbol{S}^t\boldsymbol{D}_{-e}^*\boldsymbol{S}$ は $\boldsymbol{S}^t\boldsymbol{D}_{-e}\boldsymbol{S}$ と同じ固有ベクトルを持ち，固有値はもともと0の固有値を除いて一定数だけ正の方にずれるという性質を持つ(Y. Takane[1977])．したがって，$-e_{ij}$ に加える正定数は $\boldsymbol{S}^t\boldsymbol{D}_{-e}^*\boldsymbol{S}$ の固有値がもともと0の固有値以外はすべて正になるよう選ぶとよい．

数量化法4類はデータが類似性データで簡単に距離と線形関係を持つような非類似性データに変換できないとき，非計量MDSの近似解としてよく用いられる．ただし，データの最適単調変換を求めることはしないので，計算が簡単なのと引き換えに，解にゆがみが生じることがあるので注意しなければならない．

最後に，数量化法全般の参考文献を掲げておく．林知己夫・樋口伊佐夫[1970]，駒沢勉[1978]，西里静彦[1982]．このうち駒沢[1978]には数量化法のプログラムが掲載されている．岩坪秀一[1987]は数量化法の数学的基礎を整理してまとめたものである．西里[1982]には簡単な例題が豊富に収められている．