

# 動物の種—個体数組成の類似度にもとづく 多変量解析法とその適用例

品川 汐夫

## 1. 緒言

海域環境調査の一環として行われる底生動物調査や漁獲物組成調査の資料では、一般に調査地点数よりも出現した動物の種類数の方がはるかに大きい。また個体数は種によって大きな差があり、1地点における種—個体数組成は、大量に出現する少数の種（優占種）と、ごく僅か出現したり、しなかったりする多数の種からなるのが普通である。動物の個体数の変動は、理化学的環境要因だけではなく、種間、種内の生物学的要因が複雑に作用した結果であるから、極めて非線型性が強く、加えて分布のパッチ性のため、採集誤差による変動も大きい。

これらの理由により、動物の種—個体数組成データの解析には、変量間の直線関係を想定した主成分分析法は適用できず、また、度数データの解析に広く応用される双対尺度解析などの数量化法も適用し難い。

データの性質に制約を必要としない解析の方法としてクラスター解析があり、動物群集の解析にもよく用いられている。この場合、対象間の類似度または汎距離をどのように定義するかが、結果を左右する。

動物群集の解析では、類似度として Kimoto<sup>1)</sup>の群集間類似度  $C_{\pi}$  を用い、群平均法によりデンドログラムを作成する方法が、一般的に行われている。しかし  $C_{\pi}$  では、少数種の分布が多数種に比べて軽視される事や、1地点における総個体数の相違が考慮されない事などの欠点があり、またデンドログラムの作成の過程では、所属のあいまいな地点が悪影響を及ぼすことがある。

そこで著者は、 $C_{\pi}$  を修正した相関指数  $R_{\pi}$  を新たに定義し、さらに関係類似度  $R_{\pi}$  を導入して、関係類似度行列の固有値分解と、フェジイクラスタリング<sup>2)</sup>の手法を取り入れた解析方法を開発した<sup>3)</sup>。これは、関係類似度行列を固有値分解することにより誤差の成分を除き、情報を集約できるという点で、主成分分析と同等の効果をもたらす解析方法であり、そのことは、乱数を用いた人工データによる数値実験で検証された。

## 2. 解析方法

### 2・1. 地点間類似度、相関指数、関係類似度の定義

2・1・1. 地点間類似度：2つの地点  $X$  と  $Y$  における種-個体数組成をベクトル  $x$  と  $y$  によって次式で表す。

$${}^t x = [x_1, x_2, x_3, \dots, x_S]$$

$${}^t y = [y_1, y_2, y_3, \dots, y_S]$$

$S$ ：全地点で出現した種類数

但し添字  $t$  は列ベクトルを転置した行ベクトルであることを示す。

次に平均ベクトル  $\bar{x}$  と  $\bar{y}$  を次式で定義する。

$${}^t \bar{x} = [\bar{x}, \bar{x}, \dots, \bar{x}], \quad \bar{x} = \frac{1}{S} \sum_i x_i$$

$\bar{y}$  も同様

$S = 2$  の場合、 $x$ 、 $\bar{x}$ 、 $y$ 、 $\bar{y}$  は図-1 に示す関係にある。そこで  $X$  と  $Y$  の類似性を、次の3つの要因に分解できる。

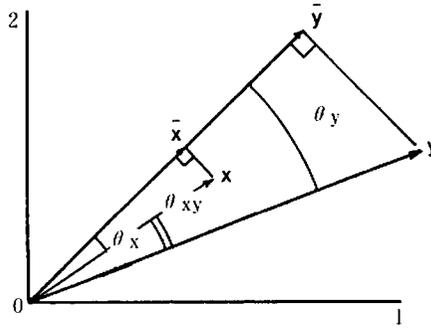


図1.  $S = 2$  の場合の  $X$  と  $Y$  の関係

- 1)  $x$  と  $y$  のなす角度  $\theta_{xy}$
  - 2)  $\bar{x}$  と  $\bar{y}$  の長さ  $\|\bar{x}\|$ 、 $\|\bar{y}\|$
  - 3)  $x$  と  $\bar{x}$ 、 $y$  と  $\bar{y}$  のなす角度  $\theta_x$ 、 $\theta_y$
- これら3つの要因を考慮した類似度として次式を定義できる。

$$C_{\pi'} = \epsilon_1 \epsilon_2 \gamma$$

但し  $\gamma = \cos \theta_{xy} = {}^t x \cdot y / \|x\| \|y\|$

$$\epsilon_1 = 2 d_x d_y / (d_x^2 + d_y^2)$$

$$d_x = \cos \theta_x = \|\bar{x}\| / \|x\|, \quad d_y \text{ も同様}$$

$$\begin{aligned} \epsilon_2 &= 2 \|\bar{x}\| \|\bar{y}\| / (\|\bar{x}\|^2 + \|\bar{y}\|^2) \\ &= 2 N_x N_y / (N_x^2 + N_y^2) \end{aligned}$$

$$N_x = \sum_i x_i, \quad N_y \text{ も同様}$$

これらは  $S > 2$  の場合にも拡張され、一般の類似度として用いられる。 $\epsilon_1 \gamma$  は Kimoto の  $C_{\pi}$  に一致する。

2・1・2. 地点間相関指数  $R_s$ ：図-1 から分かるように、 $C_{\pi'}$  は個体数が大きい種

の影響を大きく受ける。そのような影響を除くために、第  $i$  種の、出現した地点についての個体数の平均を  $m_i$  とし、 $\alpha_i = x_i / m_i$ 、 $\beta_i = y_i / m_i$ 、 $\bar{\alpha} = \frac{1}{S} \sum_i \alpha_i$ 、 $\bar{\beta} = \frac{1}{S} \sum_i \beta_i$ 、 $\overline{\alpha \beta} = \frac{1}{S} \sum_i \alpha_i \beta_i$  とする。

このとき、 $S$  が十分大きければ、 $X$  と  $Y$  に何らかの関連性があるとき  $\overline{\alpha \beta} > \bar{\alpha} \bar{\beta}$  または  $\overline{\alpha \beta} < \bar{\alpha} \bar{\beta}$  となり、関連性がなければ  $\overline{\alpha \beta} = \bar{\alpha} \bar{\beta}$  となることが期待される。

そこで、地点間相関指数  $R_s$  を次式で定義する。

$$R_s = \epsilon_1 \epsilon_2 \Delta / |1 - \epsilon_1 \epsilon_2 (\gamma - \Delta)| \quad \text{for } \Delta \geq 0$$

$$= \Delta / (\gamma - \Delta) \quad \text{for } \Delta < 0$$

但し、

$$\Delta = S m^2 (\overline{\alpha \beta} - \bar{\alpha} \bar{\beta}) / \| \mathbf{x} \| \| \mathbf{y} \|$$

$$m^2 = \frac{1}{S} \sum_i m_i^2$$

$R_s$  には次のような性質がある。

- 1) 種-個体数組成が全く等しいときにだけ 1 となり、共通種が全くないときにだけ -1 となる。
- 2) その中間では、符号は  $\Delta$  と同じで、同じ  $\Delta$  の値に対しては、 $\gamma$  が大きいほど大きく、また同じ  $\gamma$  の値に対しては、 $\Delta$  の絶対値が大きいほど絶対値が大きい。
- 3)  $\alpha_i$  と  $\beta_i$  が互いに独立な確率分布をするとき、期待値が 0 となる。

