

底生生物の分布および種類組成を類別する 統計的手法の検討

赤嶺 達郎¹⁾

On Statistical Methods in order to Classify the Distributions of Species and the Species Compositions of Stations on benthic organisms

TATSURO AKAMINE

Abstract

As the data for the purpose of this analysis, 20 species of benthic organisms caught with beam trawl at 20 stations off the coast of Tottori Pref. are used. Statistical methods compared in this investigation are as follows:

1) Principal component analysis ... Species and Stations are classified with score and factor loading calculated from the variance and covariance matrix or the correlation matrix of species or stations.

2) Dendrogram ... Clusters are combined by a minimum method. Similarity indices as distance are correlation coefficient, standardized Euclidean distance, Morishita's C_i index, Kimoto's C_{II} index, Odam's PD , and Whittaker's PS .

3) Polar ordination with Phytagorean distance.

According to those methods, the following results are obtained;

1) The correlation matrix should be applied to the classification. An addition of O-point in relation to score and minimum spanning tree are necessary for the correct judgement. A logarithm transformation was not necessary for this data. Distributions of condition factors on a 2-dimensional diagram of score or factor loading of stations make it easy to judge the influence of condition factors.

2) The correlation coefficient and C_{II} index are better for the similarity index.

3) Polar ordination is too subjective for analysis. Thus, it may not be applicable for the purpose.

I. 緒 言

この研究の目的は貝桁網調査のデータから、1次解析としてすみやかに漁場や魚種の全体像をつかむための簡易手法を検討することである。検討した統計的手法は主成分分析、デンドロ

1) 〒 951 新潟市水道町1丁目5939-22 日本海区水産研究所
(Japan Sea Regional Fisheries Research Laboratory, Suido-cho, Niigata 951, Japan)

グラム、極座標づけ (polar ordination) である。

動物の個体群や群集の類似性を解析する統計的手法は、木元 (1976) や Jeffers (1982) によって概括的に述べられているが、水産研究に応用しにくい問題点があり、以下にこれらを摘記する。

1. 個体数で判定していること……昆虫等では個体間、種間における生物量の差が比較的小さく、個体数=生物量としてもさしつかえないが、貝桁網等の採集物は魚種、年令等によつて生物量の差が著しい。

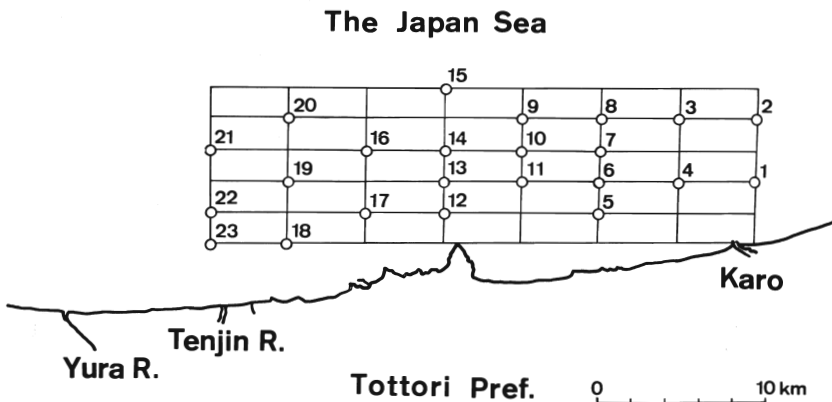
2. 漁具効率……陸上では昆虫や植物を杵取り法で100%採集できるが、貝桁網等の手法では漁具効率が低く、しかも魚種や採集状態によつて著しく異なる。

3. 環境要因の判定……デンドログラムでは類別は明確に行なうことができるが、環境要因の判定が容易ではない。産業との関係が強い水産ではこれが容易に行なえる手法が必要である。

本研究では出現数の多い20種を摘出し、魚種ごとの寄与に差がでないように検討することによつて1., 2. の影響を少なくした。3. についてはWhittaker (1979) が傾度分析の一手法としてピタゴラス距離による極座標づけを紹介している。座標づけにより環境要因の判定は容易になるが、この手法は主観的要素が強い。従つてここでは客観的に座標づけする手法として主成分分析を採用した。

主成分分析を中心とした解析は、鉄 (1977) が相模湾内定置網漁獲物、大森 (1980) が油谷湾底曳網漁獲物について行なつているが、いずれも対数変換した後に相関行列を用いて、主として因子負荷量から分析している。また、水戸 (1977) はベーリング海の底魚群集について recurrent group 分析法、群 (クラスター) 分析法、因子分析法 (セントロイド法) によつて検討を行ない、因子分析法で最も明確に分類できるとした。さらに相関係数を近接率と順位相関係数と比較して、類似度指数として有効であるとした。

本研究では主成分分析の他に、クラスター分析として最短距離法によるデンドログラム、座標づけの手法として極座標づけを行ない比較検討した。またこれと関連して最小距離樹形図 (minimum spanning tree)、類似度指数の比較検討、環境要因の判定を行なつた。最短距離法 (minimum method, shortest distance method) は最近隣法 (nearest neighbor method)、単一連関・単 (純) 連結法 (single linkage method) と同一のものである。

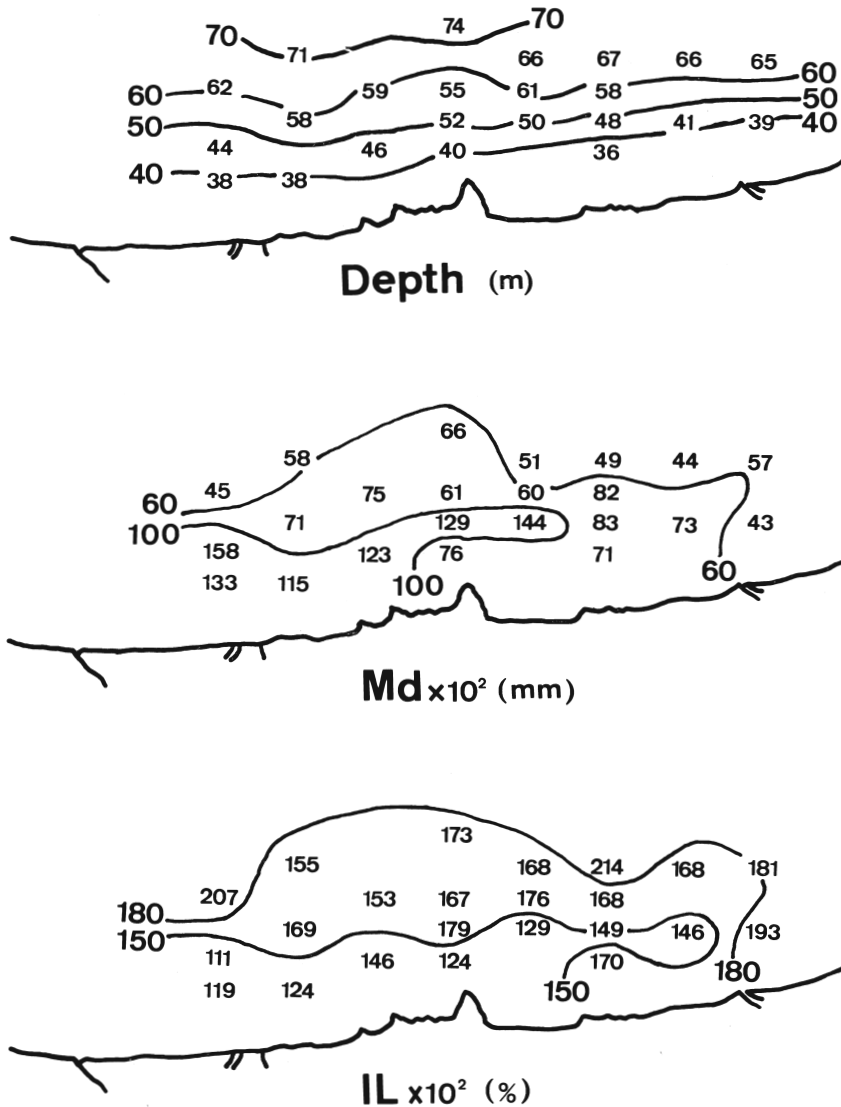


第1図 貝桁網の調査点
Fig. 1. Sampling stations of beam trawl.

II. 材料と方法

調査は1981年7月26~28日に鳥取県沿岸域のイタヤガイ漁場（第1図）において調査船みずほ丸で貝桁網を用いて行なつた。使用した貝桁網は間口2 m×20cm, つめの長さ20cm, 目合1.5 cmの網で, 1.5kt 10分間曳きで行なつた。同時に光研式採泥器で底質を採取した。各調査点の水深, 中央粒径値 (Md), 強熱減量 (IL) の値と等量線を第2図に示す。なお, St. 9~11はしいら漬けの漁具がはいつており曳網できなかつた。St. 1は網破損のため採集個体数が著しく少なかつた。

採集した生物は177種2778個体であつたが, この中から出現数の多い20種2077個体を摘出し



第2図 環境要素の値
Fig. 2. Condition of stations.

第 1 表 解析に使用した種

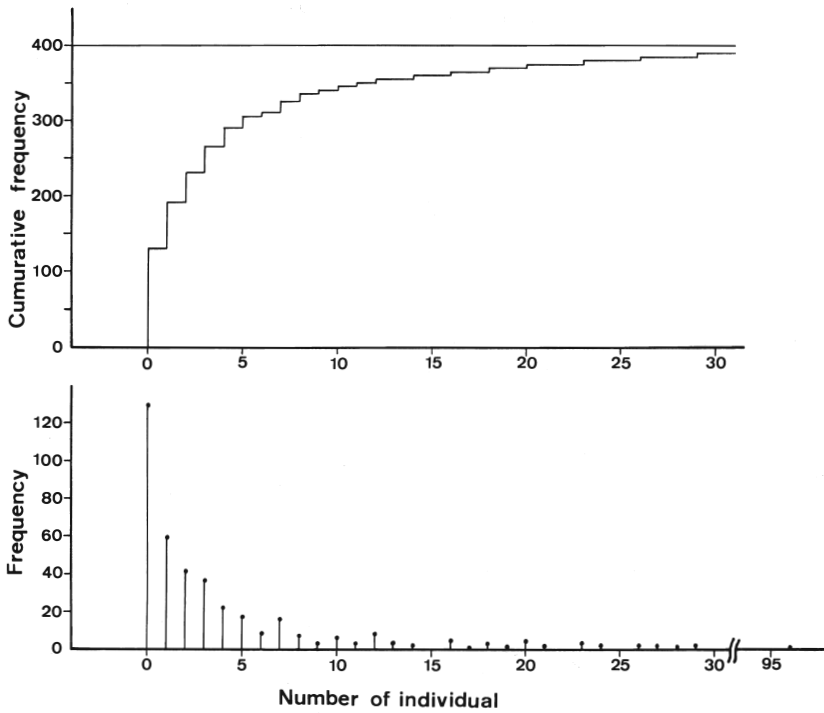
Table 1. Selected speices for this analysis.

