

主座標分析のS—P表への応用

平井安久

1. はじめに

n 個体の v 変量について観測されたデータを、個体と個体の間の関係をもとにして、多変量データを低い次元で要約する統計的手法の中に主座標分析法がある。これは Gower³⁾によって始められたもので、いくつかの例が報告されている^{2,4)}。特にデータの値が1と0のみの場合にも適用できるため、S—P表(佐藤⁶⁾)に用いられるような教育データへの応用を考えてみたい。

S—P表の考えは、 n 個体の v 項目のテスト問題に対する得点を1, 0のデータで表わし、授業分析や生徒と問題の関連の分析に役立てようというものである。ただ、その処理結果がデータの相互依存性から得られるものである点では主座標分析の考えと共通の性質をもつものである。したがって個体間の関係に注目すれば、各個体のデータを低次元でプロットする処理を行ない、その互いの位置関係を知ることで個体どうしの関連を把握できると考えられる。以下いくつかのS—P表のデータに主座標分析を応用してみることにする。

2. 主座標分析

主座標分析の計算手順は以下の通りである。

- (1) n 個体についての v 変量の観測値が与えられたとする。個体 i と個体 j の関係を表わす数 m_{ij} を求める。(この m_{ij} は一般にユークリッド距離, 相関係数, 一致係数などで表わす。)
- (2) m_{ij} から $a_{ij} = m_{ij} - \bar{m}_i - \bar{m}_j + \bar{m}$ への変換を行なう。ただし \bar{m}_i は行平均で \bar{m} は全平均である。 $M = [m_{ij}]$, $\alpha = [a_{ij}]$ とおくとこの変換は M から $\alpha = (I - N)M(I - N)$ への変換である。ただし I は単位行列, N は全要素が $1/n$ の $n \times n$ 行列とする。
- (3) 行列 α の固有値, 固有ベクトルを求める。ここで固有値 λ_k について, $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ とし, 対応する固有ベクトルの要素を $[a_{1k}, a_{2k}, \dots, a_{nk}]'$ とし, それらの2乗和が $a_{1k}^2 + a_{2k}^2 + \dots + a_{nk}^2 = \lambda_k$ となるように標準化をする。
- (4) t 次元での点 P_i の座標は $(a_{i1}, a_{i2}, \dots, a_{it})$ で与えられる。

	λ_1	λ_2	\dots	λ_n
点 P_1	a_{11}	a_{12}	\dots	a_{1n}
点 P_2	a_{21}	a_{22}	\dots	a_{2n}
\vdots	\vdots	\vdots	\vdots	\vdots
点 P_n	a_{n1}	a_{n2}	\dots	a_{nn}

ここで点 P_i と P_j のユークリッド距離を $d(P_i, P_j)$ とすると, $t = n$ のとき

$$\{d(P_i, P_j)\}^2 = m_{ii} + m_{jj} - 2m_{ij} \quad \dots (1)$$

という性質をもつ。ただし $t < n$ のときは

$$\{d(P_i, P_j)\}^2 \approx m_{ii} + m_{jj} - 2m_{ij} \quad \dots (2)$$

となる。

以上が主座標を求める手順の概略であるが、各個体についての観測値が1または0で与えられる場合には、第4節で述べるように一致係数を用いて個体間の関係を表わすことができる。

3. S-P表

佐藤⁶⁾によるS-P表のもとになるデータは、テスト問題の各項目に対して各生徒が正答であるか誤答であるかを1, 0で表わしn人の生徒のv個の項目に対する正答, 誤答をn×vの1, 0行列に表わしたものである。正答には1, 誤答には0を与えると, n行v列の得点表が得られる。第i行の要素は番号iの生徒の結果であり, 第j列の要素は番号jの問題についての各生徒の得点の様子を表わしている。この得点表の行と列を並べかえて, 上の行に合計得点の多い生徒が来て左の列に正答者数の多い問題が来るようにする。佐藤⁶⁾によるS-P表の例を表1に示す。(ただし, 次節以降での比較の都合上成績の上位の生徒から順に番号をつけなおしてある。) S曲線とは各生徒の合計得点の分だけ左端から距離をとり区切を入れ全生徒についての区切線をつないだもの。同様の方法で各問題について上端からの距離をとることによってP曲線が得られる。なお合計得点の同じ生徒, 正答者数が同じ問題については, 下のような注意係数CS_i, CP_jの小さい順に並べる。

表1. S-P表の例
(問題)

	2	3	7	4	9	1	6	5	10	8	CS _i
1.	1	1	1	1	1	1	1	1	1	1	10テン 0.000
2.	1	1	1	1	1	1	1	1	1	0	9テン 0.000
3.	1	1	1	1	1	0	1	1	0	1	8テン 0.571
4.	1	1	1	1	1	1	1	0	0	0	7テン 0.000
5.	1	1	0	1	1	0	1	1	0	0	6テン 0.500
6.	1	1	1	0	1	0	1	0	1	0	6テン 0.500
7.	1	1	1	1	0	1	0	0	0	0	5テン 0.000
8.	1	1	1	0	1	0	0	0	1	0	5テン 0.400
9.	1	1	0	1	0	0	1	0	0	1	5テン 0.700
10.	1	0	0	1	0	1	0	1	1	0	5テン 1.000
11.	0	1	1	1	0	1	0	0	0	0	4テン 0.400
12.	1	0	0	0	0	1	0	1	0	1	4テン 1.200
13.	1	0	1	0	1	0	0	0	0	0	3テン 0.333
14.	0	0	1	0	0	1	0	0	0	0	2テン 0.714
15.	0	1	0	0	0	0	0	0	0	0	1テン 0.250
	12	11	10	9	8	8	7	6	5	4	
CP _j	0.1000	0.3529	0.5526	0.1538	0.1500	0.7500	0.0000	0.2857	0.3750	0.5526	

佐藤⁶⁾より

注意係数 (生徒……CS_i, 問題……CP_j)

$$CS_i = \frac{\sum_{j=1}^{v_i} (1-y_{ij})y_{.j} - \sum_{j=\sum_{i'=1}^{i-1} v_{i'}+1}^v y_{i'j} \times y_{.j}}{\sum_{j=1}^{v_i} y_{.j} - y_{.i} \times \mu'}$$

$$CP_j = \frac{\sum_{i=1}^{v_j} (1-y_{ij})y_{.i} - \sum_{i=\sum_{j'=1}^{j-1} v_{j'}+1}^n y_{ij} \times y_{.i}}{\sum_{i=1}^{v_j} y_{.i} - y_{.j} \times \mu}$$

- y_{ij}: i行目 j列目のデータ
- y_{.i}: 生徒 i の合計得点
- y_{.j}: 問題 j の正答者数
- μ: 平均得点
- μ': 平均正答者数

ここで, 注意係数とは, 「ある変量のパターンの注意係数とは, その変量の区切線を境に1と0の入れかわりのない完全パターンを想定して, これを基準にしてその変量のパター