2・1・3. 関係類似度  $R_{sn}$  : 地点間相関指数行列  $[R_s]$  の第  $i$  列ベクトルを  $r_i$  とすると、 $r_i$  は地点  $i$  と他の地点との関係を表している。そこで 2 地点  $i$  と  $j$  の、他の地点との関係の類似度として、次式を定義する。

$$R_{snij} = r_i \cdot r_j / \| r_i \| \| r_j \|$$

$R_{snij}$  を成分とする行列を  $[R_{sn}]$  とすると、

$$[R_{sn}] = [N] [R_s]^2 [N]$$

但し  $[N]$  は  $1 / \| r_i \|$  を第  $i$  対角成分とする対角行列である。この式から、 $[R_{sn}]$  は正定値 (全ての固有値が正) であることが分かる。

底生動物の個体数データは、僅かな採集条件の違いなどにより大きく変動することがあり、それが数理解析を困難にする理由の一つである。関係類似度を用いることにより、偶然性の影響を小さくし、後述する情報集約率を高めることができる。

## 2・2. 関係類似度行列の固有値分解による共通要因の抽出と地点の座標空間への配置

$[R_{sn}]$  は正定値対称行列であるから、その固有値を  $\lambda_i$ 、対応する長さ 1 の固有ベクトルを  $e_i$  とすると、次のように分解される。

$$[R_{sn}] = \sum_{i=1}^r \lambda_i e_i \cdot e_i, \quad \lambda_i > 0$$

$r$  は  $[R_{sn}]$  のランク数で、 $n \times n$  個の関係の原因となる要因数といえる。実際には、地点

の独自要因（誤差）があるからランク数は地点数に等しい。そこで  $\lambda_i$  の大きい順に適当な  $r$  の値をとれば、 $[R_{ij}]$  から独自要因を除いて、共通要因を取り出すことができる。

このとき、 $\sum_{i=1}^r \lambda_i \leq n$  であるから、 $\sum_{i=1}^r \lambda_i / n$  は情報集約率で、共通要因の大きさの目安とできる。また、ベクトル  $\lambda_i \mathbf{e}_i$  の第  $j$  成分は、第  $j$  地点における第  $i$  要因の大きさを表していると考えられるから、これを第  $i$  成分とするベクトルを、地点  $j$  の要因ベクトルとして、第  $j$  地点を  $r$  次元の座標をもつ点に対応させることができる。

### 2・3. 出現種の座標空間への配置

第  $j$  種の、 $n$  地点における出現個体数を、 $n$  次元のベクトル  $Z_j$  で表す。このとき、第  $j$  種と第  $i$  要因との関連度  $r_{ij}$  を次式で定義する。

$$r_{ij} = |Z_j \cdot \mathbf{e}_i| / \|Z_j\|$$

これにより、第  $j$  種を  $r$  次元の座標をもつ点に対応させることができる。

### 2・4. 地点と出現種のファジイクラスタリング

以上のようにして、地点と出現種を共通要因の空間内の点に対応させることができるから、この空間内での距離を用いて、地点と出現種をクラスタリングできる。

群平均法によるクラスタリングは、最も近い距離にある点を1つのクラスターにまとめ、クラスター間の距離はクラスターを構成する要素点間の平均距離として、順次クラスターを形成するもので、要素点はあるクラスターに所属するか否かのどちらかである。

これに対して、ファジイ  $k$ -means 法によるクラスタリングは、予め定めた数のクラスターへの要素点の所属率を、クラスターの重心にある要素点の所属率が1、重心から遠ざかる程0に近くなるようにして、クラスターとしてのまとまり（クラスター内分散）が最適となるクラスターの重心と要素点の所属率を、繰り返し計算によって求めるものである。

この場合、クラスター数は与えなければならないが、群平均法によるデンドログラムを参考として、何通りかの試算を行い、クラスターのまとまりを表す尺度の分割度が最大となる計算結果を選べばよい。

この方法によると、異なった環境の境界にあるような、中間的性質の地点も正しく評価され、クラスター内での地点の順序付けや、クラスター間の重複度も計算できるから、環境要因の空間的拡がりなど、より多くの情報を得ることができる。

こうして地点と出現種をいくつかの群にクラスタリングすれば、出現種の分布と各地点の種-個体数組成の特徴が明らかとなり、これと種群を構成する個々の種に関する生態学的知見を併わせれば、調査域の環境について、いろいろな考察が可能となる。

また、出現種の個体数を各種群への所属率に応じて合計すれば、その個体数データの確率分布は正規分布に近づくと考えられるから、理化学的環境データとの正準相関分析や、回帰分析などの適用も可能となる。