mark	class	species
G1	GASTROPODA	<i>Fusinus perplexus perplexus</i>
G2		<i>Doxander japonicus</i>
G3		<i>Philine argentata</i>
G4		<i>Niotha variegata</i>
G5		<i>Murex sobrinus</i>
B1	BIVALVIA	<i>Glycymeris vestita</i>
B2		<i>Eucrassatella nana</i>
B3		<i>Megacardita ferruginosa</i>
B4		<i>Pecten albicans</i>
S1	SCAPHOPODA	
E1	ECHINOIDEA	<i>Glyptocidaris crenularis</i>
E2		<i>Peronella japonica</i>
C1	CRUSTACEA	<i>Metapeaeopsis dalei</i>
C2		<i>Diogenes edwardsi</i>
C3		<i>Actumnus squamosus</i>
C4		<i>Leucosia haematosticta</i>
P1	OSTEICHTHYES	<i>Heteromycteris japonicus</i>
P2		<i>Lepidotrigla microptera</i>
P3		<i>Hypodytes rubripinnis</i>
P4		<i>Pleuronichthys cornutus</i>

第 2 表 解析に

Table 2. Data

SP	ST	1	2	3	4	5	6	7	8	12	13
	G1	0	16	6	14	27	33	12	3	10	20
G2	0	7	1	29	12	3	17	2	0	5	
G3	0	1	32	2	1	0	0	1	0	0	
G4	0	7	4	0	0	0	4	3	1	2	
G5	0	5	3	1	0	0	0	4	0	1	
B1	1	2	0	12	5	8	2	0	23	4	
B2	1	4	1	12	2	0	0	1	4	1	
B3	0	7	6	0	0	0	0	1	0	0	
B4	0	0	1	7	6	2	3	1	4	1	
S1	5	0	3	3	1	96	2	0	28	14	
E1	0	2	0	0	7	7	2	6	52	27	
E2	0	1	3	0	0	0	4	1	1	2	
C1	0	16	19	2	1	8	10	24	5	7	
C2	2	36	1	5	23	4	8	0	8	38	
C3	0	7	2	1	0	1	0	5	1	4	
D4	0	3	4	0	0	0	0	4	0	0	
P1	0	2	3	3	3	0	4	3	2	6	
P2	0	5	2	0	0	1	0	2	1	1	
P3	0	0	0	3	0	4	0	0	0	12	
P4	0	3	2	2	0	1	0	3	0	1	
Total		9	124	93	96	88	168	68	64	140	146
<i>m</i>		0.45	6.20	4.65	4.80	4.40	8.40	3.40	3.20	7.00	7.30
<i>s</i>		1.2	8.4	7.6	7.2	7.8	21.9	4.8	5.2	13.1	10.3



第3図 データの要素の累積度数分布曲線と棒グラフ
 Fig. 3. Distribution curve of cumulative frequency and bar diagram of elements of data.

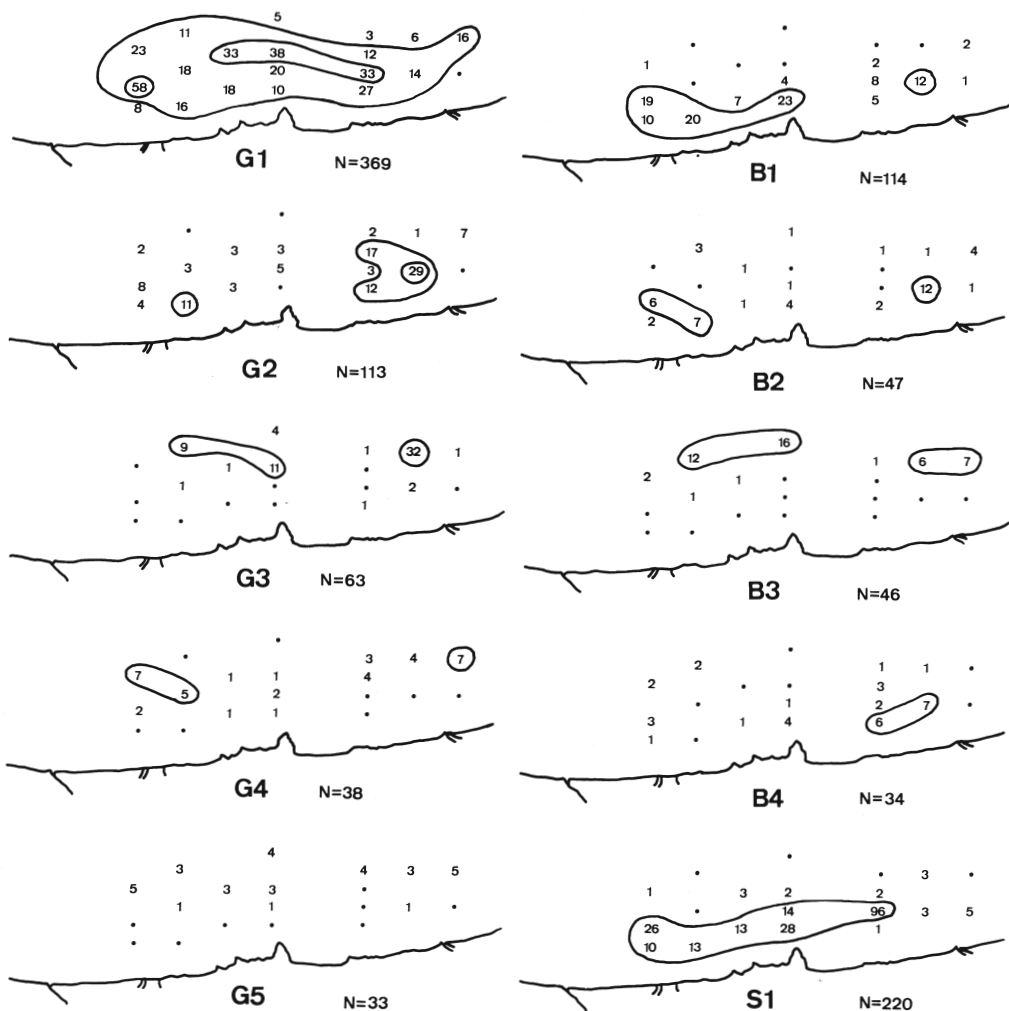
使用したデータ
 of this analysis.

14	15	16	17	18	19	20	21	22	23	Total	<i>m</i>	<i>s</i>
38	5	33	18	16	18	11	23	58	8	369	18.45	13.9
3	0	3	3	11	3	0	2	8	4	113	5.65	7.1
11	4	1	0	0	1	9	0	0	0	63	3.15	7.4
1	0	1	1	0	5	0	7	2	0	38	1.90	2.3
3	4	3	0	0	1	3	5	0	0	33	1.65	1.9
0	0	0	7	20	0	0	1	19	10	114	5.70	7.4
0	1	1	1	7	0	3	0	6	2	47	2.35	3.0
0	16	1	0	0	1	12	2	0	0	46	2.30	4.5
0	0	0	1	0	0	2	2	3	1	34	1.70	2.0
2	0	3	13	13	0	0	1	26	10	220	11.00	21.7
36	24	29	7	2	13	7	2	51	7	281	14.05	16.7
3	49	6	0	1	0	26	1	1	0	99	4.95	11.9
9	12	3	4	0	5	20	4	3	1	153	7.65	7.1
20	0	18	3	7	10	0	0	21	2	206	10.30	11.8
3	12	3	0	0	5	5	4	1	1	55	2.75	3.0
1	2	3	0	0	0	2	0	1	0	20	1.00	1.5
2	0	4	5	6	3	1	9	21	2	79	3.95	4.6
2	1	2	1	0	7	8	5	2	0	40	2.00	2.4
0	0	0	10	1	0	0	0	4	0	34	1.70	3.5
3	2	2	3	0	2	1	5	2	1	33	1.65	1.4
137	132	116	77	84	74	110	73	229	49			
6.85	6.60	5.80	3.85	4.20	3.70	5.50	3.65	11.5	2.45			
11.4	12.0	9.5	5.0	6.2	5.0	7.2	5.2	16.9	3.4			

た. 20種の種名を第1表に示す. この20種の調査点別採集個体数をもとに以下の解析を行なつた. データを第2表に, 400の要素の値の分布を第3図に, そして20種の分布図を第4図に示す.

解析手順を第5図に示す. 以下に手順を述べる.

1. 種についての分散共分散行列 V_{sp} , 相関行列 R_{sp} , 調査点についての分散共分散行列 V_{st} , 相関行列 R_{st} をつくり, それぞれについて主成分分析を行ない, スコアおよび因子負荷量によつて類別を試みた.
2. 対数変換 $\ln(1+n)$ を行なつたデータについても同様の計算を行なつた.
3. 主成分分析の欠点を補う手法として最小距離樹形図を作成した.
4. 最短距離法によるデンドログラムを作成した. 使用した類似度指数は種についての相関係数 r_{sp} , 調査点についての相関係数 r_{st} ;



第4-a図 種の分布. 種名: 第1表参照. 黒丸: 0.

Fig. 4-a. Distribution of each species. Species name: see table 1. Black circles : 0.