ンが完全パターンからどの程度差異があるかを示す値である。」⁶⁾

4. 主座標分析のS-P表への応用

S-P表において1, 0で表わされているn人の生徒のv個の小項目に対するデータをそのまま1, 0のデータによるn個体, v変量のデータとして主座標分析を行なう。まず

個体*i*と個体*j*の関係 m_{ij} を求める。個体*i*と個体*j*の間の一致する個数をそれぞれ $a_{iu}, b_{iu}, c_{iu}, d_{iu}$, ($a_{iu} + b_{iu} + c_{iu} + d_{iu} = v$) とする(表2)。ここで $m_{ij} = a_{iu} + d_{ju}$, つまり1と1および0と0の一致の個数とし、一致係数とよぶことにする。(ただし、ここで $m_{ij} = (a_{iu} + d_{ju})/v$ を一致係数としても、結果的に主座標間の相対的位置関係に影響はない。) 第2節の計算手順に従うと点 P_i の座標が得られる。さらに点 P_i と P_j 間のユーク

リッド距離 $b(P_i, P_j)$ については $i < j$ のとき(2)より

$$\begin{aligned} \{d(P_i, P_j)\}^2 &\approx m_{ii} + m_{jj} - 2m_{ij} \\ &= a_{iu} + d_{iu} + a_{ju} + d_{ju} - 2(a_{iu} + d_{ju}) \\ &= 2v - 2m_{ij} \end{aligned}$$

となることより

$$m_{ij} \approx v - \{d(P_i, P_j)\}^2 / 2 \quad \dots\dots (3)$$

となり、 $d(P_i, P_j)$ から一致係数 m_{ij} が近似的に再現される。

4.1 完全S-P表の場合

注意係数 CS_i, CP_j を求める際には実際の反応パターンが、完全反応パターンからどれ程離れているかが問題になるが、この完全反応パターンの場合にはS曲線とP曲線が全く一致して、1と0の逆転がまったく起きない。この完全S-P表を表1にある一般S-P表に対して考える。各生徒の合計得点 $\sum_{j=1}^{10} y_{ij}$ ($i = 1, 2, \dots, 15$) が表1

のS-P表の場合と等しくなるようにすれば、表3のような完全S-P表を得る。

この場合、個体*i*と個体(*i*+1)の合計得点は同じまたは近い値である。つまり一致係数 $m_{i,i+1} = a_{i,i+1} + d_{i,i+1}$ の値が大きい(あるいは $10 - m_{i,i+1} = b_{i,i+1} + c_{i,i+1}$ が小さい)。各個体間について m_{ij} を求めると表4を得る。これより α に変換して、その固有値(最大3個)、固有ベクトルを求めて表5を得る。

第1主座標の分散(寄与率)は54.07%、第3成分までの累積寄与率は83.3%である。第*j*固有値の固有ベクトルの第*i*要素が点 P_i の第*j*要素となるから、取り出

表3 完全S-P表

	2	3	7	4	9	1	6	5	10	8	点	CS_i
1.	1	1	1	1	1	1	1	1	1	1	10	0
2.	1	1	1	1	1	1	1	1	1	0	9	0
3.	1	1	1	1	1	1	1	1	0	0	8	0
4.	1	1	1	1	1	1	1	0	0	0	7	0
5.	1	1	1	1	1	1	0	0	0	0	6	0
6.	1	1	1	1	1	1	0	0	0	0	6	0
7.	1	1	1	1	1	0	0	0	0	0	5	0
8.	1	1	1	1	1	0	0	0	0	0	5	0
9.	1	1	1	1	1	0	0	0	0	0	5	0
10.	1	1	1	1	1	0	0	0	0	0	5	0
11.	1	1	1	1	0	0	0	0	0	0	4	0
12.	1	1	1	1	0	0	0	0	0	0	4	0
13.	1	1	1	0	0	0	0	0	0	0	3	0
14.	1	1	0	0	0	0	0	0	0	0	2	0
15.	1	0	0	0	0	0	0	0	0	0	1	0

した3つの主座標Ⅰ, Ⅱ, Ⅲのうち, ⅠとⅡからなる平面上に P_i をプロットすると図1が得られる。さらにⅠとⅢについてのプロット, ⅡとⅢについてのプロットをした図をそれぞれ図2, 図3に示す。それぞれの特徴をさらにはっきりさせるためにⅠ, Ⅱ, Ⅲの主座標からなる立体プロットをしてみると図4が得られる。各個体を表わす点 P_i ($i=1, \dots, 15$) は, 図中の破線に沿って位置していることがわかる。もとのデータ(表3)と比べると, 各個体の合計得点 $\sum_{j=1}^{10} y_{ij}$ の大小がⅠ軸の方向に表われていることがわかる。

表4 個体間的一致係数

		(個体番号)														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
(個体番号)	1	10	9	8	7	6	6	5	5	5	5	4	4	3	2	1
	2		10	9	8	7	7	6	6	6	6	5	5	4	3	2
	3			10	9	8	8	7	7	7	7	6	6	5	4	3
	4				10	9	9	8	8	8	8	7	7	6	5	4
	5					10	10	9	9	9	9	8	8	7	6	5
	6						10	9	9	9	9	8	8	7	6	5
	7							10	10	10	10	9	9	8	7	6
	8								10	10	10	9	9	8	7	6
	9									10	10	9	9	8	7	6
	10										10	9	9	8	7	6
	11											10	10	9	8	7
	12												10	9	8	7
	13													10	9	8
	14														10	9
	15															10

個体間の位置関係については, 同得点の個体は同じ位置に表わされ, 合計得点の近い

表5 行列 α の固有値および固有ベクトル

主座標	Ⅰ	Ⅱ	Ⅲ
固有値 (累積寄与率)	21.6286 (0.5407)	8.4326 (0.7515)	3.2702 (0.8333)
1	2.1435	1.0150	-0.6867
2	1.9453	0.7742	-0.2667
3	1.5672	0.3499	0.3164
4	1.0442	-0.1574	0.7060
5	0.4246	-0.6274	0.6639
6	0.4246	-0.6274	0.6639
7	-0.2735	-0.7998	-0.1903
8	-0.2735	-0.7998	-0.1903
9	-0.2735	-0.7998	-0.1903
10	-0.2735	-0.7998	-0.1903
11	-0.8704	-0.2134	-0.5789
12	-0.8704	-0.2134	-0.5789
13	-1.3064	0.4742	-0.2594
14	-1.6215	1.0494	0.2187
15	-1.7868	1.3757	0.5631
要素2乗加	21.6286	8.4326	3.2702

個体 i と個体 $i+1$ は近い位置にプロットされている。つまり一致係数 $m_{i+1} = a_{i+1} + d_{i+1}$ の値が大きい($10 - m_{i+1} = b_{i+1} + c_{i+1}$ が小さい)個体どうしが近い位置に存在する。