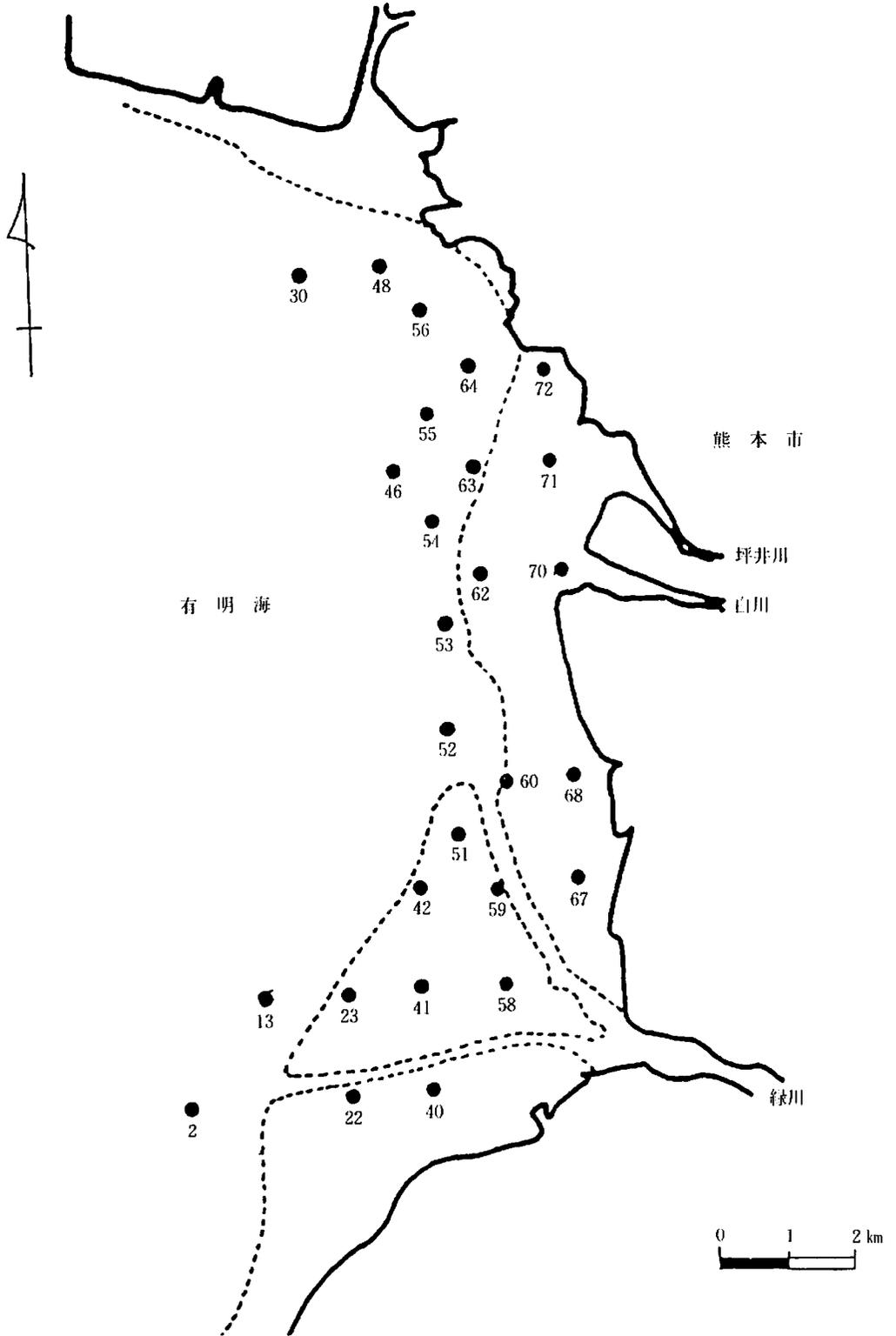


図2. 底生動物の調査地点

### 3. 底生動物調査結果への適用例

上に述べた解析方法を、干潟の底生動物調査結果に適用した例を次に示す。この調査は、熊本開発研究センターにより、平成元年8月に熊本港周辺の干潟の27地点で行われたものである。調査地点を図2に示す。

#### 3・1. 関係類似度行列の固有値分解

出現した動物の中から、複数種を含んでいると思われるものを除いて、環形動物54種、節足動物40種、軟体動物24種、棘皮動物1種を用いて関係類似度行列を計算し、これを固有値分解した。固有値を図3に示す。第4固有値までで情報集約率は82.2%と、かなり高い値となった。そこで第4成分までを以下の解析に用いた。

#### 3・2. 地点のクラスタリング

4次元の要因ベクトルにより地点間のユークリッド距離を計算し、群平均法によるデンドログラムを作成した(図4)。また、これを参考として、クラスター数を3~5としたファジイクラスタリングを行った結果、クラスター数を3とした場合が最も分割度が大きくなった。クラスターをA、B、C区とし、各地点がそれらに所属する割合を図5に示す。

さらに、所属率から計算されたA、B、C区の重複度行列を表1に示す。

デンドログラムとファジイクラスタリングの結果を比較すると、各区の境界に位置する地点13、40、52、53で相違がみられる。

表1. 地点群 A、B、C 区の重複度行列  
(分割係数0.6940、分割度0.5411)

	A 区	B 区	C 区
A 区	0.1647	0.0440	0.0733
B 区	0.0440	0.2505	0.0357
C 区	0.0733	0.0357	0.2788

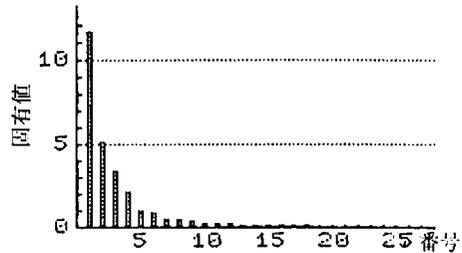


図3. 地点間関係類似度行列の固有値

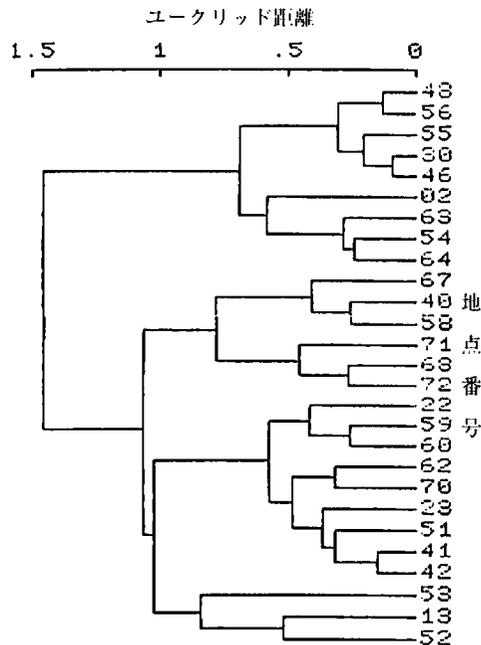


図4. 群平均法による地点のデンドログラム

デンドログラムによれば、地点13, 52, 53が1つのクラスターのようにみえるが、ファジイクラスタリングによれば、地点13は40とともにA区とC区の中間的性質で、地点52はA区の割合が大きい。

重複度行列から、B区とC区が最も異なっており、A区とC区は重複が大きいことが分かる。

### 3・3. 主要出現種のクラスタリングと種群の分布

全調査地点を通じて2個体以上出現した94種をSP-A, B, C, Dへの4群に分割するファジイクラスタリングを行った。結果を表2に、また重複度行列を表3に示す。

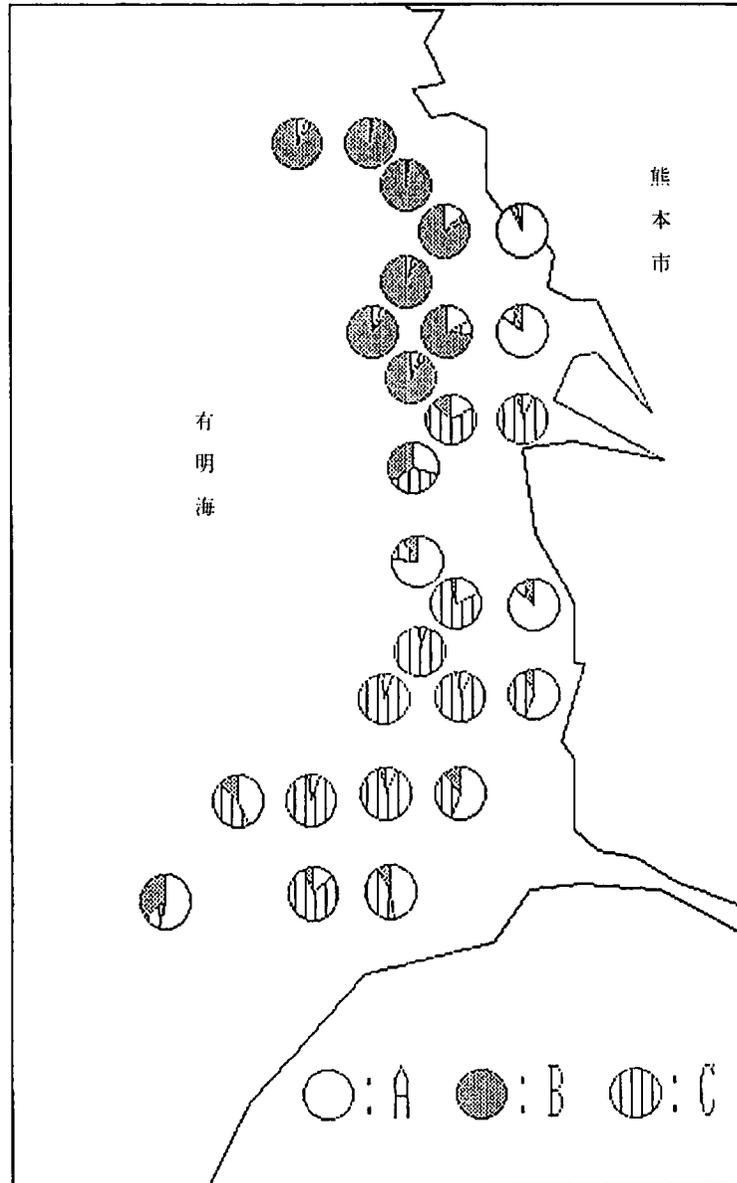


図5. ファジイk-means法による地点のクラスタリング  
(地点がA, B, Cに所属する割合を示す)