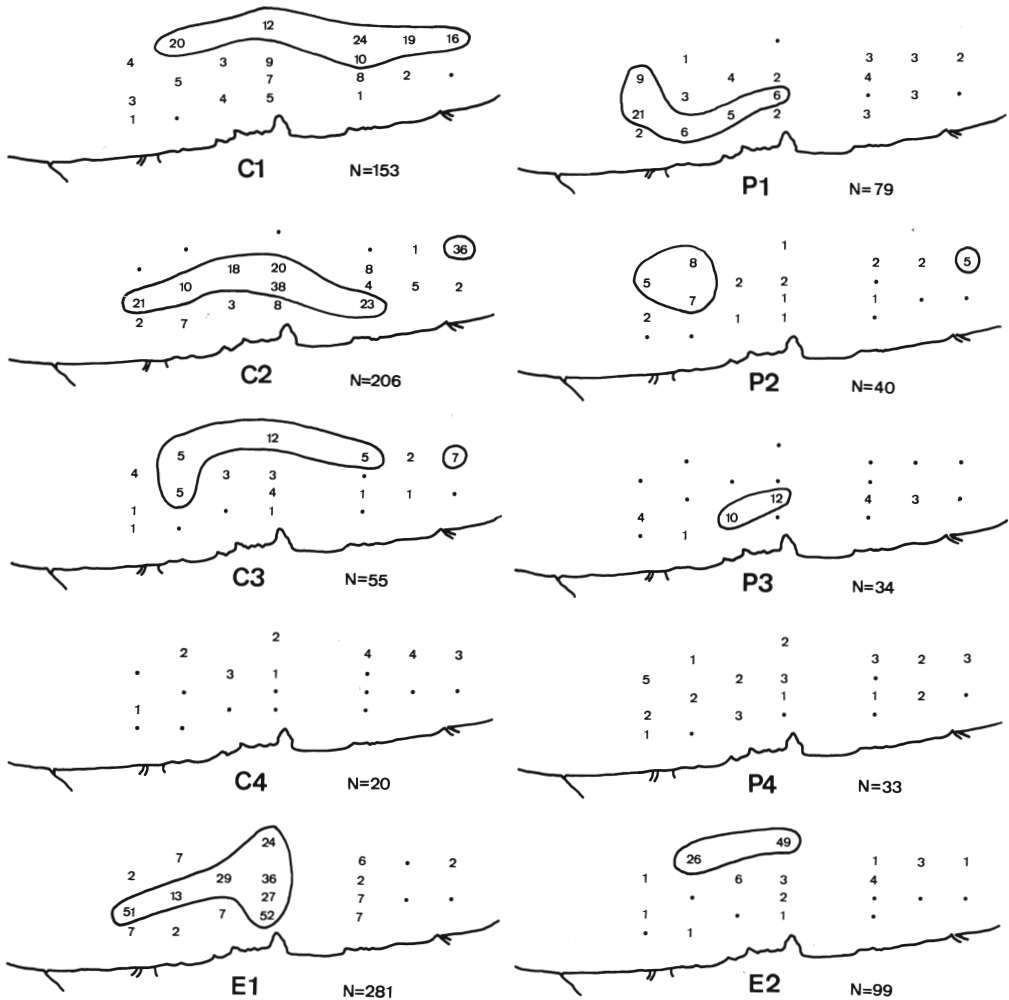
$$r_{sp} = \frac{S_{ij} - \frac{N_i N_j}{q}}{\sqrt{\left(S_{ii} - \frac{N_i^2}{q}\right)\left(S_{jj} - \frac{N_j^2}{q}\right)}}, \quad r_{st} = \frac{S_{XY} - \frac{N_X N_Y}{p}}{\sqrt{\left(S_{XX} - \frac{N_X^2}{p}\right)\left(S_{YY} - \frac{N_Y^2}{p}\right)}}$$

種についての標準ユークリッド距離 d_{sp} , 調査点についての標準ユークリッド距離 d_{st} ;

$$d_{sp} = \sqrt{\frac{\sum_X (n_{iX} - n_{jX})^2}{\left(S_{XX} - \frac{N_X^2}{p}\right) / (p-1)}}, \quad d_{st} = \sqrt{\frac{\sum_i (n_{iX} - n_{iY})^2}{\left(S_{ii} - \frac{N_i^2}{q}\right) / (q-1)}}$$

森下の C_2 指数, その修正式である木元の C_{II} 指数;

$$C_2 = \frac{\frac{S_{XY}}{N_X N_Y}}{\left(\frac{S_{XX} - N_X}{N_X^2 - N_X} + \frac{S_{YY} - N_Y}{N_Y^2 - N_Y}\right) / 2}, \quad C_{II} = \frac{\frac{S_{XY}}{N_X N_Y}}{\left(\frac{S_{XX}}{N_X^2} + \frac{S_{YY}}{N_Y^2}\right) / 2}$$



第4-b図 種の分布. 種名: 第1表参照. 黒丸: 0.

Fig. 4-b. Distribution of each species. Species name: see table 1. Black circles: 0.

Odum の PD (percentage difference);

$$PD = \frac{\sum_i |n_{iX} - n_{iY}|}{N_X + N_Y} = 1 - 2 \frac{\sum_i \text{Min}(n_{iX}, n_{iY})}{N_X + N_Y}$$

Whittaker の PS (percentage similarity);

$$PS = 1 - 2 \sum_i \left| \frac{n_{iX}}{N_X} - \frac{n_{iY}}{N_Y} \right| = \sum_i \text{Min} \left(\frac{n_{iX}}{N_X}, \frac{n_{iY}}{N_Y} \right)$$

以上 8 種である。なお、 p は種の総数、 q は調査点の総数であり、 n_{iX} は種 : i の調査点 : X における出現個体数である。

また、

$$N_i = \sum_{X=1}^q n_{iX}, \quad N_X = \sum_{i=1}^p n_{iX}$$

$$S_{ij} = \sum_{X=1}^q n_{iX} n_{jX}, \quad S_{XY} = \sum_{i=1}^p n_{iX} n_{iY}$$

である。

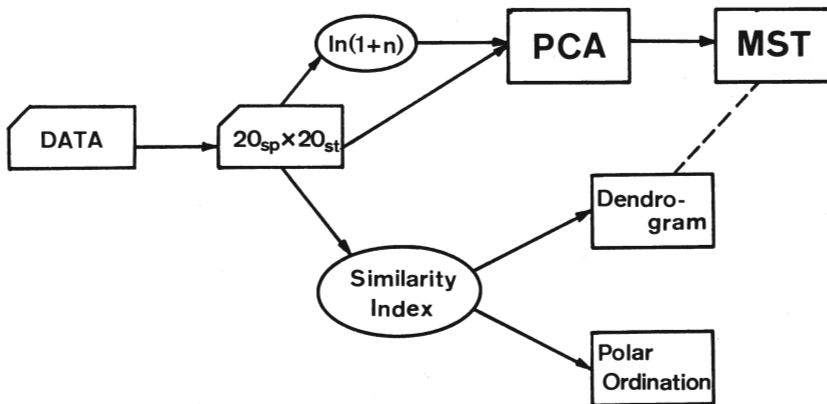
5. 上記の類似度指数について相関図から有効性を検討した。
6. PD, PS についてピタゴラス距離による極座標づけを行ない 2 次元表示を試みた。ピタゴラス距離は両端の 2 点を a, b とおくと、

$$X = \frac{L^2 + Da^2 - Db^2}{2L}$$

で与えられる。ここで l は 2 点 a, b の距離、 Da, Db は点 X から点 a, b までの距離である。距離が最大の 2 点を両端の 2 点にとることにより、他のすべての点の座標を得ることができる。

7. 主成分分析より得られた 2 次元図に環境要素として水深、中央粒径値 (Md)、強熱減量 (IL) を書き込み、等量線を描いて調査点、種との関係を考察した。

類似度指数は全データについて計算すべきものであり、またデンドログラムの作成方法として最短距離法はかなり欠点のあることが木元 (1976) 等に指適されている。しかし、1 次解析としてはこのような方法でもよいと判断した。



第 5 図 解析の手順
Fig. 5. The main process of this analysis.

また、主成分分析の計算は農林水産研究センターのクロードパッチ処理を利用し、石間のプログラム (1969) で行なった。

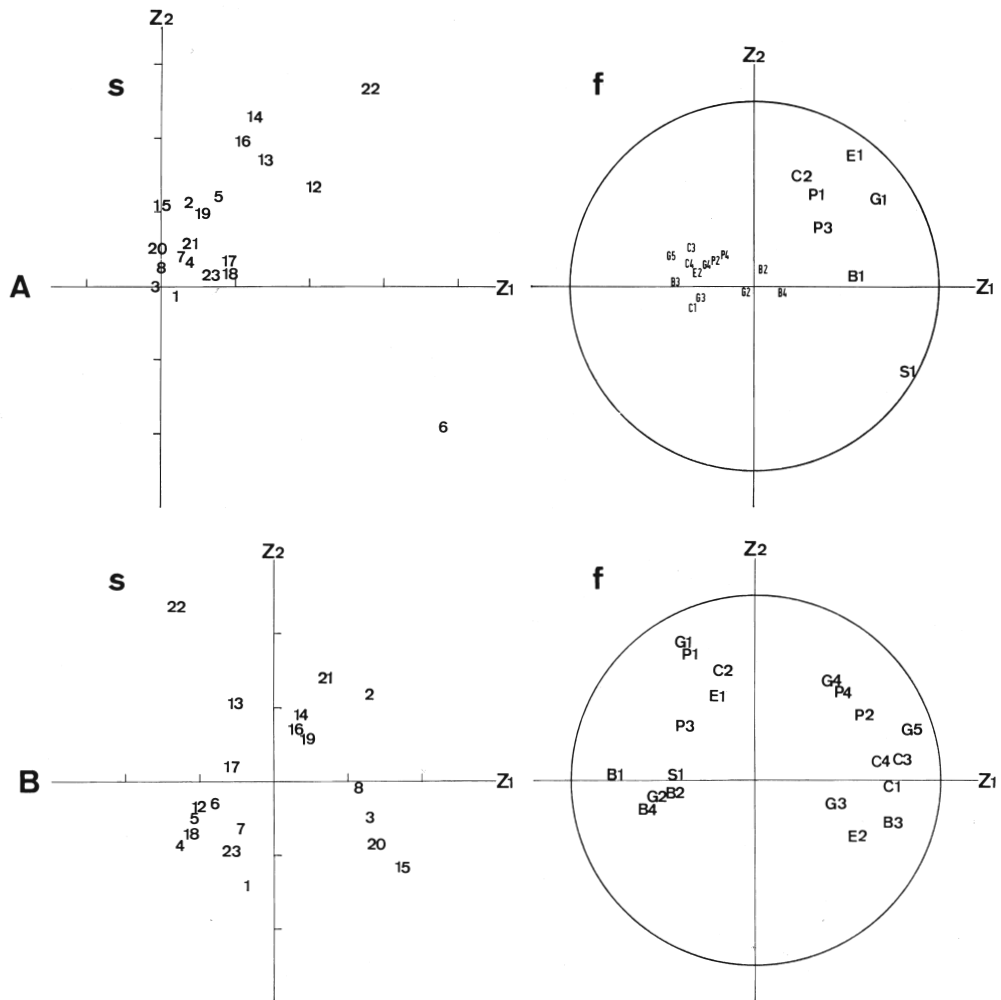
なお、図の表題では行列 M を用いた主成分分析を $PCA (M)$ と略記した。

Ⅲ. 結果と考察

1. 主成分分析

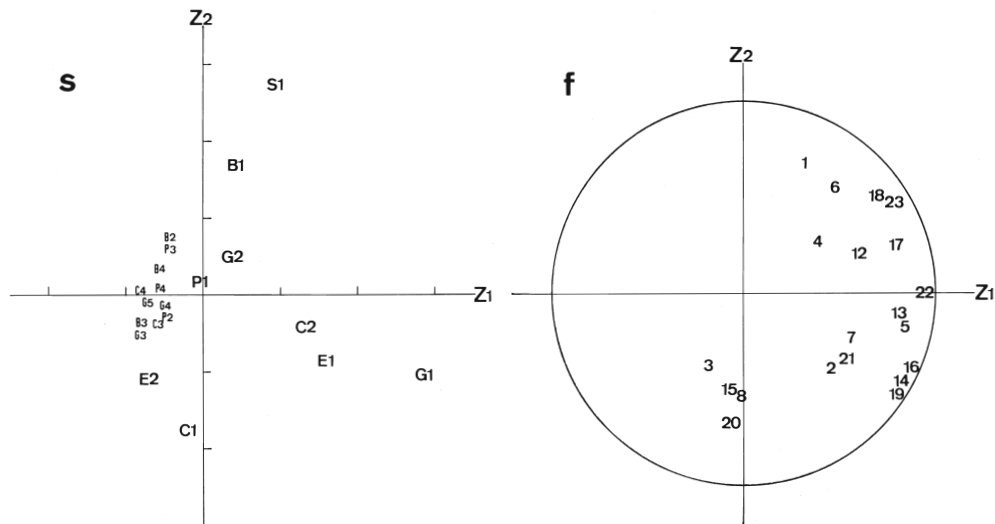
1). 分散共分散行列と相関行列

種を変数にとり、対数変換を行わない場合の結果を第6図に示す。分散共分散行列からの結果ではスコアにおいて極端にとびはなれた点の影響が強く出ていて、類別には不適當である。相関行列からの結果は比較的類別しやすいパターンになっている。調査点を変数にとつた場合も同様で、分散共分散行列からの結果は不適當である。相関行列からの結果を第7図に示す。



第6図 $PCA (V_{sp})$ (A) と $PCA (R_{sp})$ (B) のスコア (s) と因子負荷量 (f)

Fig. 6 Score (s) and factor loading (f) of $PCA (V_{sp})$ (A) and $PCA (R_{sp})$ (B).