さらに立体プロットにおける一致係数 m_{ij} の $d(P_i, P_j)$ からの再現性についてしらべる。(3)式より, $v=10, t=3$ のときの $v - \{d(P_i, P_j)\}^2/2$ の値を求めて表6に示す。これを表4の m_{ij} の値と比べると, かなりよい近似が見られる。その比 $m_{ij}/[v - \{d(P_i, P_j)\}^2/2]$ の平均値は0.934, 分散は0.003である。全体的に一致係数の高い個体どうしは近くにプロットされていることがわかる。

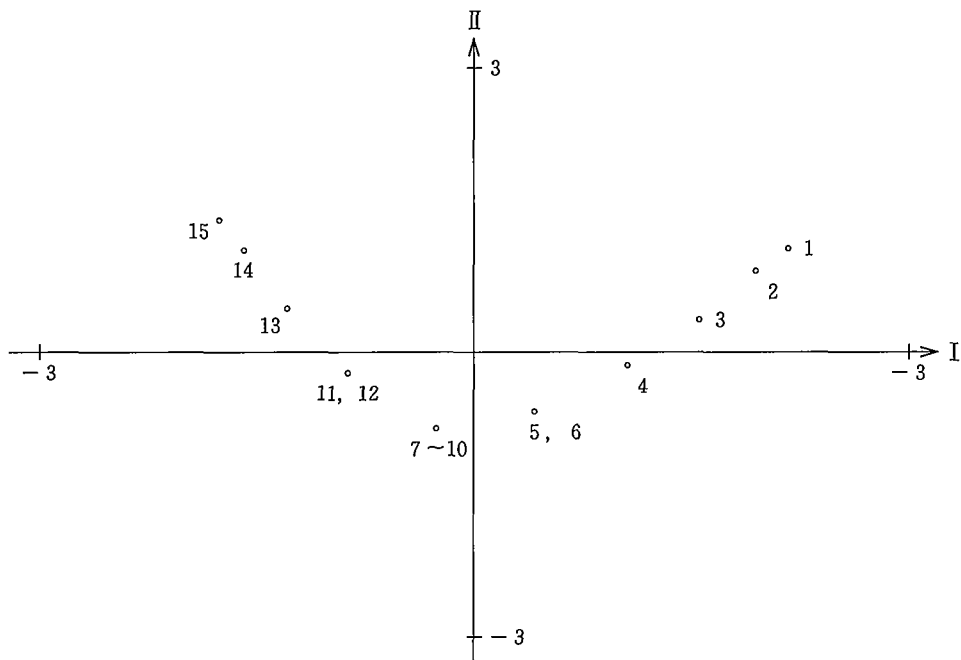


図1 完全S-P表のデータ(表3)に主座標分析を行なったときのI, II主座標に関するプロット

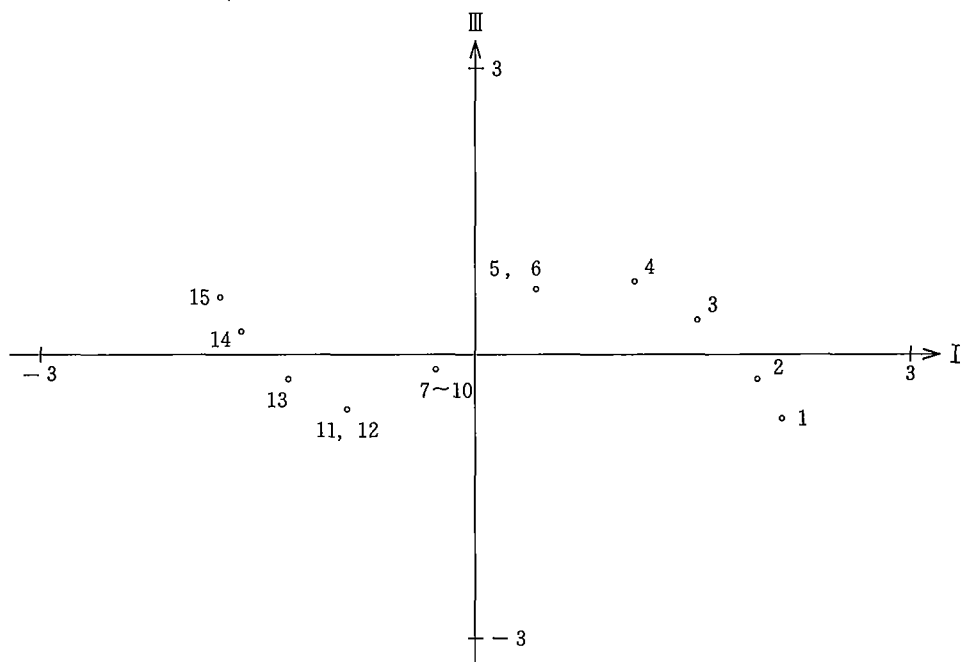


図2 完全S-P表のデータ(表3)に主座標分析を行なったときのI, III主座標に関するプロット

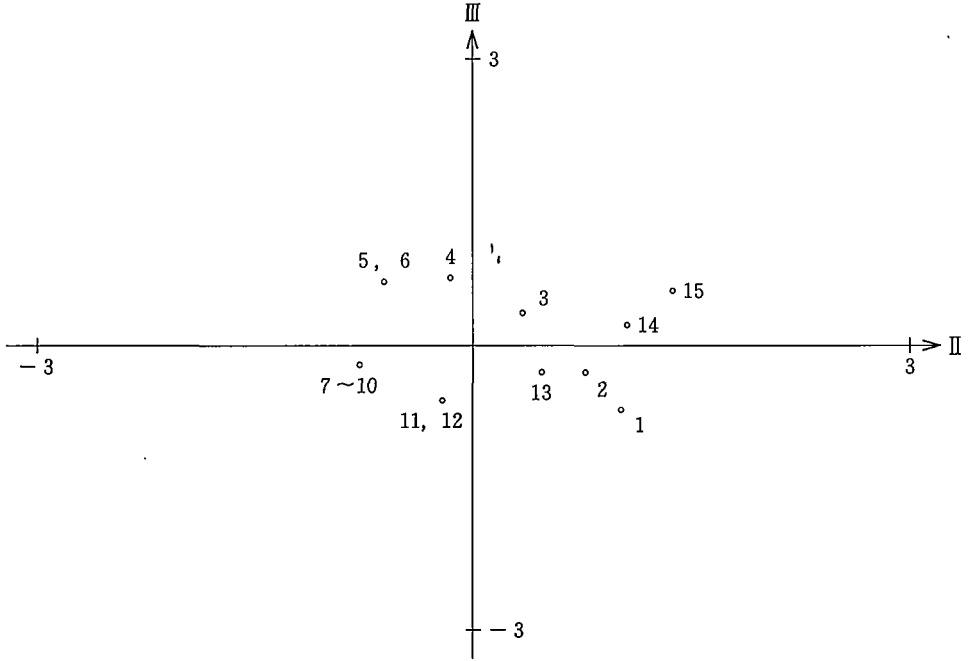


図3 完全S-P表のデータ(表3)に主座標分析を行なったときのII, III主座標に関するプロット

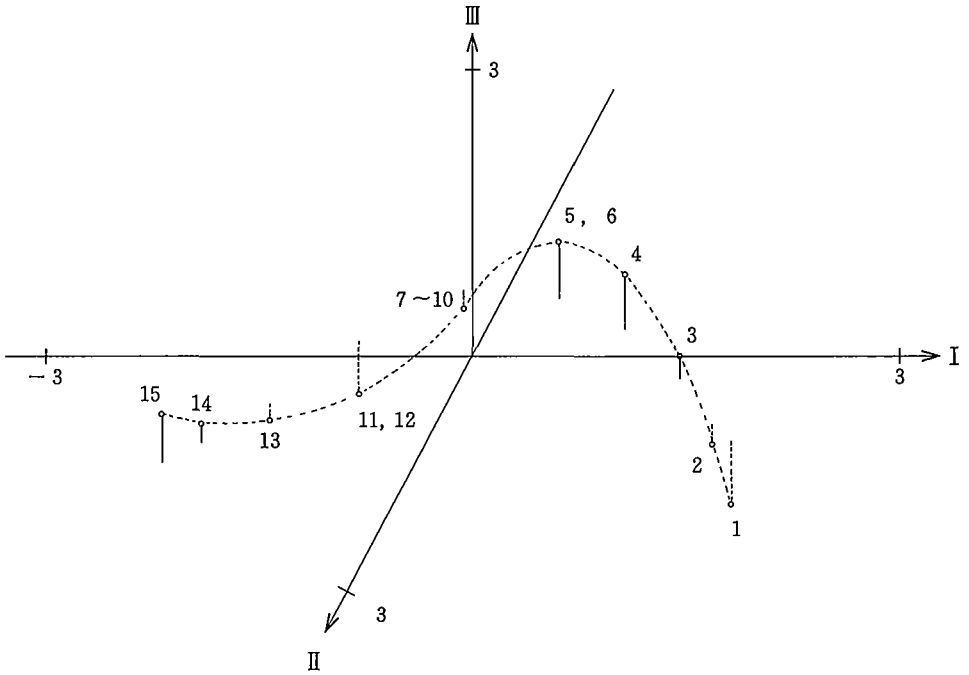


図4 完全S-P表のデータ(表3)に主座標分析を行なったときの立体プロット

表6 立体プロット(図3)での $10 - |d(P_i, P_j)|/2$ の値

	(個体番号)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	10.0000	9.8632	9.1096	7.7386	6.2619	6.2619	5.3092	5.3092	5.3092	5.3092	4.6980	4.6980	3.8117	2.5017	1.4304
2		10.0000	9.6685	8.6869	7.4284	7.4284	6.2968	6.2968	6.2968	6.2968	5.4995	5.4995	4.6683	3.4832	2.5107
3			10.0000	9.6586	8.8093	8.8093	7.5167	7.5167	7.5167	7.5167	6.4696	6.4696	5.6978	4.6666	3.8189
4				10.0000	9.6967	9.6967	8.5239	8.5239	8.5239	8.5239	7.3400	7.3400	6.5719	5.6000	4.8074
5					10.0000	10.0000	9.3767	9.3767	9.3767	9.3767	8.3035	8.3035	7.4687	6.4016	5.5436
6						10.0000	9.3767	9.3767	9.3767	9.3767	8.3035	8.3035	7.4687	6.4016	5.5436
7							10.0000	10.0000	10.0000	10.0000	9.5744	9.5744	8.6526	7.2979	6.2048
8								10.0000	10.0000	10.0000	9.5744	9.5744	8.6526	7.2979	6.2048
9									10.0000	10.0000	9.5744	9.5744	8.6526	7.2979	6.2048
10										10.0000	9.5744	9.5744	8.6526	7.2979	6.2048