表2. 出現種のファジイ  $k$ -means 法によるクラスタリング  
 (数字は各々の種が種群 SP-A, B, C, D に所属する割合を示す)

SP-A	所属率	SP-B	所属率
ウミホタルモドキ	0.8389	<i>Lumbrineris longifolia</i>	0.9590
ドロヨコエビ	0.8354	<i>Mediomastus</i> sp.	0.8731
<i>Ampelisca bocki</i>	0.8347	<i>Harmothoe</i> sp.	0.8564
<i>Prionospio bocki</i>	0.7965	イヨスダレ	0.8531
イソチロリ	0.7570	<i>Brada villosa</i>	0.8208
<i>Prionospio matmgreni</i>	0.7414	ヨコナガモドキ	0.7665
<i>Mutilus</i> sp.	0.7183	セマタコスエモノガイ	0.7652
<i>Pseudopolydora kempi japonica</i>	0.6565	アリアケケホリガイ	0.7530
<i>Euchone</i> sp.	0.6343	タマガシフサゴカイ	0.7451
<i>Photis</i> sp.	0.5956	<i>Leonnates persica</i>	0.7339
<i>Magelona japonica</i>	0.5708	<i>Praxillella pratermissa</i>	0.7217
ミナミシロガネゴカイ	0.5578	ユビナガホンヤドカリ	0.6800
マサゴウロコムシ	0.5464	シズクガイ	0.6527
ホソツツムシ	0.5404	<i>Iphinoe sagamiensis</i>	0.6469
<i>Heteromastus</i> sp.	0.4876	トウヨウムツアシガニ	0.6444
ケシトリガイ	0.4864	ダルマゴカイ	0.6256
マダラサシバ	0.4783	棘皮類	0.5608
ハナオカカギゴカイ	0.4675	<i>Glycinde</i> sp.	0.4613
<i>Magelona</i> sp.	0.4580	ヒメカノコアサリ	0.4521
ヨコヤマキセワタ	0.4400	チヨノハナガイ	0.4367
<i>Chaetozone setosa</i>	0.4376	ヨコヤマキセワタ	0.4191
<i>Neomysis japonica</i>	0.4275	<i>Chaetozone setosa</i>	0.4003
<i>Pectinaria okudai</i>	0.4169	<i>Apscudes</i> sp.	0.3939
チヨノハナガイ	0.3933	ホソツツムシ	0.3342
<i>Paraonis gracilis minuta</i>	0.3795	<i>Magelona japonica</i>	0.3313
ハヤテシロガネゴカイ	0.3583	<i>Heteromastus</i> sp.	0.3256
<i>Chone</i> sp.	0.3433	マサゴウロコムシ	0.3222
<i>Glycinde</i> sp.	0.3361		
<i>Ouvnia</i> sp.	0.3282		
<i>Neriniides</i> sp.	0.3113		

SP-C	所 属 率	SP-D	所 属 率
<i>Corophium</i> sp.	0.9478	<i>Aricidea eximia</i>	0.8416
ホソハサミソコエビ	0.9212	サルボウ	0.8122
<i>Gastrosaccus formosensis</i>	0.8553	トゲワレカラ	0.7982
<i>Pontogeneia</i> sp.	0.8533	コオニスビオ	0.7916
カマテワレカラ	0.8533	アシビキツバサゴカイ	0.7907
ヒサシソコエビ	0.8459	アラムシロ	0.7801
ニホンドロソコエビ	0.7979	ヘイケガニ	0.7725
<i>Ceratonereis erythraeensis</i>	0.7912	<i>Anaitides</i> sp.	0.7436
カギメリタヨコエビ	0.7746	ミズヒキゴカイ	0.7417
<i>Eteone</i> sp.	0.7441	<i>Phylo felix asiaticus</i>	0.7354
カギツメピンノ	0.7251	クビナガスガメ	0.7062
ヤマトスビオ	0.6895	ホソヨコエビ	0.6812
マルソコエビ	0.6702	<i>Stenothoe gallensis</i>	0.6695
<i>Dimorphostylis asiatica</i>	0.6635	サザナミガイ	0.6695
ツノメエビ	0.6604	キセワタ	0.6428
シオフキ	0.6557	スゴカイイソメ	0.6333
ウミナナフシ	0.6503	チマキゴカイ	0.5466
マドカスビオ	0.5946	<i>Montacutona</i> sp.	0.5150
ニッポンモバヨコエビ	0.5637	ホトトギスガイ	0.5082
マテガイ	0.5582	ヒメネジガイ	0.5022
<i>Rhynchospio</i> sp.	0.5029	<i>Amaeana trilobata</i>	0.4866
アサリ	0.4688	<i>Nerinides</i> sp.	0.4597
<i>Gammaropsis</i> sp.	0.4152	<i>Capitella capitata</i>	0.4183
<i>Owenia</i> sp.	0.3635	ハヤテシロガネゴカイ	0.4166
ヒメネジガイ	0.3531	アサリ	0.3523
ホトトギスガイ	0.3419	<i>Paraonis gracilis minula</i>	0.3393
		<i>Rhynchospio</i> sp.	0.3337

表 3. 種群 SP-A, B, C, D の重複度行列  
(分割係数0.5068, 分割度0.3423)

	SP-A	SP-B	SP-C	SP-D
SP-A	0.1129	0.0549	0.0381	0.0457
SP-B	0.0549	0.1257	0.0289	0.0303
SP-C	0.0381	0.0289	0.1400	0.0487
SP-D	0.0457	0.0303	0.0487	0.1281

各種群への所属率が0.5以上の種の個体数を各地点別に合計して1種当りに換算した値を種群の個体数とし、一方、地点をA区、C区、B区の順に、かつそれらへの所属率が大きい順に並べて、種群の個体数を棒グラフで表した。ただし、どの区の所属率も0.5未満のあいまいな地点53、13、40は末尾に並べた(図6)。さらにこれらを、各地点における相対的比率を表す特化係数<sup>4)</sup>に換算して、同様の棒グラフで表した(図7)。