第7図 PCA (R_{st}) のスコア (s) と因子負荷量 (f)
 Fig. 7. Score (s) and factor loading (f) of PCA (R_{st}).

相関行列はもとのデータを変数ごとに平均0, 分散1になるように, $X' = (X - \bar{X})/s$ で標準化した値の分散共分散行列に等しい. 標準化後の各要素の分布状態を第8図に示す. これによりすべての値が-1.5~4.0の間におさまっていることがわかる. 調査点の平均および標準偏差の変動係数の方が種のそれより小さいため, 調査点について標準化した場合の方が変換後の値の幅が小さくなっている.

2). 対数変換

対数変換 $\ln(1+n)$ したデータについて, 種を変数にとつた場合の結果を第9図に示す. いずれも類別しやすいパターンとなっており, 対数変換しない場合の相関行列からの結果とかなりの共通性がみられる. 調査点を変数にとつた場合も同様である. 対数変換後の各要素の分布状態を第10図に示す. これにより0~4.6の間に1対1対応で変換されていることがわかる.

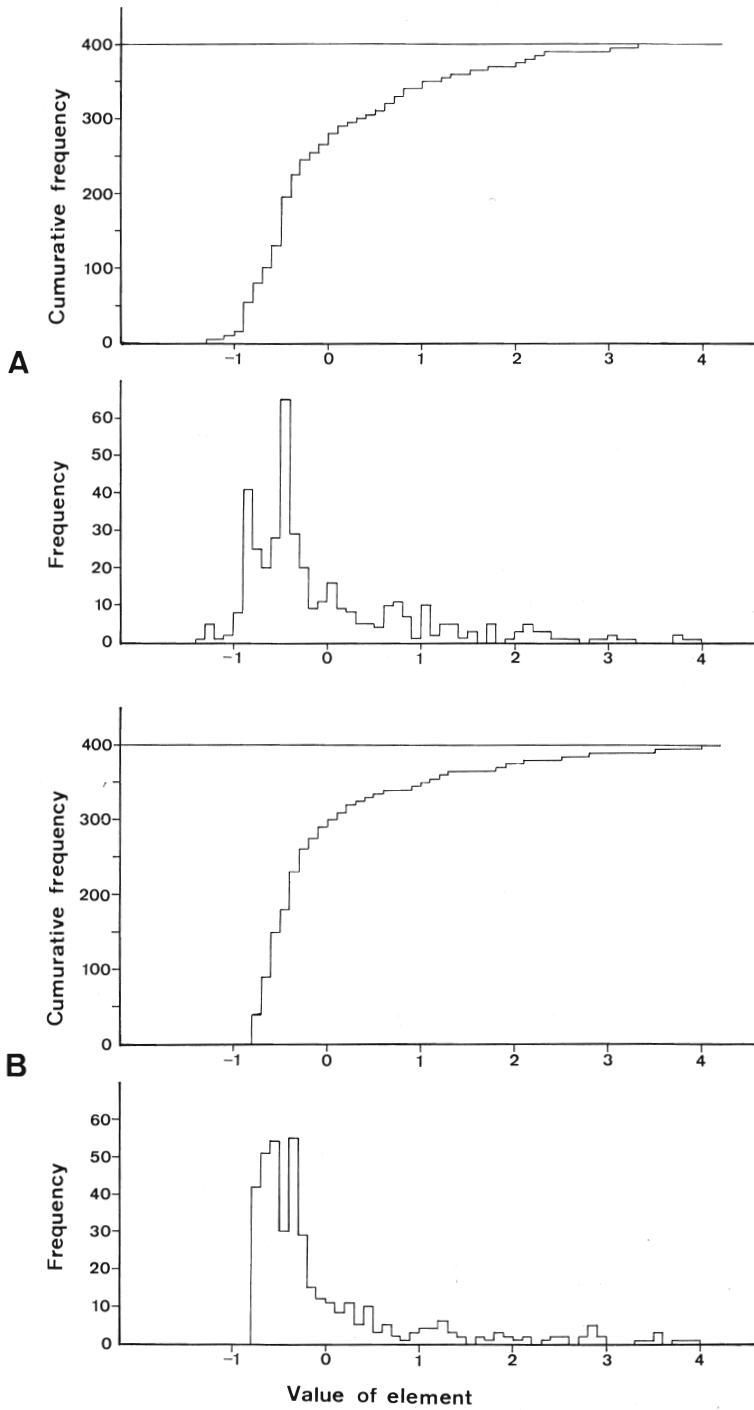
対数変換では大きな値を押さえる反面, 小さな値の差を大きく評価してしまう. 今回のデータのように小さな値が多く, しかも漁具効率が低い場合には, 小さな値の差を過大評価することはよくない. 今回の例では結果において著しい差はなかつたが, 相関行列からのスコアで見ると, 対数変換しない場合では採集個体数の多い St. 22 が, 少ない St. 1 より全体からはなれているのに対し, 対数変換した場合には逆転している. 対数変換は一般には有効な手法であるが, 今回のデータに関しては有効性に乏しいと判断して, 以後の論議にはすべて対数変換しない場合のみを用いた.

3). スコアと因子負荷量

一般の主成分分析では多数のサンプルを比較的少数の変数で解析するため, スコアによる類別がほとんどで, 因子負荷量による類別は少ない. 因子負荷量は相関係数なので採集個体数に関係しない. 逆にスコアでは採集個体数が直接影響する. 漁具効率が不安定な場合にはこのことは最大の欠点となる.

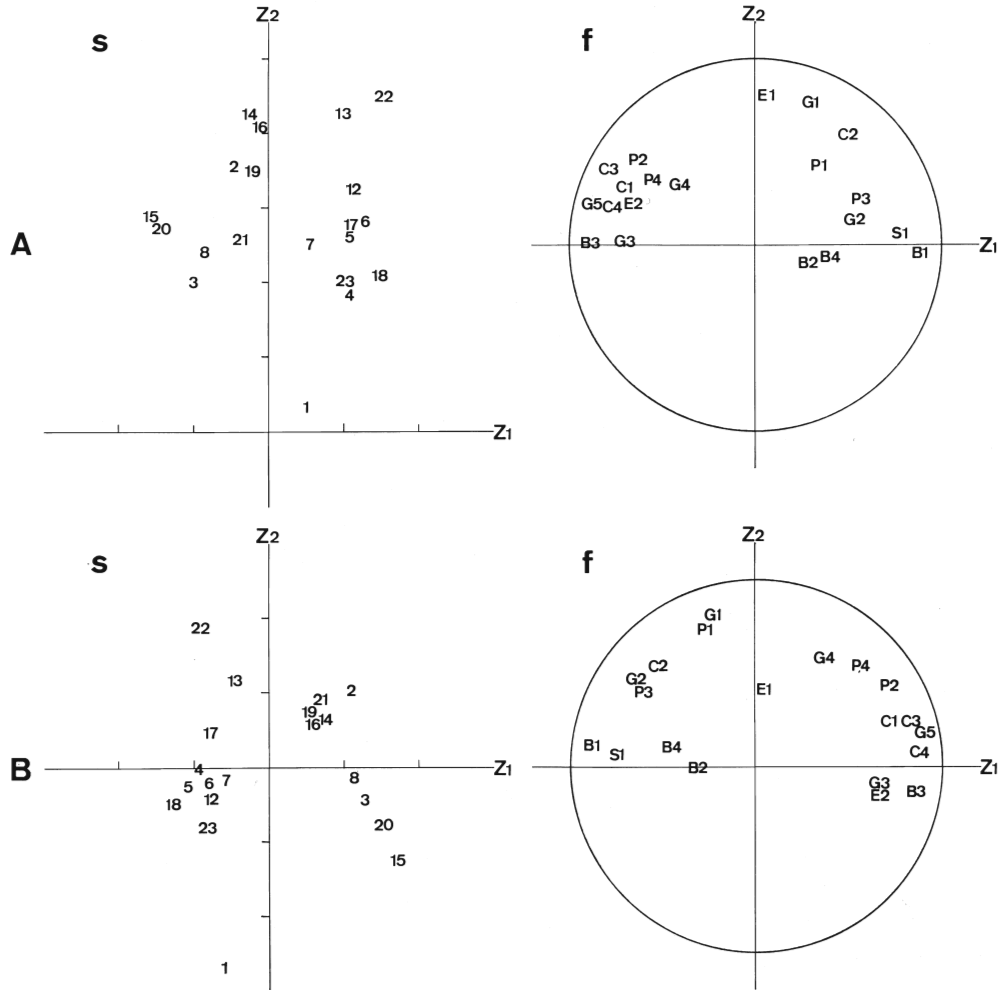
スコアのこの欠点を補うため0点の追加を行なった. $P_1 = (n_{1i}), P_2 = (n_{2i})$ ですべての i について $n_{2i} = hn_{1i}$ の場合を考える. 分散共分散行列を用いる場合には

$$z_{2k} = \sum_i a_{ki} n_{2i} = h \sum_i a_{ki} n_{1i} = h z_{1k}$$



第 8 図 標準化したデータの要素の累積度数分布曲線とヒストグラム。
(A) 種について標準化. (B) 調査点について標準化.

Fig. 8. Distribution curve of cumulative frequency and histogram of elements of standerzied data. (A) Standerzied about species. (B) Standerzied about stations.