11											10.0000	10.0000	9.6175	8.6025	7.6655
12												10.0000	9.6175	8.6025	7.6655
13													10.0000	9.6706	9.1400
14														10.0000	9.8738
15															10.0000

4.2 一般のS-P表の場合

表1のS-P表のデータは一般の場合であるから、このデータに対して主座標分析を行なう。第4.1節での完全S-P表と、各個体の合計得点 $\sum_{j=1}^{10} y_{ij}$ は同じであるが、各問題の正答者数 $\sum_{i=1}^{15} y_{ij}$ は異なり、各個体の注意係数 CS_i も多くは0以上の値となる。一致係数

表7 一般S-P表(表1)の一致係数

	(個体番号)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	10	9	8	7	6	6	5	5	5	5	4	4	3	2	1
2		10	7	8	7	7	6	6	4	6	5	3	4	3	2
3			10	7	6	8	5	5	7	3	4	4	5	2	3
4				10	7	7	8	6	6	4	7	3	6	5	4
5					10	6	5	9	5	3	4	2	7	4	5
6						10	5	5	7	5	4	4	5	2	5
7							10	6	6	6	9	5	6	7	6
8								10	4	4	5	3	8	5	6
9									10	4	5	5	4	3	6
10										10	5	7	4	5	4
11											10	4	5	8	7
12												10	5	6	5
13													10	7	6
14														10	7
15															10

表8 一般S-P表(表1)における行列 α の固有値および固有ベクトル

主座標	I	II	III
固有値 (累積寄与率)	20.5085 (0.3076)	13.0458 (0.5033)	10.0023 (0.6533)
1	-1.7611	-0.6465	-0.6907
2	-1.5100	-0.1684	-1.1261
3	-1.5437	-0.1979	0.7488
4	-0.7319	0.4108	-0.6272
5	-0.7472	1.5973	0.1163
6	-1.1259	-0.2860	1.0414
7	0.6688	-0.1357	-0.7683
8	0.0052	1.5796	-0.0874
9	-0.2957	-0.5938	1.5565
10	0.4744	-1.5807	-0.6219
11	1.1343	-0.0527	-0.8515
12	1.0258	-1.6802	0.5631
13	0.7673	1.1602	0.3117
14	2.0088	0.1460	-0.6665
15	1.6308	0.4479	1.1018
要素2乗和	20.5085	13.0458	10.0023

m_{ij} を求めると表7が得られ、行列 α の固有値、固有ベクトルを表8に示す。第2固有値までの寄与率は50.33%、第3固有値までの寄与率が65.33%であるから完全S-P表の場合より低い。I、II主座標によるプロット、およびI、II、III主座標による立体プロットを図5、図6に図示する。個体7、8、9、10の合計得点はいずれも5点であるが、正答した問題の組が異なるため互いに離れた位置関係でプロッ