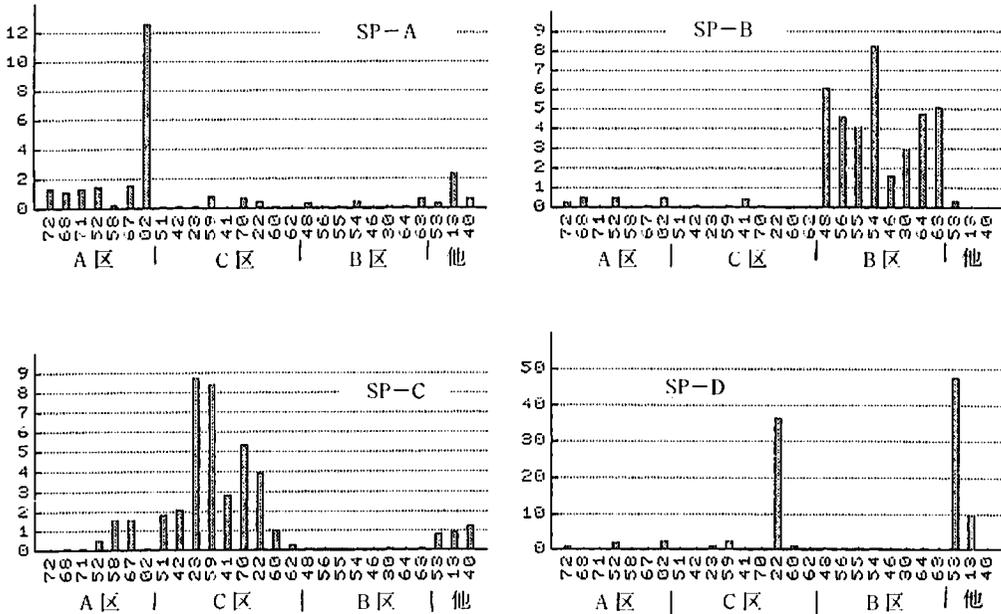


図6. 種群の個体数と地点群の特徴

(種群 SP-A, B, C, D に、所属率が0.5以上の種の個体数を合計し、1種当りに換算した値を示す)

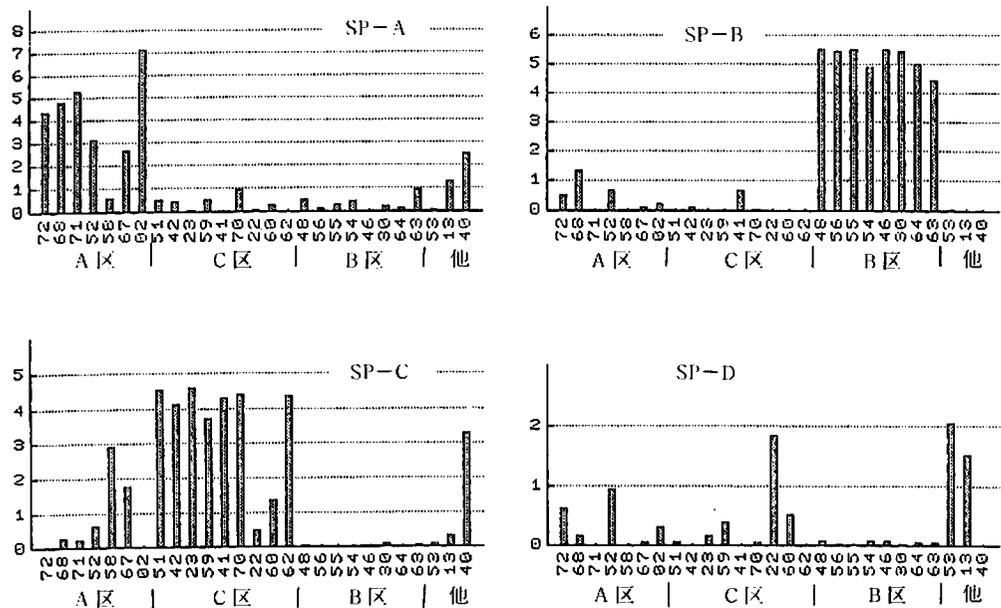


図7. 種群 SP-A, B, C, D の特化係数と地点群の特徴

図から、SP-A、B、CはそれぞれA区、B区、C区で相対的比率が大きいことが分かる。SP-Dは所属のあいまいな地点53、13とA区、C区に散在しているが、その個体数は小さい。

各種群の構成種をみると、SP-BとCが対照的である。すなわち、SP-Bには *Lumbrineris logifolia*、イヨスタレ、シズクガイなど、安定した泥底域に生息する種が多いのに対し、SP-Cには流動的な環境に生息する端脚類などの小型甲殻類が多く、また、ヤマトスピオ、*Rhynchospio* sp. など河口域干潟でよくみられる種が含まれている。SP-Aの種構成には、これらの中間的な様相がみられ、適度に流動する泥底域の群集と考えられる。

これらのことから、B区は安定した泥底域、C区は流動的な干潟域、A区はC区よりは安定し、B区よりは流れの影響を受けた泥底域と推察される。

図8に、動物門別の、情報量による多様度  $D_n$  を上と同様の棒グラフで示す。これから、A区とB区は環形動物が豊富であり、C区は節足動物が豊富な区域であることが分かる。このことも、上に述べたことを裏付けている。

図9~12に、種群を構成する個別の種の分布を示す。これからも分かるように、個別の種の個体数は変動が大きいため、その分布からは上に述べたような推察は困難である。関係類似度行列の固有値分解とファジイクラスタリングの手法を用いて地点を順序づけ、分布の似た種をクラスタリングすることによって、これらの推察が可能となった。