第9図 対数変換したデータに行なつたPCA $V_{(sp)}$ (A)とPCA $R_{(sp)}$ (B)のスコア (s) と因子負荷量 (f)

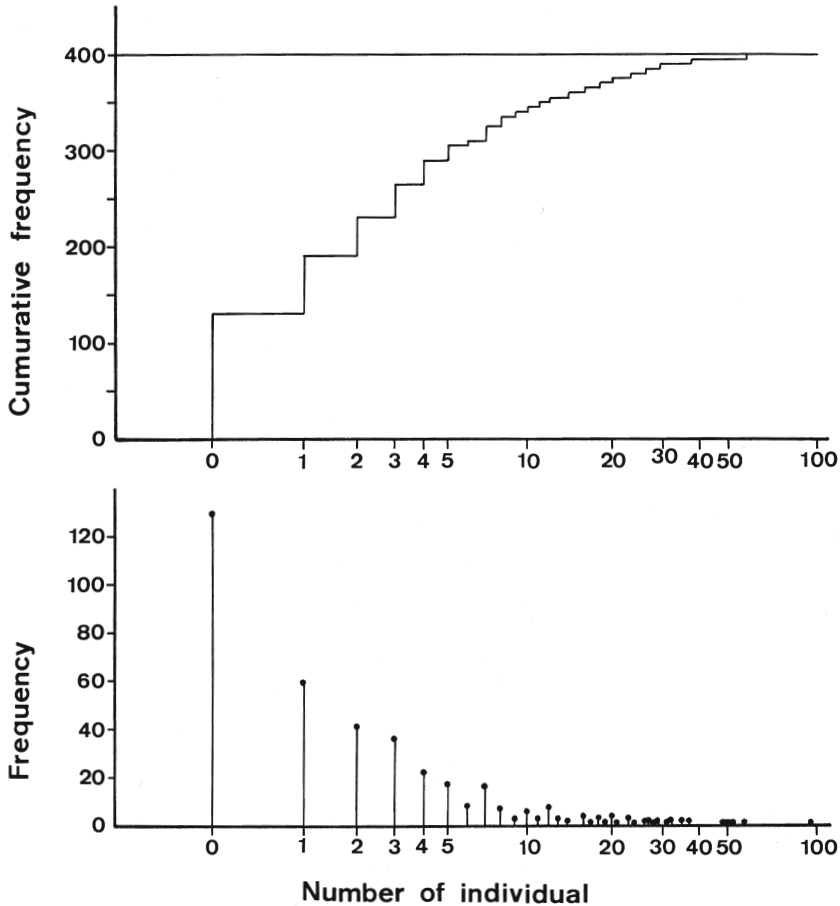
Fig. 9. Score (s) and factor loading (f) of PCA ($V_{(sp)}$) (A) and PCA ($R_{(sp)}$) (B) about transformed data into logarithm.

となり、2点 P_1, P_2 は Z_1-Z_2 平面上で原点0を通る直線上にある。相関行列を用いる場合には

$$z_{2k} = \sum_i a_{ki} \frac{n_{2i} - \bar{n}_i}{S_i} = \sum_i a_{ki} \frac{n_{2i}}{S_i} - \sum_i a_{ki} \frac{\bar{n}_i}{S_i}$$

なので、 $P_0 = (\sum_i a_{ki} \bar{n}_i / S_i)$ とおくと $P_a + P_0 = (\sum_i a_{ki} n_{ai} / S_i)$ となる。したがつて2点 $P_1 + P_0, P_2 + P_0$ は原点0を通る直線上にある。故に3点 $P_1, P_2, -P_0$ は同一直線上にある。ただし、 Z_1-Z_2 平面は n 次元空間の射影なので、同一直線上にあるからといって $n_{2i} = h n_{1i}$ の関係があるとは限らない。しかし、スコアの Z_1-Z_2 平面に0点として $-P_0 = (-\sum_i a_{ki} \bar{n}_i / S_i)$ を書き加えることにより、採集体数数の影響をある程度考慮することができる。

第6図と第7図のスコアに0点 ($-P_0$) を書き加えた例を第11図に示したが、St. 22, G1, S1, E1, C2等の採集体数が多いことがわかる。また $-P_0$ に近づけることにより、他の点と



第10図 対数変換したデータの要素の累積度数分布曲線と棒グラフ
 Fig. 10. Distribution curve of cumulative frequency and bar diagram of elements of transformed data into logarithm.

の関係も考察できる。

4). 寄 与 率

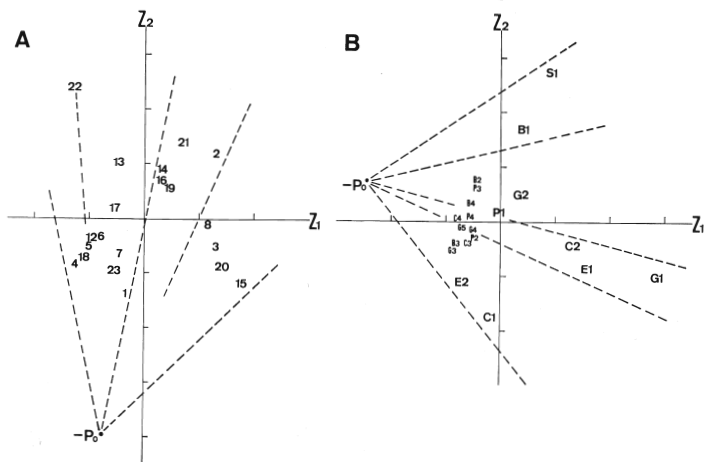
一般に主成分分析では寄与率が80%を越すまで成分をとる必要がある。寄与率を無視した次元の著しい縮小化は危険である。今回の結果では寄与率は Z_1, Z_2 で50~60%程度となっており、80%を越すには Z_3 くらいまで必要である。

Z_2 までで不十分な場合は、 Z_3 までとつて3次元表現することが多いが、理解し難く、環境要素の書き込み等にも不便である。したがって2次元表示において、寄与率の低さを補う手法が必要とされる。今回は、スコアでは標準ユークリッド距離 d による最小距離樹形図を採用した。因子負荷量では相関係数 r を使用した。

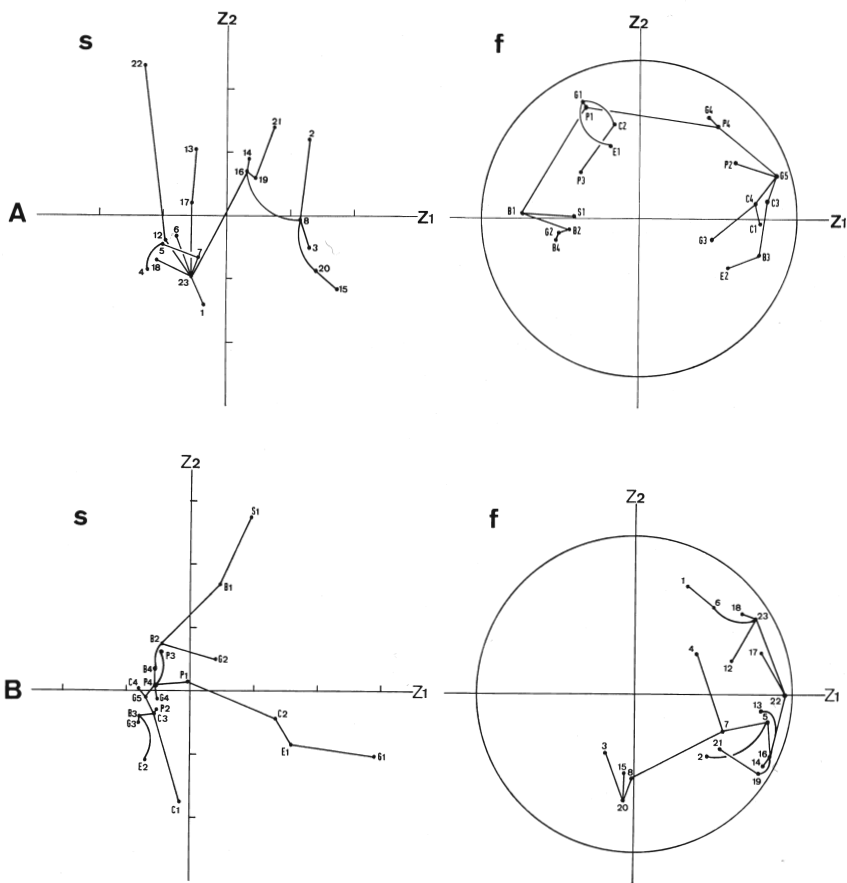
第6図と第7図に適用した例を第12図に示す。これにより2次元に射影したことによる折りたたみの程度を知ることができる。 d, r の値はデンドログラムの節に掲載している(第3表, 第4表)。

5). 変 数 の 選 択

調査点を変数にとる方法は各調査点ごとに標準化してしまうので、種の分布が他種の出現数によって歪められてしまう。逆に種を変数にとる方法では漁具効率が不安定な場合に問題があ



第11図 PCA (R_{sp}) (A) と PCA (R_{st}) (B) のスコアに加えた0点 ($-P_0$)
 Fig. 11. O-point ($-P_0$) on score of PCA (R_{sp}) (A) and PCA (R_{st}) (B).



第12図 PCA (R_{sp}) (A) と PCA (R_{st}) (B) のスコア (s) と因子負荷量 (f) の最小距離樹形図
 Fig. 12. Minimum spanning tree about score (s) and factor loading (f) of PCA (R_{sp}) (A) and PCA (R_{st}) (B).

る。従つて、現時点ではスコアと因子負荷量では前述のように表現が異なるので、可能であれば両方を行ない比較検討するのが最良の方法であると考える。

2. デンドログラム

1). デンドログラム

類似度指数 r_{st} , r_{sp} , d_{st} , d_{sp} , C_2 , C_{II} , PD , PS の値を第3～6表に示す。またそれぞれについて最短距離法により作成したデンドログラムを第13図に示す。

調査点についてみると r_{st} , C_2 , C_{II} は類似した結果となつており、沖合の4点 (St. 3, 8, 15, 20) がかなり明確に区別されている。 PS はこれらと同様の傾向を示しているが、あまり明確ではない。 PD は採集個体数の少ない St. 1 がとびはなれてしまつてゐる。 d_{st} は採集個体数の差を直接反映する類似度指数なので、これらとはかなりちがう結果となつてゐる。

種についてみると r_{sp} では C1, B3, E2 等の沖寄りに生息するグループ、 G2, B1 等の岸寄りに生息するグループ、そのどちらにも属さないものの3つに大別できる。 d_{sp} ではすべて同一のクラスターになつてしまい類別が困難である。

2). 類似度指数の検討

類似度指数間の主な相関図を第14図に示す。

(1) 相関係数 r

相関係数は類似度指数として広く使用されているが、指数としては不適當だと言われている。木元 (1976) によると“種ごとの個体数は正規分布しないため、相関係数はどんな母数の推定値にもならない”とされている。しかし、水戸 (1977) もすでに指摘しているが、類似度指数の妥当性を統計的推定の面から判断するのは不適當である。相関係数の問題点は、本来、相関係数は2変数間の線型関係を測る量なので負の値の解釈が類似度指数として困難な点と、(0, 0) の要素の影響が強い点の2点である。