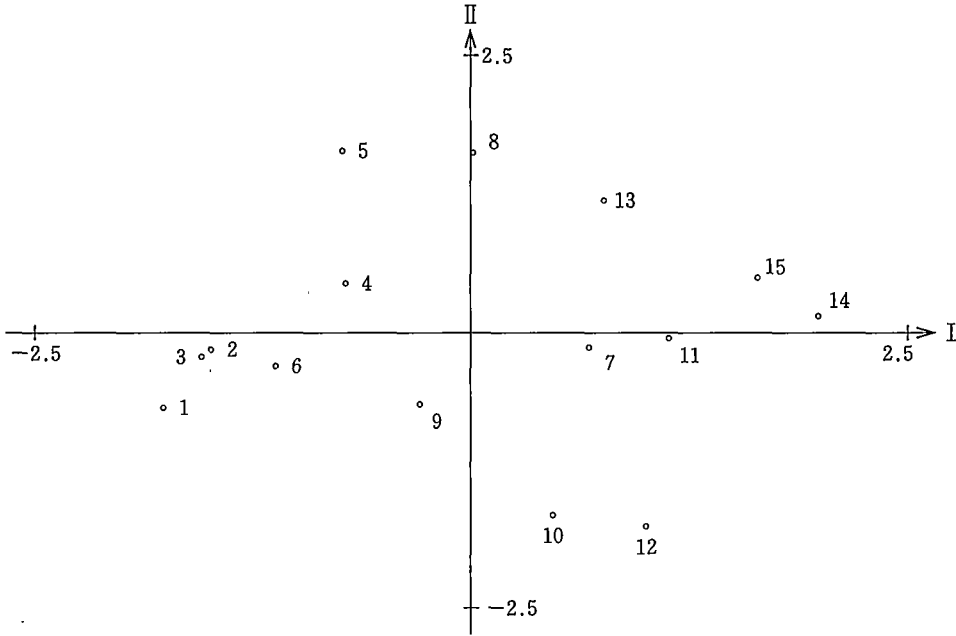


図5 一般S-P表(表1)のデータに対するI, II座標プロット

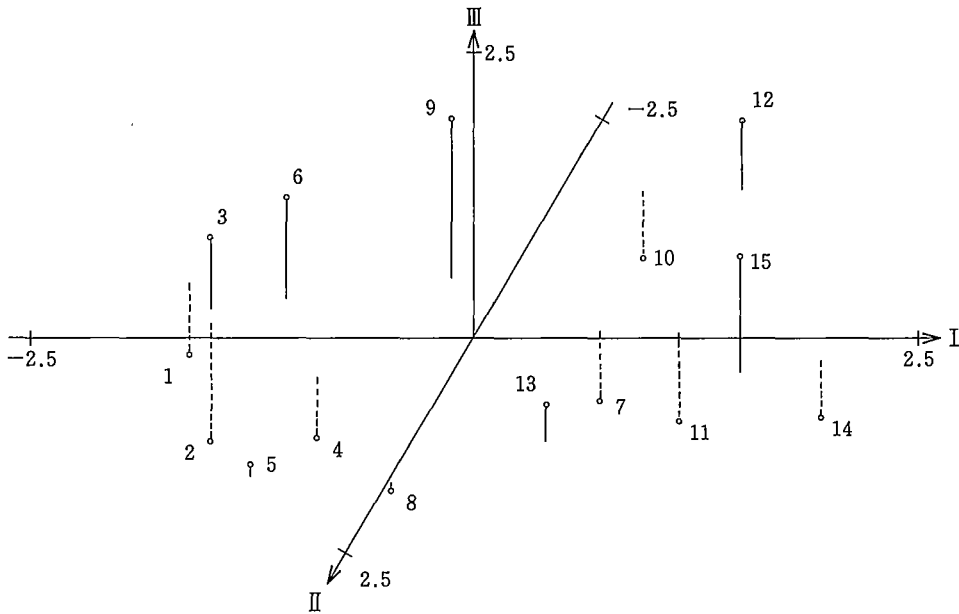


図6 一般S-P表(表1)のデータに対するI, II, III座標プロット

トされている。しかし、4.1節の場合同様、I軸の方向に合計得点の大小が表わされていることがわかる。

また一致係数 m_{ij} の再現性については、立体プロットにおける $10 - \{d(P_i, P_j)\}^2/2$ の値を表9に示した。比 $m_{ij}/[10 - \{d(P_i, P_j)\}^2/2]$ の平均値は0.770、分散は0.014であり、完全S-P表の場合よりはバラツキが大きくなっている。

表9 立体プロット(図6)での $10 - |d(P_i, P_j)|^2/2$ の値

	(個体番号)														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	10.0000	9.7594	8.8397	8.9094	6.6431	8.2333	6.9144	5.7804	6.4000	7.0626	5.6190	4.7963	4.6691	2.5796	2.0421
2		10.0000	8.2413	9.4050	7.3783	7.5703	7.5619	6.7848	5.5740	6.9068	6.4593	4.2153	5.4905	3.6539	2.3957
3			10.0000	8.5386	7.8714	9.8661	6.3998	6.8711	8.8167	6.0683	5.1231	5.5831	6.3119	2.6294	4.6905
4				10.0000	9.0196	8.2877	8.8598	8.8997	7.0160	7.2894	8.2160	5.5607	8.1546	6.2085	5.7134
5					10.0000	7.7271	7.1046	9.6960	6.4606	3.9315	6.4003	2.9574	8.7384	4.8427	6.0263
6						10.0000	6.7409	6.9831	9.4753	6.4981	5.6270	6.5988	6.8961	3.5353	5.9292
7							10.0000	8.0770	6.7276	8.9263	9.8847	7.8572	8.5723	9.0573	7.6182
8								10.0000	6.2417	4.7534	7.7384	3.9545	9.5419	6.7976	7.3312
9									10.0000	6.8436	5.9318	8.0432	7.1221	4.6002	7.4984
10										10.0000	8.5885	9.1408	5.7651	7.3310	5.7880
11											10.0000	7.6692	8.5206	9.5808	7.8436
12												10.0000	5.9013	7.0935	7.4075
13													10.0000	8.2367	9.0614
14														10.0000	8.3195
15															10.0000

4.3 $n = 30, v = 10$ の場合

表10のような得点分布を考える。このデータより作られるS-P表を表11に示す。4.1節と同様に完全反応パターンを求めて表12に示す。このデータに対して主座標分析を行なう。行列 α の第1~第3固有値および固有ベクトルを表13に示す。このとき第3