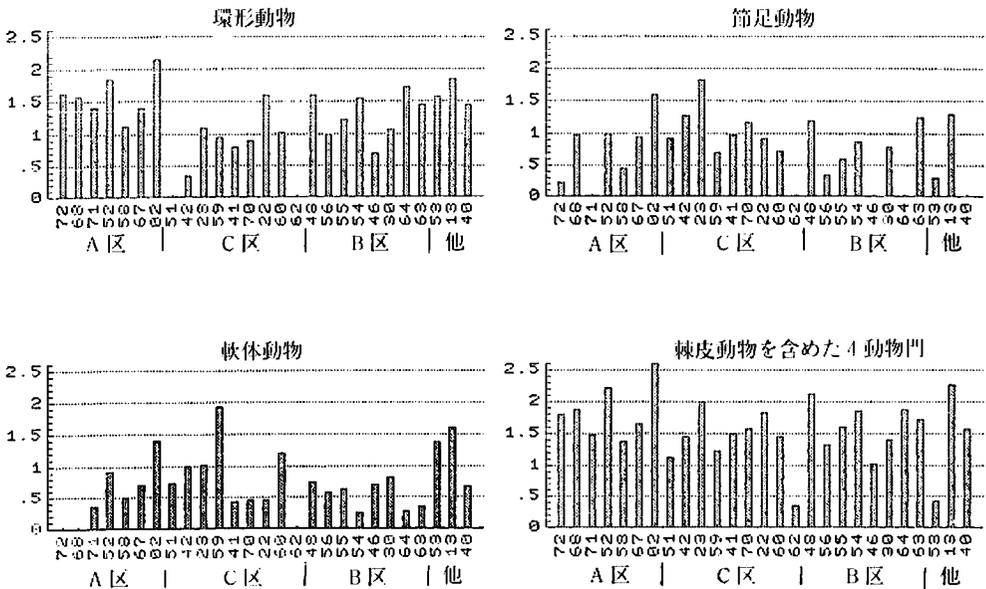


図8. 動物門別多様度指数  $D_n$  と地点群の特徴

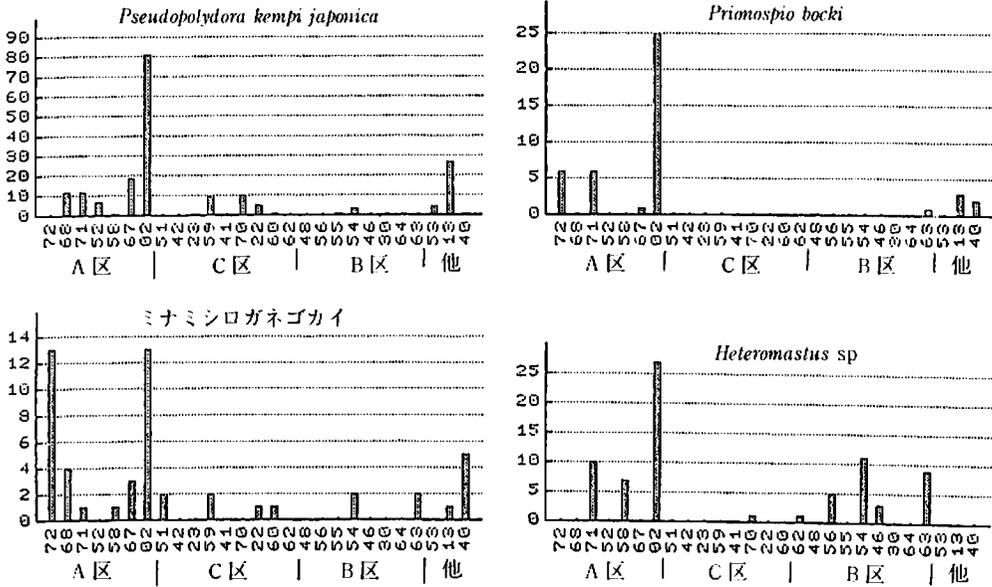


図9. SP-Aの構成種の分布と地点群の特徴

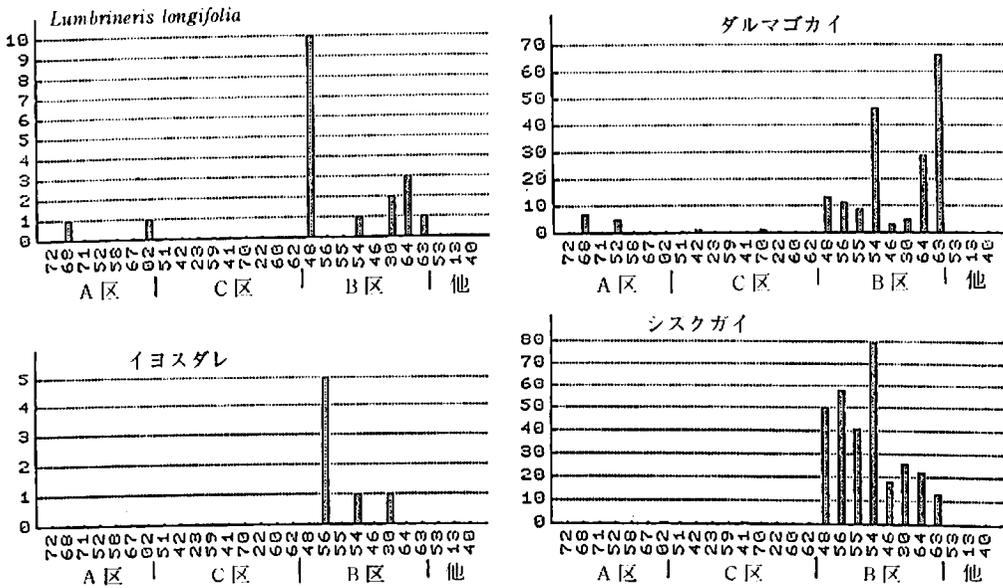


図10. SP-Bの構成種の分布と地点群の特徴

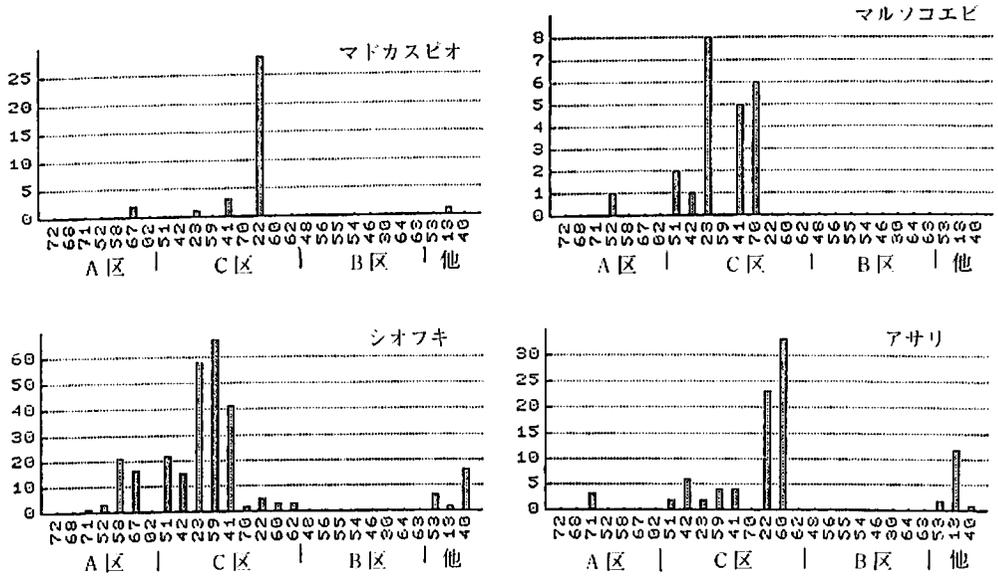


図11. SP-Cの構成種の分布と地点群の特徴

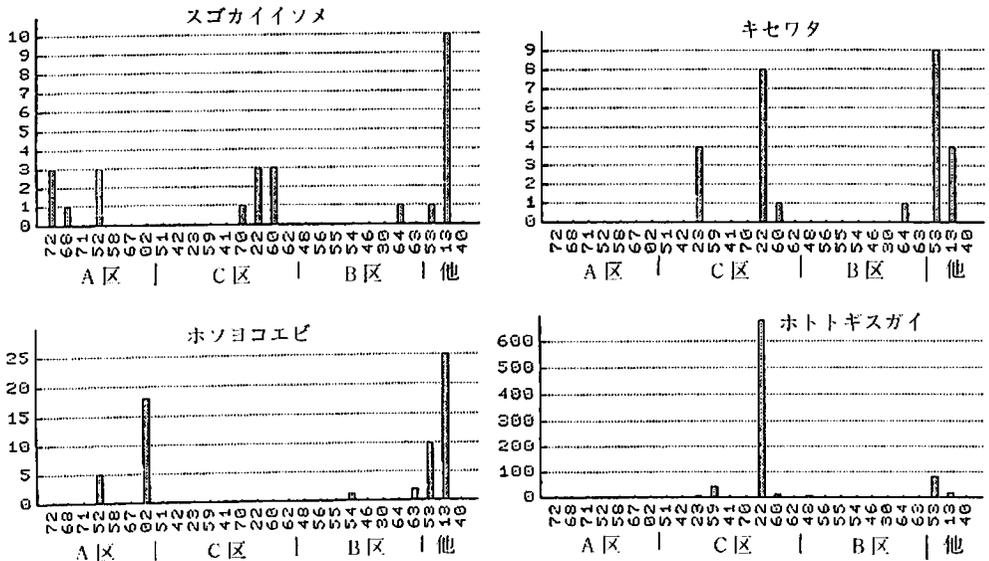


図12. SP-Dの構成種の分布と地点群の特徴

#### 4. 魚類の胃内容物調査結果への適用例

柳井港において昭和56年～57年の約2年間にわたって行われた魚類の胃内容物調査結果<sup>10)</sup>に、上に述べた解析方法を適用した。

##### 4・1. 類別胃内容物間の関係類似度行列の固有値分解

調査期間を7～8月、9～10月、11～1月の3期間に分け、それぞれの期間について、同一魚種の、類別に識別された胃内容物の度数を集計した。同一魚種でも調査期間の異なる場合は異なる魚種として扱った。調査された魚種は7～8月は88種、9～10月は86種、11～1月は53種である。