2点間の類似性が低くなると共通種が減少し、片方だけ0の要素が多くなる。この場合相関係数は負の値をとる。しかし、相関係数が-1に近い場合は線型性が強く、共通種が多いことになる。これは類似性を正しく反映していない。実例についてみると相関係数は負の大きな値をとることはない。しかし、下限がはつきりしない点は欠点なので、正の小さい値や負の値の解釈には十分な注意が必要である。

今回のデータでは C_{II} とかなりよい相関を示し、デンドログラムもほぼ一致したので問題はないと考える。一方、最短距離法では最大値で判定するので、この欠点を補つていると判断される。

次に(0, 0)の要素の影響について考える。類似性が高い場合には問題はないが、低い場合には負の値となるので前述のように注意が必要である。しかし、注意すべき点は、(0, 0)の要素の影響が強いため選択する種によつて値が大きく変動する可能性が強い点である。 C_{II} や PS では(0, 0)の要素は値に影響しない。したがつて今回のような抽出データでも全個体数の75%を占めているため、全体像をよく反映していると推定される。しかし、相関係数ではこれらの指数と比較して種の選択の影響が強く、全体像をある程度歪めている点に注意する必要がある。

(2) PD と PS

PD , PS はよく似た指数であるが、 PD が直接個体数の差をとつてゐるのに対し、 PS は百分率の差をとつてゐる。したがつて PD は採集個体数の差の影響を強く受ける。今回のデータのように漁具効率が不安定なデータに対しては、 PS を使う方がよい。

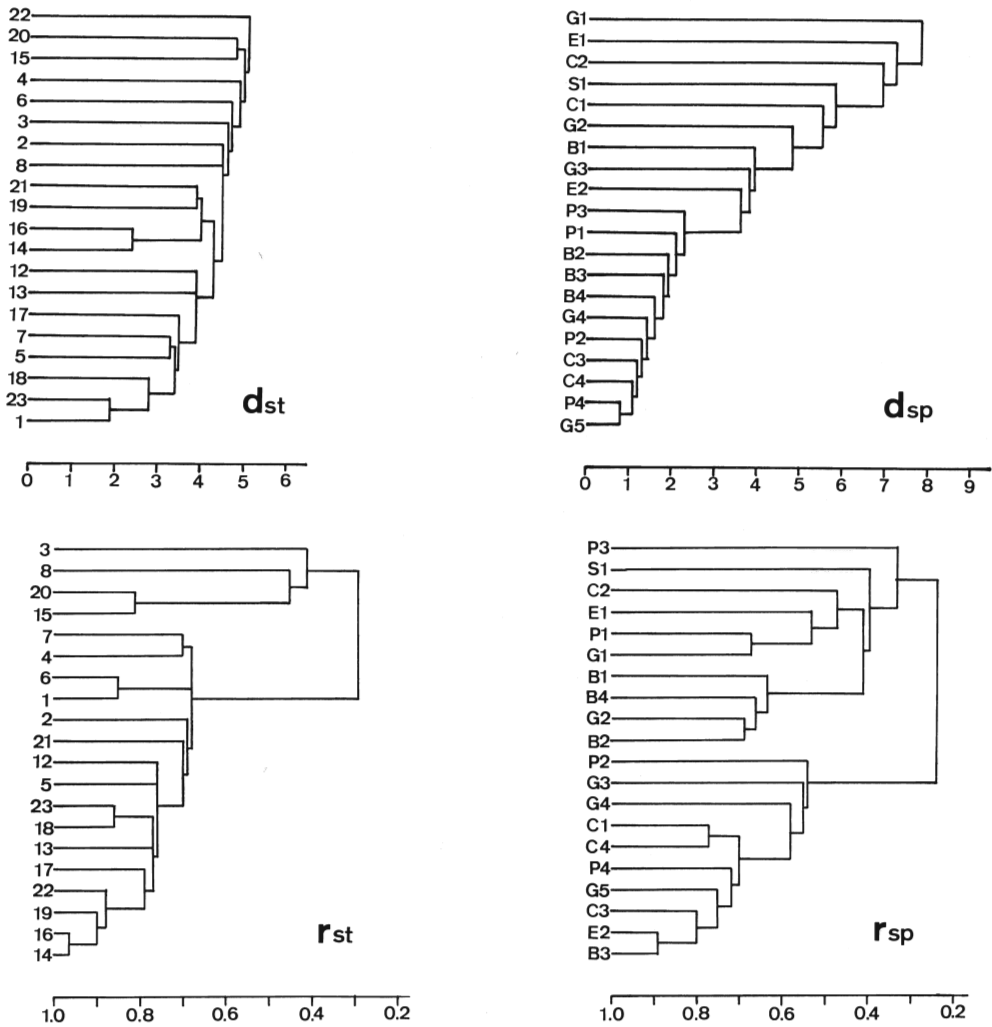
しかし、 PS , PD は相関係数や C_{II} ほど明確に類別できない。これは計算過程で積の要素が

第5表 C_λ と C_Π の類似マトリックス Table 5. Similarity matrix of C_λ and C_Π .

ST	1	2	3	4	5	6	7	8	12	13	14	15	16	17	18	19	20	21	22	23
1	1.253	.269	.088	.210	.253	.868	.174	.006	.185	.436	.114	.003	.178	.446	.523	.116	.012	.035	.357	.569
2	.321	1.054	.339	.423	.768	.163	.682	.472	.234	.775	.610	.218	.621	.429	.413	.725	.406	.488	.494	.352
3	.091	.357	1.053	.178	.174	.152	.330	.559	.106	.177	.362	.288	.206	.235	.140	.291	.592	.302	.173	.158
4	.250	.147	.188	1.060	.658	.234	.800	.189	.253	.380	.353	.058	.373	.527	.769	.405	.181	.435	.480	.610
5	.297	.807	.183	.694	1.049	.290	.762	.180	.394	.762	.790	.137	.821	.647	.681	.792	.249	.640	.770	.593
6	.972	.167	.155	.240	.297	1.010	.266	.119	.526	.395	.293	.069	.316	.630	.542	.249	.127	.292	.502	.651
7	.209	.734	.354	.862	.814	.275	1.097	.508	.273	.582	.566	.270	.577	.577	.626	.647	.447	.597	.573	.532
8	.007	.504	.596	.203	.192	.123	.353	1.081	.277	.306	.396	.378	.325	.306	.122	.452	.625	.397	.277	.230
12	.562	.243	.110	.263	.108	.535	.287	.291	1.027	.666	.662	.365	.660	.609	.544	.539	.237	.241	.787	.803
13	.518	.813	.186	.399	.826	.402	.622	.326	.689	1.043	.811	.292	.835	.701	.555	.794	.291	.434	.812	.634
14	.169	.636	.377	.369	.823	.298	.600	.408	.682	.842	1.034	.403	.975	.666	.464	.913	.456	.651	.882	.593
15	.003	.226	.300	.061	.112	.070	.285	.398	.375	.303	.416	1.030	.435	.185	.092	.334	.839	.220	.299	.209
16	.210	.650	.215	.392	.858	.322	.615	.345	.682	.870	1.012	.450	1.042	.691	.510	.919	.137	.683	.913	.619
17	.539	.462	.252	.568	.691	.650	.633	.333	.640	.719	.706	.195	.736	1.098	.779	.674	.316	.701	.859	.839
18	.624	.471	.148	.820	.721	.557	.679	.131	.568	.587	.487	.097	.539	.844	1.071	.484	.184	.522	.730	.908
19	.110	.779	.312	.436	.845	.257	.709	.492	.568	.848	.968	.353	.979	.739	.524	1.096	.460	.785	.839	.563
20	.014	.430	.626	.192	.263	.130	.483	.671	.247	.307	.478	.876	.460	.342	.197	.497	1.065	.438	.321	.233
21	.041	.522	.323	.467	.681	.301	.653	.430	.253	.462	.689	.232	.726	.766	.563	.857	.472	1.087	.694	.498
22	.420	.513	.179	.501	.799	.509	.607	.292	.807	.839	.908	.308	.944	.908	.765	.887	.335	.732	1.025	.845
23	.694	.386	.173	.669	.644	.680	.595	.255	.859	.691	.640	.224	.672	.940	1.002	.630	.258	.554	.912	1.142

第6表 PD と PS の類似マトリックス Table 6. Similarity matrix of PD and PS .

ST	1	2	3	4	5	6	7	8	12	13	14	15	16	17	18	19	20	21	22	23
1	.271	.054	.306	.313	.627	.176	.016	.397	.352	.161	.008	.190	.312	.370	.433	.135	.027	.027	.314	.397
2	.940	.373	.537	.268	.545	.510	.268	.596	.504	.350	.521	.350	.521	.370	.358	.617	.494	.509	.383	.368
3	.902	.530	.273	.174	.195	.381	.545	.208	.282	.377	.353	.344	.271	.173	.334	.553	.368	.224	.233	.233
4	.867	.627	.746	.335	.603	.219	.379	.403	.332	.123	.357	.176	.612	.226	.322	.461	.226	.322	.461	.568
5	.897	.472	.823	.457	.823	.362	.590	.246	.353	.600	.373	.148	.619	.531	.554	.555	.236	.460	.609	.522
6	.910	.740	.824	.689	.609	.366	.184	.414	.440	.360	.145	.349	.470	.591	.470	.345	.219	.346	.500	.543
7	.870	.500	.615	.439	.436	.703	.380	.310	.498	.470	.217	.496	.470	.479	.544	.531	.363	.472	.489	.456
8	.973	.479	.401	.788	.763	.802	.621	.802	.350	.369	.423	.389	.359	.359	.176	.493	.511	.491	.300	.305
12	.880	.720	.803	.636	.649	.571	.654	.784	.529	.485	.284	.284	.182	.551	.521	.376	.247	.232	.637	.710
13	.884	.400	.732	.636	.436	.573	.561	.686	.469	.630	.336	.665	.665	.636	.473	.632	.296	.390	.688	.580
14	.945	.487	.600	.708	.449	.607	.590	.632	.516	.364	.111	.871	.871	.509	.361	.415	.511	.511	.659	.459
15	.986	.648	.636	.877	.864	.847	.790	.602	.721	.669	.591	.412	.412	.211	.081	.412	.680	.269	.270	.244
16	.904	.483	.656	.660	.392	.606	.576	.611	.531	.328	.162	.581	.404	.523	.404	.740	.372	.529	.715	.500
17	.791	.622	.718	.514	.479	.494	.531	.645	.539	.354	.589	.809	.523	.511	.628	.548	.273	.511	.743	.699
18	.806	.606	.819	.378	.453	.627	.461	.838	.473	.513	.701	.917	.630	.366	.379	.379	.172	.352	.643	.750
19	.952	.414	.653	.624	.469	.669	.479	.507	.617	.145	.422	.670	.358	.444	.608	.412	.608	.638	.630	.489
20	.983	.504	.448	.767	.747	.777	.663	.448	.752	.711	.595	.347	.628	.711	.814	.565	.472	.489	.256	.268
21	.951	.626	.614	.657	.565	.685	.532	.518	.756	.571	.552	.756	.492	.480	.631	.361	.607	.662	.488	.368
22	.924	.626	.826	.618	.308	.531	.684	.809	.295	.403	.368	.784	.409	.519	.195	.624	.805	.662	.488	.685
23	.690	.711	.746	.517	.518	.613	.590	.681	.534	.569	.710	.823	.648	.317	.398	.593	.723	.672	.647	.647



第13-a 類似度指数によるデンドログラム
 Fig. 13-a. Dendrograms of similarity index.