表10 $n = 30, v = 10$ の場合のデータ

	(問題番号)										
	1	2	3	4	5	6	7	8	9	10	
1	1	1	1	1	1	1	1	1	1	1	10テン
2	1	1	1	1	1	1	1	1	1	0	9テン
3	1	1	1	1	1	1	1	1	0	1	9テン
4	1	1	1	1	1	1	1	0	1	1	9テン
5	1	1	1	1	1	1	0	1	1	1	9テン
6	1	1	1	1	1	1	1	1	0	0	8テン
7	1	1	1	1	1	1	1	0	0	1	8テン
8	1	1	1	1	1	0	1	0	1	1	8テン
9	1	1	1	1	1	1	0	0	0	0	7テン
10	1	1	1	1	1	0	1	0	0	1	7テン
11	1	1	1	1	1	1	0	0	0	0	6テン
12	1	1	1	1	1	0	0	1	0	0	6テン
13	1	1	1	0	1	0	1	0	1	0	6テン
14	1	1	1	0	0	1	0	1	1	0	6テン
15	1	1	1	1	1	0	0	0	0	0	5テン
16	1	1	1	1	0	1	0	0	0	0	5テン
17	1	1	0	0	1	0	1	0	0	1	5テン
18	1	1	1	1	0	0	0	0	0	0	4テン
19	1	1	1	0	1	0	0	0	0	0	4テン
20	1	1	1	0	0	0	1	0	0	0	4テン
21	1	1	1	0	0	0	0	0	0	0	3テン
22	1	0	1	1	0	0	0	0	0	0	3テン
23	1	1	0	1	0	0	0	0	0	0	3テン
24	0	1	0	1	0	1	0	0	0	0	3テン
25	1	1	0	0	0	0	0	0	0	0	2テン
26	0	1	1	0	0	0	0	0	0	0	2テン
27	1	0	0	0	1	0	0	0	0	0	2テン
28	1	0	0	0	0	0	0	0	0	0	1テン
29	0	1	0	0	0	0	0	0	0	0	1テン
30	0	0	1	0	0	0	0	0	0	0	1テン

表11 表10のデータからのS-P表

	(問題)										CS _i
	1	2	3	5	4	7	6	10	8	9	
1	1	1	1	1	1	1	1	1	1	1	10テン 0.000
3	1	1	1	1	1	1	1	1	1	0	9テン 0.000
2	1	1	1	1	1	1	1	0	1	1	9テン 0.116
4	1	1	1	1	1	1	1	1	0	1	9テン 0.116
5	1	1	1	1	1	0	1	1	1	1	9テン 0.581
6	1	1	1	1	1	1	1	0	1	0	8テン 0.000
7	1	1	1	1	1	1	1	1	0	0	8テン 0.000
8	1	1	1	1	1	1	0	1	0	1	8テン 0.247
9	1	1	1	1	1	1	1	0	0	0	7テン 0.000
10	1	1	1	1	1	1	1	0	1	0	7テン 0.126
11	1	1	1	1	1	1	0	1	0	0	6テン 0.035
12	1	1	1	1	1	1	0	0	1	0	6テン 0.141
13	1	1	1	1	0	1	0	0	0	1	6テン 0.352
14	1	1	1	0	0	0	0	1	0	1	6テン 0.739
15	1	1	1	1	1	1	0	0	0	0	5テン 0.000
16	1	1	1	1	1	0	0	0	0	1	5テン 0.281
17	1	1	0	1	0	1	0	1	0	0	5テン 0.625
19	1	1	1	1	0	0	0	0	0	0	4テン 0.000
18	1	1	1	0	1	0	0	0	0	0	4テン 0.033
20	1	1	1	0	0	0	1	0	0	0	4テン 0.196
21	1	1	1	0	0	0	0	0	0	0	3テン 0.000
23	1	1	0	1	0	0	0	0	0	0	3テン 0.213
22	1	0	1	1	0	1	0	0	0	0	3テン 0.319
24	0	1	0	1	0	1	0	1	0	0	3テン 0.745
25	1	1	0	0	0	0	0	0	0	0	2テン 0.000
26	0	1	1	0	0	0	0	0	0	0	2テン 0.144
27	1	0	0	0	1	0	0	0	0	0	2テン 0.385
28	1	0	0	0	0	0	0	0	0	0	1テン 0.000
29	0	1	0	0	0	0	0	0	0	0	1テン 0.000
30	0	0	1	0	0	0	0	0	0	0	1テン 0.288

表12 表11のS-P表に対する
完全反応パターン

1.	1	1	1	1	1	1	1	1	1	1
2.	1	1	1	1	1	1	1	1	1	0
3.	1	1	1	1	1	1	1	1	1	0
4.	1	1	1	1	1	1	1	1	1	0
5.	1	1	1	1	1	1	1	1	1	0
6.	1	1	1	1	1	1	1	1	0	0
7.	1	1	1	1	1	1	1	1	0	0
8.	1	1	1	1	1	1	1	1	0	0
9.	1	1	1	1	1	1	1	0	0	0
10.	1	1	1	1	1	1	1	0	0	0
11.	1	1	1	1	1	1	0	0	0	0
12.	1	1	1	1	1	1	0	0	0	0
13.	1	1	1	1	1	1	0	0	0	0
14.	1	1	1	1	1	1	0	0	0	0
15.	1	1	1	1	1	0	0	0	0	0
16.	1	1	1	1	1	0	0	0	0	0
17.	1	1	1	1	1	0	0	0	0	0
18.	1	1	1	1	0	0	0	0	0	0
19.	1	1	1	1	0	0	0	0	0	0
20.	1	1	1	1	0	0	0	0	0	0
21.	1	1	1	0	0	0	0	0	0	0
22.	1	1	1	0	0	0	0	0	0	0
23.	1	1	1	0	0	0	0	0	0	0
24.	1	1	1	0	0	0	0	0	0	0
25.	1	1	0	0	0	0	0	0	0	0
26.	1	1	0	0	0	0	0	0	0	0
27.	1	1	0	0	0	0	0	0	0	0
28.	1	0	0	0	0	0	0	0	0	0
29.	1	0	0	0	0	0	0	0	0	0
30.	1	0	0	0	0	0	0	0	0	0

固有値までの累積寄与率は85.43%である。I, II主座標によるプロットおよびI, II, III主座標による立体プロットを図7, 図8に図示するが, このときのプロットの様子は前述の $n = 15, v = 10$ の完全S-P表の場合に共通していることがわかる。(ただし, 固有ベクトルの各要素の符号が逆になる場合には同じデータに対しても軸に関して裏返しの状態でプロットされる。)

m_{ij} の再現性については, 立体プロットにおいて, 比 $m_{ij}/[10 - \{d(P_i, P_j)\}^2/2]$ の平均値は0.929, 分散は0.003である。