全期間を通じて度数が10未満の類別項目と消化物、その他、および空胃を除いた17項目について関係類似度行列を計算し、これを固有値分解した。用いた17項目の記号と名称を表4に示す。これらの項目が出現しない魚種は除いたため、計算に用いた魚種は7～8月は76種、9～10月は75種、11～1月は45種の累計196種である。

表4. 類別胃内容物の記号と名称

記号	類別名	記号	類別名
Lf	Larvae of fish	Al	Algae
St	Stomatopoda	Ga	Gastropoda
Ma	Macrura	Bi	Bivalvia
Br	Brachyura	Os	Ostracoda
Op	Ophiuroidea	Co	Copepoda
Po	Polychaeta	Cl	Cladocera
Is	Isopoda	Le	Eggs of fish
Xc	Xenopheusia	Di	Diatom
Am	Amphipoda		

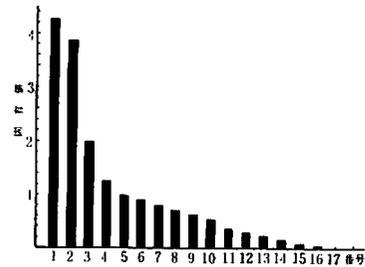


図13. 胃内容物間関係類似度行列の固有値

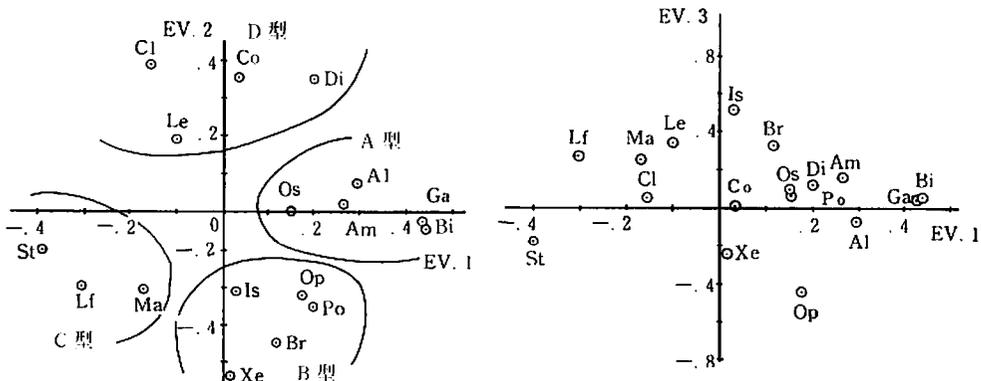


図14. 関係類似度行列の第1～第3固有空間への類別胃内容物の配置

固有値を図13に、第1～第3固有ベクトルによる類別胃内容物の配置を図14に示す。

固有値は第2成分までで情報集約率が48%とやや低いですが、図14から、第3成分では、類別項目の有意と思われる配置がなされていないので、第2成分までを以下の解析に用いた。

#### 4・2. 類別胃内容物のファジクラスタリング

2次元の要因ベクトルを用いて、類別胃内容物をA～Dの4つの型に分類するファジクラスタリングを行った。結果を表5に、また重複度行列を表6に示す。

表5. 類別胃内容物のファジ*k*-means法によるクラスタリング  
(数字は各項目がA,B,C,D型に所属する割合を示す)

	Lf	St	Ma	Br	Op	Po	Is	Xe	Am
A型	・27E-02	0.023	0.0418	0.0247	0.0736	0.0857	0.0694	0.084	0.9291
B型	・67E-02	0.0428	0.145	0.9441	0.882	0.866	0.8064	0.7194	0.0359
C型	0.9871	0.8956	0.7723	0.0219	0.0269	0.0291	0.0907	0.1528	0.0117
D型	・34E-02	0.0385	0.0409	・92E-02	0.0175	0.0192	0.0334	0.0437	0.0233

	Al	Ga	Bi	Os	Co	Cl	Le	Di
A型	0.9252	0.8852	0.8594	0.545	0.018	0.0504	0.0861	0.2256
B型	0.0315	0.063	0.08	0.2355	・9E-02	0.0367	0.0711	0.0743
C型	0.0124	0.0194	0.0234	0.0835	・98E-02	0.0583	0.1173	0.0594
D型	0.0308	0.0324	0.0371	0.136	0.9632	0.8545	0.7255	0.6407

表6. 類別胃内容物A,B,C,D型の重複度行列  
(分割係数0.7282,分割度0.6376)

	A型	B型	C型	D型
A型	0.2134	0.0368	0.0124	0.0273
B型	0.0368	0.2171	0.0264	0.0168
C型	0.0124	0.0264	0.1433	0.0163
D型	0.0273	0.0168	0.0163	0.1544

クラスタリングされたA～D型を構成する項目に共通する特性をみると、A型は静止型あるいは運動力の小さいベントスであり、B型は歩行型あるいはやや運動力のあるベントスであり、C型は遊泳型あるいは運動力の大きい動物で、D型は浮遊型であることが分かる。すなわち、餌生物の運動性が魚類の食性に関与していることが推察される。また、重複度行列から、A型とB型は重複する度合いが大きいことや、C型とD型はクラスターとしてのまとまりが小さいことなどが分かる。