ないためと考えられるが、本来もつと大きなスケールでの比較に用いるべき指数であると推測する。

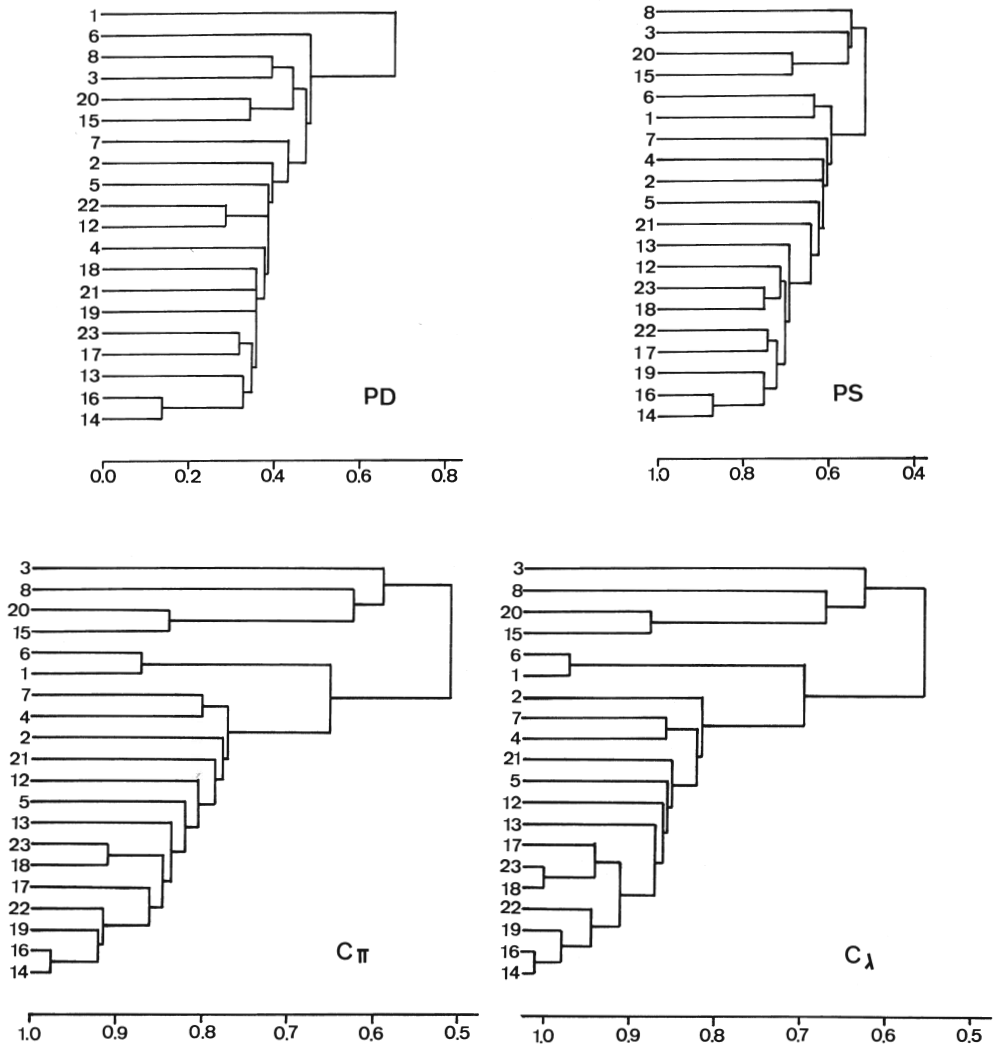
(3) C_λ と C_N

C_N は C_λ の修正式であり、1 より大きな値をとらない。また C_λ は採集個体数の影響を若干受けるが、 C_N は受けない。たとえば C_N では一方のデータの値をすべて10倍しても変化しない。したがって C_N に統一すべきであると考えられる。今回のデータでは相関図で判断する限りではもつとも明確に類別している。

(4) 標準ユークリッド距離 d

d は個体数の差を直接反映する指数なので、今回のデータのように漁具効率が不安定なデータには不適當である。

以上の論議から、類似度指数としては C_N と相関係数が適切であると結論できる。しかし、



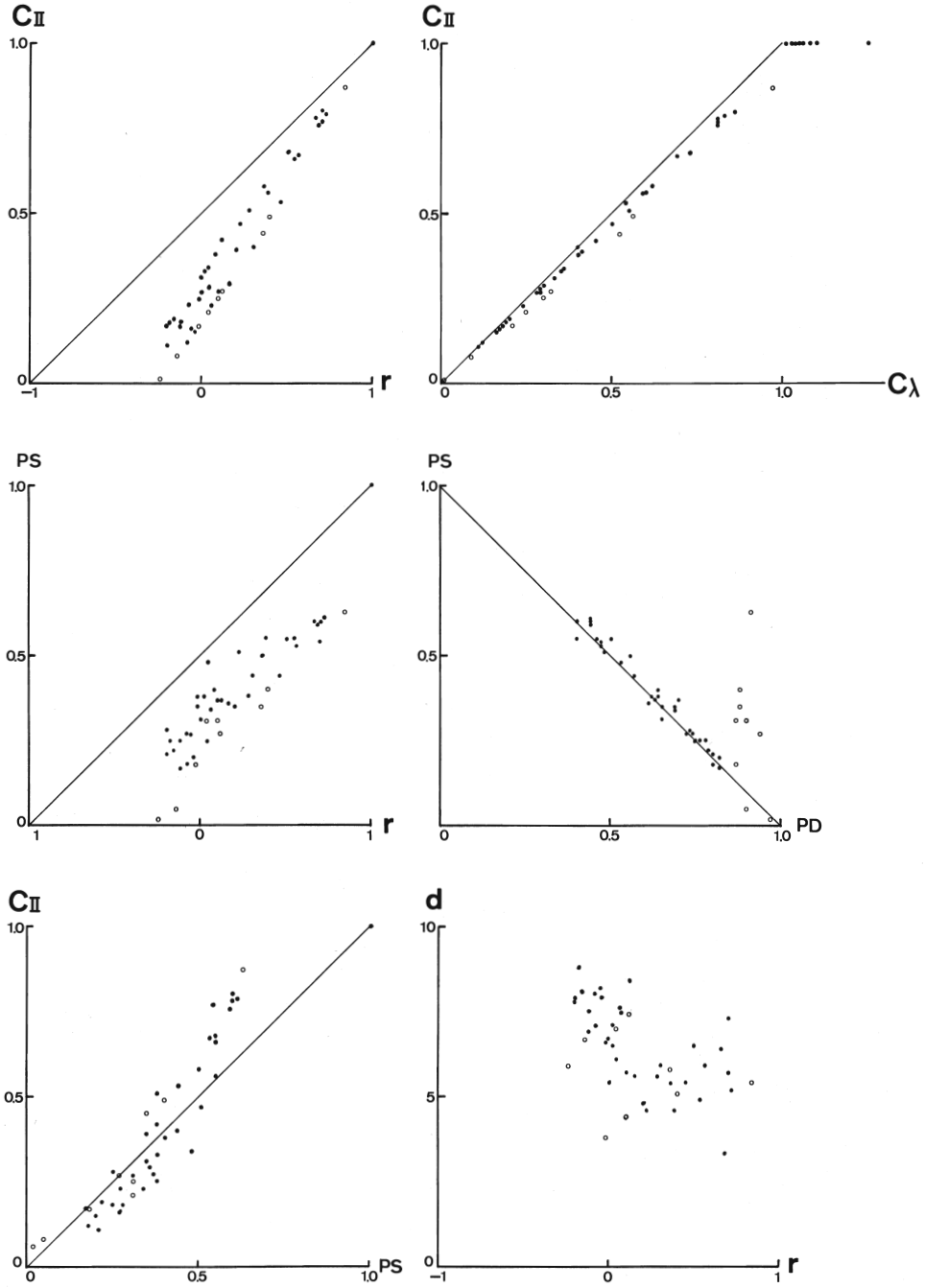
図第13-b 類似度指数によるデンドログラム

Fig. 13-b. Dendrograms of similarity index.

C_{II} では個体数はすべての種を通じて同等に扱われるので、採集個体数の多い種の寄与が大きい。これを改善するには、生物量に変換する、種ごとに基準化する等の手法が必要である。それらの点を全く考慮しない解析ではかえって相関係数の方が安全な指数であると考えられる。また種の分布についての類似度指数はあまり使用されていないが、第13図に示すように、相関係数は有効な指数である。

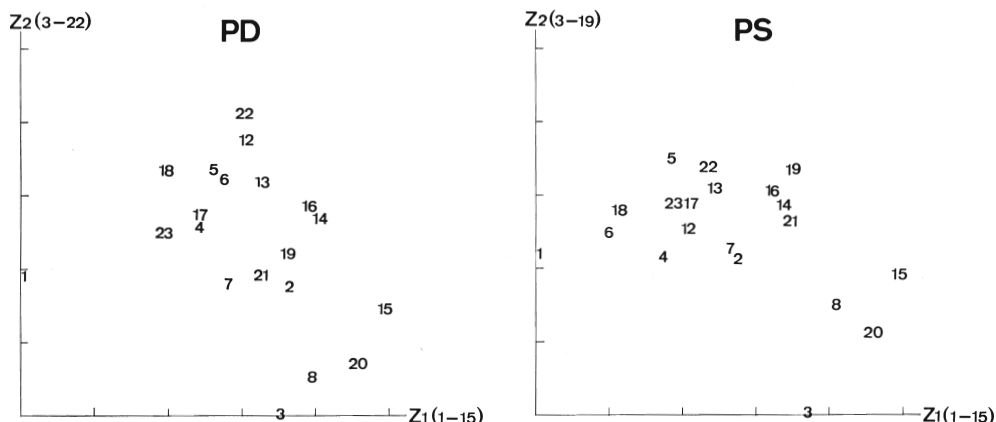
3. 極座標づけ

PD , PS のピタゴラス距離による極座標づけの結果を第15図に示す。 Z_1 の両端の2点はもつとも相違する2点を採用した。 PD , PS とも St. 1 と St. 15 であつた。 Z_2 の両端の2点は Whittaker (1979) に従つて、 Z_1 の座標が近くてかつもつとも相違する2点を採用したが、その際 Z_1 と相関が強くなる2点は棄却した。このようにして PD では St. 3 と St. 22 を、 PS では St. 3 と St. 19 を採用した。



第14図 類似度指数の相関図. 白丸: St. 1 と St. 1—13. 黒丸:
St. 2—13 と St. 2—13

Fig. 14. Correlation diagrams of similarity indices. Whits circles: St.1 and
St. 1-13. Black circles: St.2-13 and St. 2-13.