表13 表12のデータにおける
行列 α の固有値および固有ベクトル

主座標	I	II	III
固有値	57.8295	15.0605	6.8443
(累積寄与率)	(0.6196)	(0.7810)	(0.8543)
1	-1.9942	0.9122	0.4420
2	-1.9252	0.7910	0.3129
3	-1.9252	0.7910	0.3129
4	-1.9252	0.7910	0.3129
5	-1.9252	0.7910	0.3129
6	-1.5899	0.2497	-0.1820
7	-1.5899	0.2497	-0.1820
8	-1.5899	0.2497	-0.1820
9	-1.0896	-0.3911	-0.5173
10	-1.0896	-0.3911	-0.5173
11	-0.5140	-0.9280	-0.5503
12	-0.5140	-0.9280	-0.5503
13	-0.5140	-0.9280	-0.5503
14	-0.5140	-0.9280	-0.5503
15	0.1327	-0.9720	0.0599
16	0.1327	-0.9720	0.0599
17	0.1327	-0.9720	0.0599
18	0.7657	-0.6287	0.6176
19	0.7657	-0.6287	0.6176
20	0.7657	-0.6287	0.6176
21	1.3192	-0.0350	0.6339
22	1.3192	-0.0350	0.6339
23	1.3192	-0.0350	0.6339
24	1.3192	-0.0350	0.6339
25	1.6902	0.5773	-0.0908
26	1.6902	0.5773	-0.0908
27	1.6902	0.5773	-0.0908
28	1.8859	0.9597	-0.7359
29	1.8859	0.9597	-0.7359
30	1.8859	0.9597	-0.7359
要素2乗和	57.8295	15.0605	6.8443

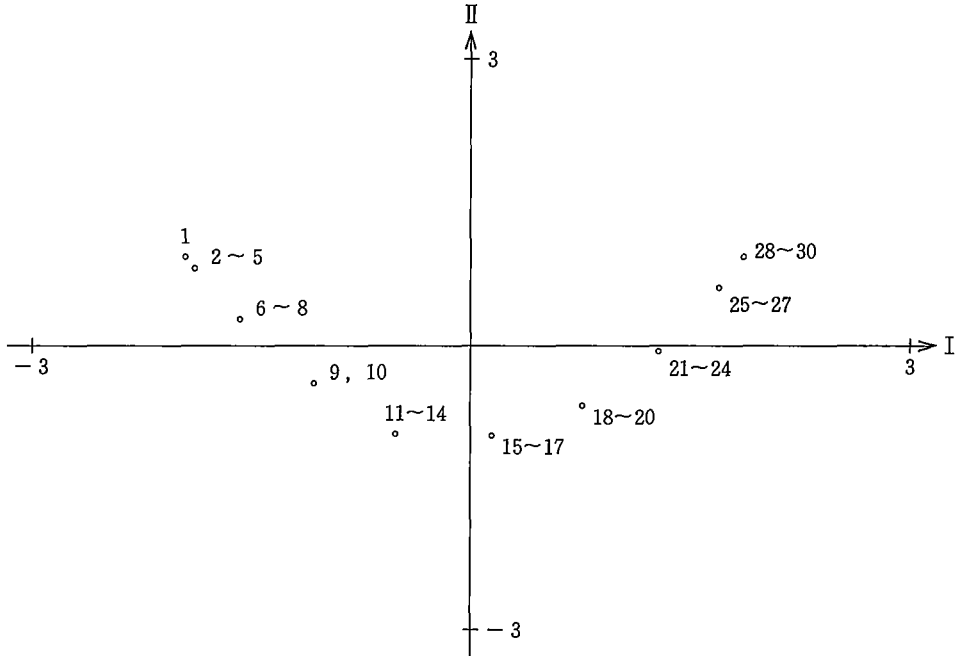


図7 完全S-P表(表12)のデータに対するI, II主座標プロット

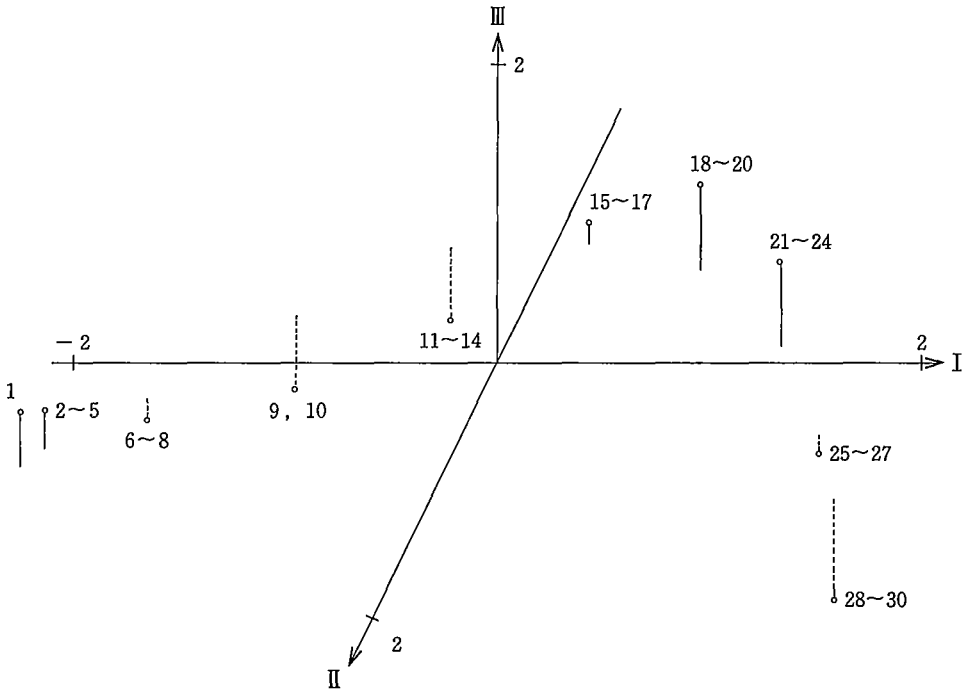


図8 完全S-P表(表12)のデータに対するI, II, III主座標プロット

次に、表11のS-P表のデータに主座標分析を行なう。表14はI, II, III主座標と固有ベクトルである。累積寄与率は第3固有値までで63.23%である。I, II主座標プロットおよびI, II, III主座標プロットを図9, 図10に示す。4.2節同様、各個体の合計得点の大小を表わしていることがわかる。さらに個体1, 3, 6, 7, 9, 15, 19, 21, 25, 28, 29は表11のS-P表では注意係数 $CS_i = 0$ であるが、それらの個体はI軸の近くに位置していることがわかる。 m_{ij} の再現性については、立体プロットにおいて、比 $m_{ij}/[10 - \{d(P_i, P_j)\}^2/2]$ の平均値は0.798, 分散は0.009である。

5. 結 語

2次元または、3次元プロットによる表現方法により各個体間の関係が明らかになり、類似した個体とそうでない個体を見わけることが、一致係数をほぼ再現しながら可能となることがわかる。また完全S-P表の形のデータに対しては特徴ある一定の形のプロットを示すことがわかる。