#### 4・3. 魚種のファジクラスタリングと各魚種群の餌生物選択特性

計算に用いた魚種をSP-A、B、C、Dの4群に分類するファジクラスタリングを行い、

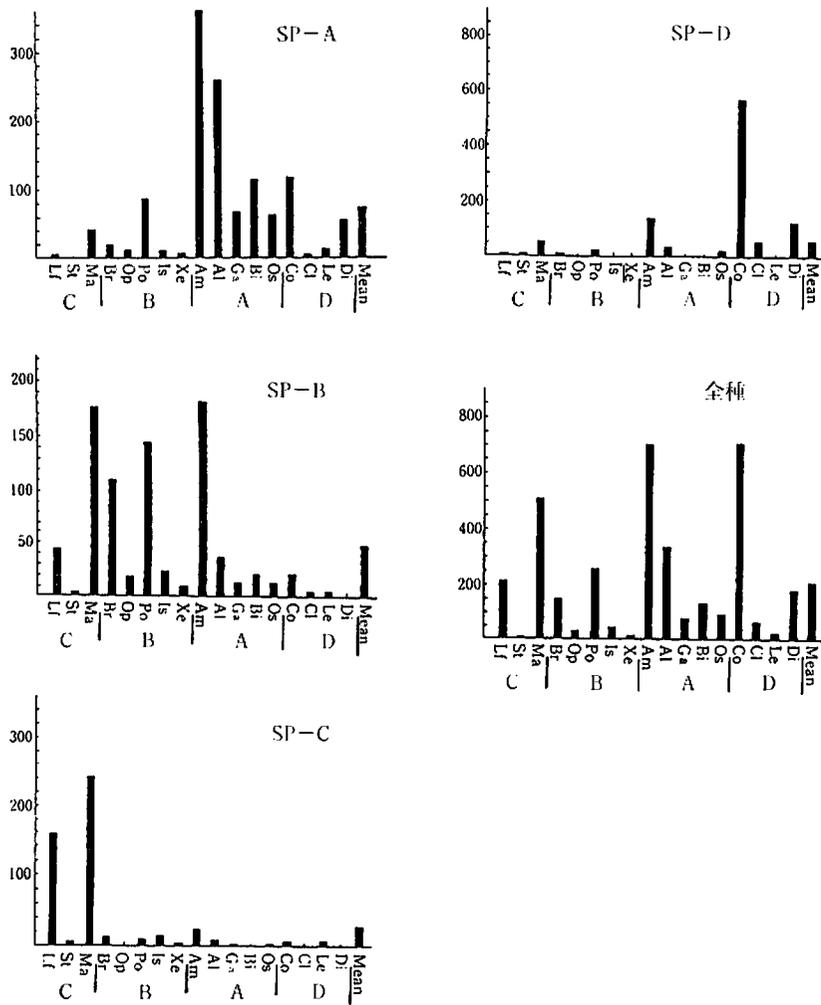


図15. 魚種群 SP-A, B, C, D における類別胃内容物の出現度数  
(所属率0.5以上の魚種について、出現度数を合計した値を示す)

各種群ごとに所属率が0.5以上の魚種の胃内容物の類別出現度数を合計し、一方、類別胃内容物をC型、B型、A型、D型の順に、かつそれらへの所属率が大きい順に並べて棒グラフで表した。結果を図15に示す。

これから、SP-AとBはAm（端脚類）で、SP-BとCはMa（長尾類）で、SP-AとDはCo（撓脚類）でそれぞれ重複が大きいことが分かる。SP-CとDではほとんど重複はない。

ところで、これらの値は環境中の餌生物の量に影響されるから、各魚種群ごとの餌生物に対する選択特性を表すものではない。そこで各魚種群ごとに、全項目の平均度数に対する比を計算し、さらにそれらを、全種群で計算した対平均比で除して、特化係数に換算した。結果を図16に示す。

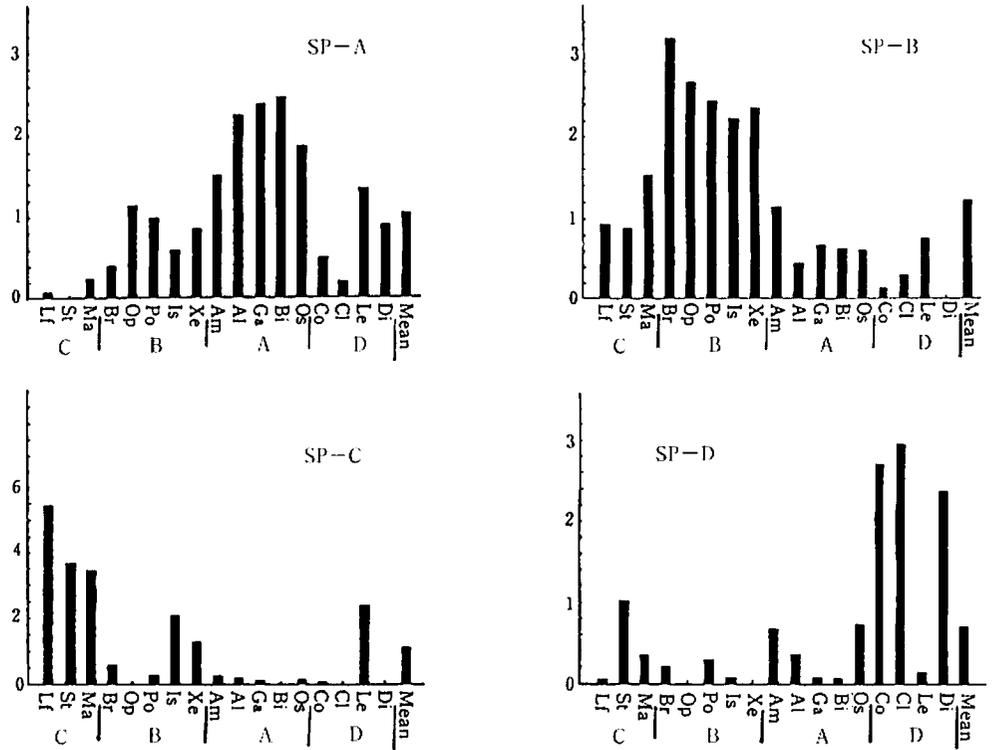


図16. 魚种群 SP-A, B, C, D における類別胃内容物の特化係数

これから、魚种群 SP-A, B, C, D は、それぞれ A 型、B 型、C 型、D 型の餌生物に対する選択度が大きいことが分かる。ただし Le (魚卵) は、SP-A で選択度が大きいことなど、他とは異なっている。これは、関係類似度行列の第 2 成分まででは情報集約率が小さく、誤差成分が大きいため、出現度数が小さい魚卵については、クラスタリングにおいて他との相違が十分反映されなかったためと考えられる。

これらの結果は、昭和 56、57 年度に下関市で行われた同様の調査データを解析した結果<sup>6)</sup>ともよく一致しており、類別段階での魚類の食性の一般的特徴をとらえたものといえるであろう。

## 5. 結言

先に述べたように、動物の種-個体数データには、一般的な多変量解析法は適用し難い。

環境調査を目的として行われる底生動物調査においては、出現する動物の種類数が多いことに加えて、種ごとの個体数の変動が大きいことなどのために、生データから環境についての情報を得ることは困難を伴う。そこで類似した組成の調査地点と類似した分布をする種をクラスタリングすることが必要となる。その場合、個体数が小さい種の分布も環境を考える上では重

要な意味をもつこともあるから、それらが十分反映されるようにしなければならない。

ここで述べた関係類似度行列の固有値分解とファジィ  $k$ -means法によるクラスタリングを用いた解析方法は、そのような必要性を満足するものである。

乱数を用いた数値実験の結果によれば、 $R_s$ の値は正の側に偏る傾向がある。それは、個体数が大きい種ほど出現し易いことが共通の要因として作用するためである。また、全ての種の個体数が出現した調査地点の平均値に等しい場合、 $R_s$ の値は0/0となるために、平均的な地点間の $R_s$ の値は誤差変動が大きくなる。関係類似度 $R_{sm}$ の計算は、 $R_s$ のこのような問題点を解決することにもなっている。

こうして得られた地点のクラスタリングの結果は、解析例に示したように、種群の分布の特化係数の相違をよく反映したものとなる。

特化係数は、情報量理論からも導かれる指数であり、クラスターの有意性の統計的検定に用いることができる。

この解析方法は、データに何ら前提条件を設定していないので、探索的データ解析の第一歩として、広く応用が可能である。

## 文献

- 1) S. Kimoto, *Esakia*, 6, 27-54 (1967)
- 2) 大隅 昇: 数理科学, 17, No. 5, 34-41 (1979)
- 3) 品川 汐夫: 日本ベントス研究会誌, 26, 49-65 (1984)
- 4) 上田 尚一: データ解析の方法, 朝倉書店, 東京, 1982, pp. 5~42
- 5) 柳井湾水産振興協議会: 柳井湾海域の漁業振興に関する調査報告書, 1984, pp. 116-135
- 6) 下関市外海漁業共励会: 下関市南部水域資源診断事業に関する報告書, 1983, pp. 27-29