第15図 PD と PS の極座標づけ
 Fig. 15. Polar ordination of PD and PS.

PD, PS とも 3 ~ 4 のグループに分けることができ、主成分分析やデンドログラムの結果と同様の傾向を示している。しかし、全体的にみて Z_1 と Z_2 に負の相関がみられる。

極座標づけは類似度指数を使用して多次元表示できる点で、主成分分析とデンドログラムの中間に位置する手法である。今回は PD, PS を使用したが、 C_{II} 等の類似度指数を使用すればさらに良い結果が得られると推察される。しかし、 Z_2 の両端の 2 点の選び方に客観性が欠け、しかも Z_1 と無相関になるように選ぶことが困難である。また、主成分分析とちがつて寄与率が与えられない点も欠点である。2次元表示により最小距離樹形図や環境要素の書き込み等の手法が可能になるが、使用の際には上記の欠点に充分注意する必要がある。

4. 環境要素の書き込み

主成分分析の結果の調査点の位置に環境要素（水深、中央粒径値、強熱減量）の値を書き込み、等量線を描いた結果を第16図に示す。

水深は種を変数にとつた場合では Z_1 と相関が強く、 Z_2 とはほとんど相関がない。調査点を変数にとつた場合では Z_1, Z_2 ともに相関がある。そしてこれに対応する種の図からそれぞれの種の分布水深が推定できる。結果はデンドログラムの結果とほぼ同様であつた。

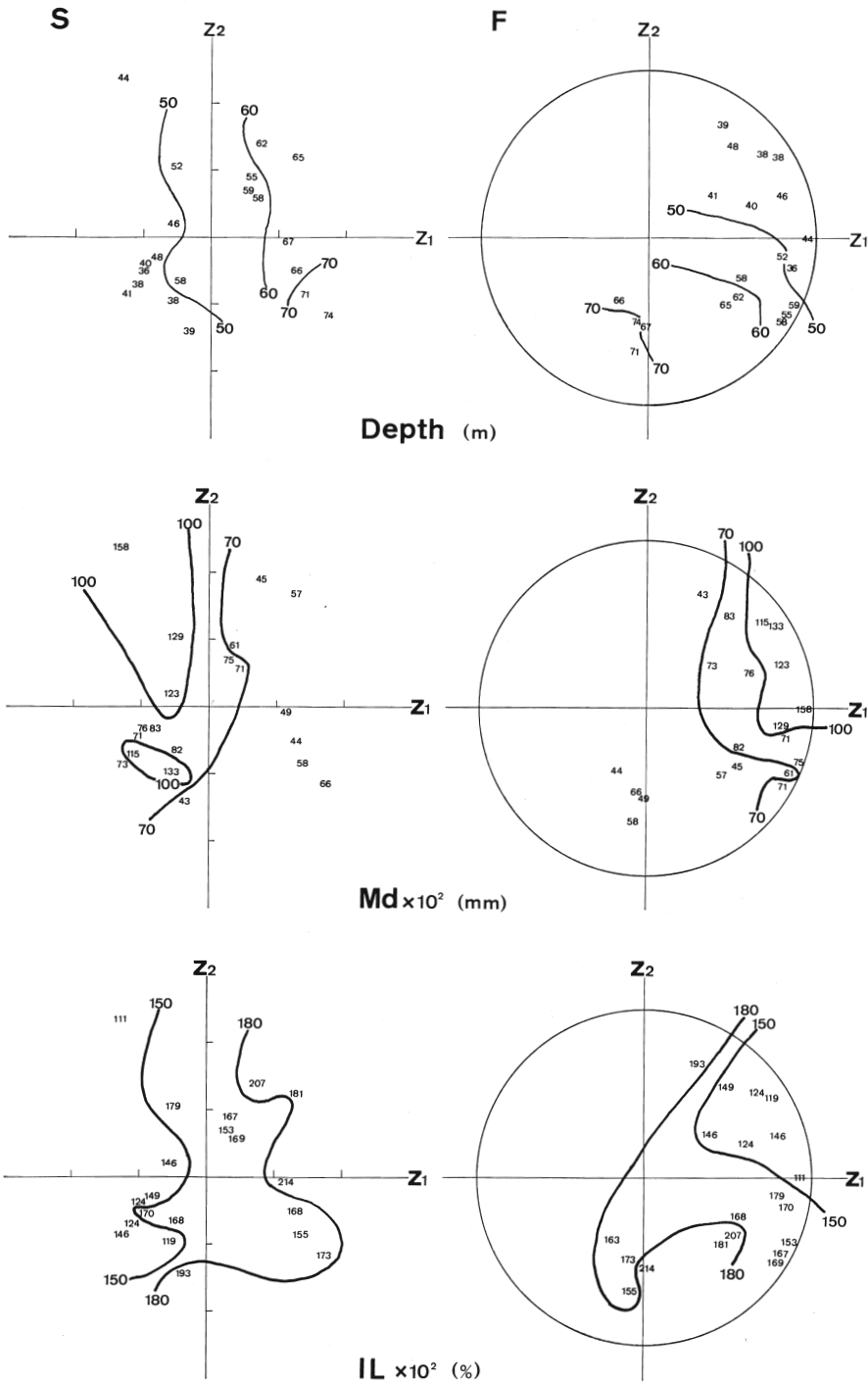
中央粒径値と強熱減量の図についても水深と同様の傾向を示しているが、水深のように単純なパターンとはなっていない。

この手法により環境要因の判定が容易にできる。今回のデータでは水深、中央粒径値、強熱減量の 3 つの環境要素のうち水深が最も影響力が強いと推察できた。

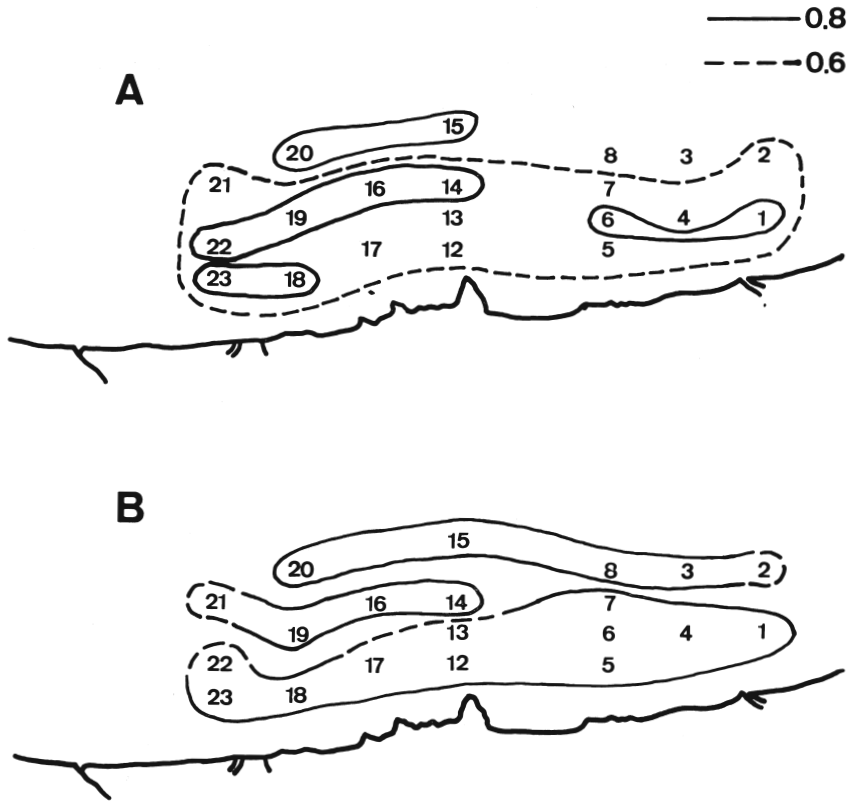
相関係数のデンドログラムと主成分分析のそれぞれのクラスターを地図に書き込んだものを第17図に示す。これは基本的な手法で第2図や第16図と比較することにより互いに補足できる。

IV. 要 約

1981年7月、鳥取県沿岸イタヤガイ漁場で行なつた貝桁網調査の採集個体数のデータに種々の統計的手法をほどこして、種および調査点の類別を検討した。手法として適当であると判断したのは、相関行列からの主成分分析で、スコアでは 0 点 ($-P_0$) の追加、最小距離樹形図の作成等が必要であるが、環境要素の書き込みが可能で全体像を容易につかむことができる。



第16図 PCA (R_{sp}) のスコア(S)と PCA (R_{st}) の因子負荷量(F)に書き込んだ環境要素の値の分布
Fig. 16. Distribution of condition factors on score of PCA (R_{sp}) (S) and factor loading of PCA (R_{st}) (F).



第17図 r_{st} のデンドログラムによる調査点のクラスター (A) と, PCA (R_{sp}) のスコアによる調査点のクラスター (B)

Fig. 17. Clusters of station on dendrogram of r_{st} (A) and score of PCA (R_{sp}) (B).

データによっては対数変換も有効であると考えられる。

最短距離法によるデンドログラムも有効で、類似度指数としては C_{II} と相関係数がよい。 C_{II} は全データに使用できるので、種の選択による偏りを防止するのに有効である。しかし、種ごとの寄与に差がある。

生物相は水深にそつて分布しており、今回の結果は外海に面した平坦な砂浜域の特徴であると考えられる。

本文を終えるにあたり調査に御助力いただいた日本海区水産研究所浅海開発部第3研究室田中邦三室長、中西孝技官、興石裕一技官、みずほ丸吉田伸夫船長はじめ乗組員の方々に深謝します。

文 献

- 石間紀男 (1969). 主成分分析. 農林研究計算センター報告A, (5) : 75-92.
- Jaffers, John N. R. (1982). 生態学のためのシステム分析入門 (古在豊樹, 高倉直, 仁科弘重訳). 朝倉書店, 東京, 209pp.
- 木元新作 (1976). 動物群集研究法 I, 多様性と種類組成. 生態学研究法講座, 14, 共立出版株式会社, 東京, 192pp.

- 鉄 健司 (1977). 相模湾内定置網漁獲物の組成に関する統計的考察. 東海水研報告 (89): 1—15.
- 水戸啓一 (1977). ベーリング海底魚群集における食物関係. I. 群集構造および魚種の分布パターンについて. 北大水産北洋研業績集特別号: 205-258.
- 大森迪夫 (1981). 油谷湾における魚類分布の統計的解析, マダイの生態的地位解明のための一試み (予報). 昭和55年度漁業資源研究会議西日本底魚部会会議報告: 56-68.
- Whittaker R. H. (1979). 生態学概説, 生物群集と生態系 (第2版) (宝月欣二訳). 培風館, 東京, 363pp.