さらに、単に名義尺度としての1, 0データの場合のみならず、何段階かの値をもつデータをもつS-P表に対しても主座標分析の適用が可能である。

表14 表11のデータにおける
行列 α の固有値および固有ベクトル

主 座 標	I	II	III
固 有 値 (累 積 寄 与 率)	47.0383 (0.4046)	14.6981 (0.5310)	11.7779 (0.6323)
1	2.2140	0.2486	-0.3270
2	1.7366	0.7885	-0.4621
3	1.8271	0.0978	0.2546
4	1.8395	-0.4878	0.2182
5	1.6501	0.9333	-0.2668
6	1.3496	0.6376	0.1195
7	1.4526	-0.6386	0.7998
8	1.2983	-1.1029	-0.1277
9	0.9752	-0.0988	0.6648
10	0.9114	-1.2538	0.4539
11	0.4113	0.5858	0.7250
12	0.2445	0.7071	-0.1662
13	0.3012	-0.8344	-1.2679
14	0.0603	1.5634	-1.1917
15	-0.1300	-0.0293	0.3791
16	-0.2752	0.4358	-1.1714
17	0.0357	-1.7639	-0.2098
18	-0.7226	0.3323	0.5944
19	-0.6496	-0.3006	-0.6261
20	-0.6784	-0.6236	-0.4710
21	-1.2423	0.0610	-0.4108
22	-0.9695	0.2899	0.6212
23	-1.0787	0.0935	0.9358
24	-0.7868	0.8908	1.5028
25	-1.5984	-0.1779	-0.0693
26	-1.4917	0.2432	-0.1897
27	-1.2525	-0.5818	-0.2578
28	-1.8452	-0.2202	-0.0425
29	-1.8477	0.0043	0.1517
30	-1.7385	0.2008	-0.1629
要 素 2 乗 和	47.0383	14.6981	11.7779

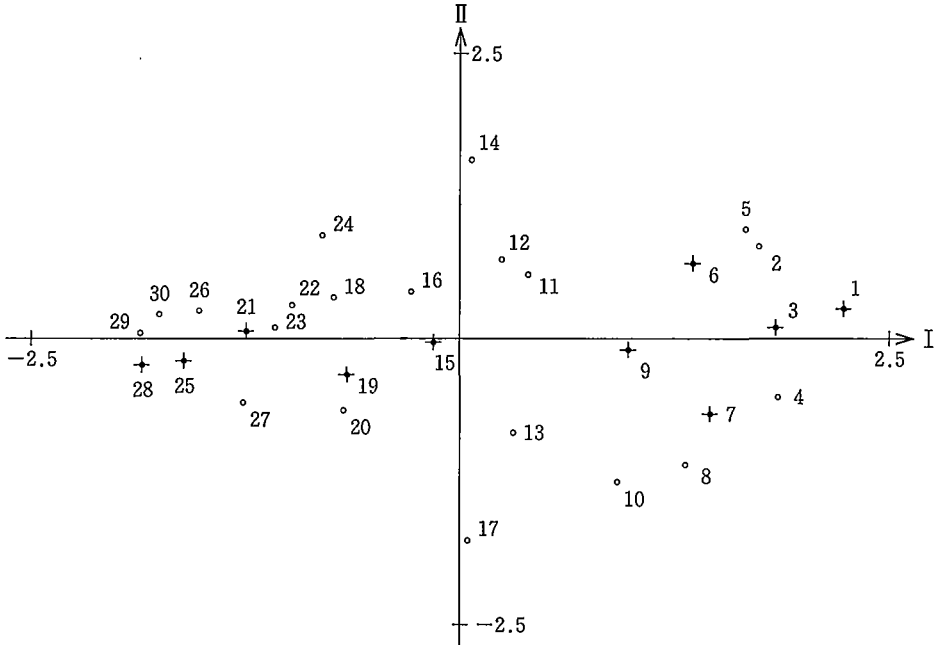


図9 一般S-P表(表11)のデータに対するI, II主座標プロット
+印はS-P表において注意係数 $CS_i = 0$ のもの

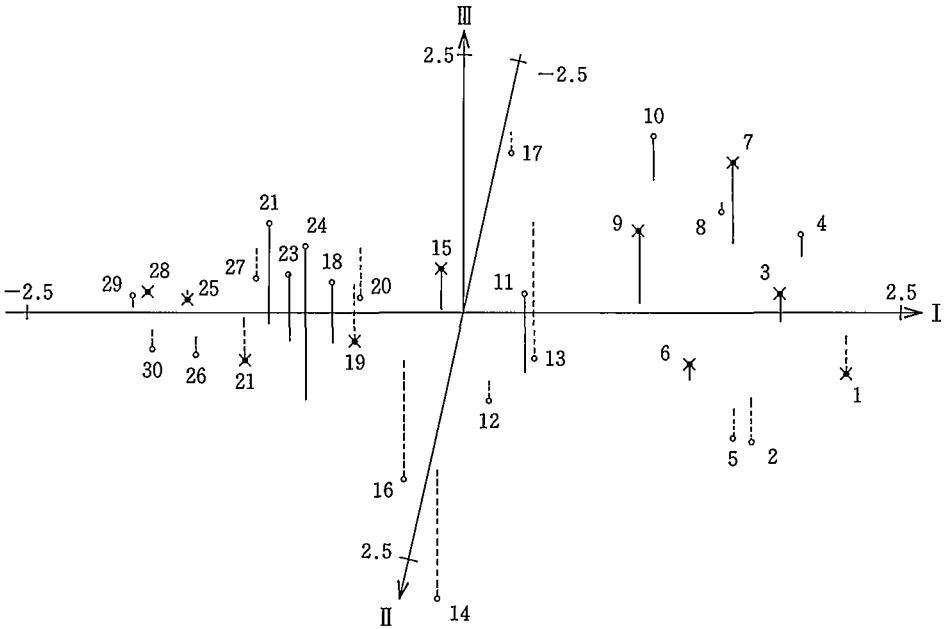


図10 一般S-P表(表11)のデータに対するI, II, III主座標プロット
×印はS-P表において注意係数 $CS_i = 0$ のもの

参考文献

- 1) Banfield, C. F. and Gower, J. C. (1980), A Note on the Graphical Representation of Multivariate Binary Data. *Appl. Statist.* 29, 238-245.
- 2) Barnett, V. (1981), *Interpreting Multivariate Data*. New York: Wiley.
- 3) Gower, J. C. (1966), Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53, 325-338.
- 4) Maindonald, J. H. (1984), *Statistical Computation*. New York: Wiley.
- 5) 田中 豊・脇本和昌, 多変量統計解析, 現代数学社, 1983.
- 6) 佐藤隆博, S-P表の作成と解釈, 明治図書, 1975.
- 7) ———, 授業設計と評価のデータ処理技法, 明治図書, 1980.
- 8) 林知己夫・鮑戸 弘, 多次元尺度解析法, サイエンス社, 1976.

(昭和59年7月16日受理)