

東北大学 正員 松本 順一郎
シ 学生員 〇大村 達夫

1) はじめに 近年、河川や湖沼の水質を評価する手法として、主成分分析が用いられている。¹⁾²⁾³⁾ 主成分分析は、種々な水質をそれぞれ一つの変量と考え、変量間の相関や、その変量によって作られる新しい変量として主成分を抽出することにより、水域の汚染度を総合的に判断するのに用いられる。

本報告は、筆者らが昭和49年5月から50年4月までの一年間に実施した宮城県女川港の水質データについて主成分分析を適用した結果の報告である。図-1に女川港の概略と調査地質を示す。調査項目は、水温、透明度、pH、 Cl^- 、BOD、COD、大腸菌群数、腸球菌数、飽和度である。解析には、A~D点のデータを用いた。



図-1 女川港概略
A~F 調査地質

2) 水質の分布様式について 主成分分析を行うためには、各水質ごとにデータを正規化する必要があるため、各水質の分布様式を調べた。その結果を表-1に示す。この結果は、有意水準5%で検定を行ったものであり、pH、 Cl^- 、飽和度は正規分布を示し、透明度、BOD、COD、大腸菌群数、腸球菌数はそれぞれ対数正規分布としてよいことがわかった。しかし、水温だけは有意水準5%で正規分布として受容できなかったが、正規確率紙上にプロットした結果より擬似正規分布と考えた。以上より、主に人為的汚染と考えられる水質項目に関しては、対数正規分布をするものと考えられる。対数正規分布をする水質項目に関しては、主成分分析を行う場合データの対数を用いて分析を行った。

水質	分布様式
水温	対数正規分布
透明度	正規分布
pH	〃
Cl^-	〃
BOD	対数正規分布
COD	〃
大腸菌群数	〃
腸球菌数	〃
飽和度	正規分布

表-1 水質の分布様式

3) 水質の相関について 表-2に水質の相関行列を示す。右上半分は、全部のデータ(77個)を用いて求めたものであり、左下半分は異常値を除いた(57個)のデータから求めたものである。ここで、異常値とは各調査地質ごとに、それぞれの水質が表-1にあげたような分布をするものとして、有意水準5%で受容できないデータの値である。表-2より、すべてのデータより求めた相関係数は、異常値を除いたものより求めた相関係数より全体的に絶対値において小さな相関係数を示した。主成分分析が、相関行列から出発することを考えると、水質の異常値の取捨により、主成分分析結果が相異なることになる。したがって、主成分分析や他の統計処理においては、異常値の取り扱いや、それに付随する調査地質の決定などを、その調査目的に従って事前に十分検討すべきであると考ええる。

	1	2	3	4	5	6	7	8	9
1. 水温		-.341	-.110	-.370	-.260	-.175	-.374	-.115	-.207
2. 透明度	-.423		-.191	-.508	-.423	-.374	-.485	-.278	-.193
3. pH	-.246	-.320		-.487	-.413	-.448	-.538	-.525	-.583
4. Cl^-	-.476	-.536	-.570		-.156	-.227	-.407	-.213	-.195
5. BOD	-.254	-.370	-.431	-.160		-.775	-.540	-.588	-.085
6. COD	-.200	-.372	-.400	-.248	-.789		-.518	-.576	-.140
7. 大腸菌群	-.447	-.529	-.621	-.538	-.466	-.456		-.735	-.286
8. 腸球菌	-.142	-.278	-.553	-.377	-.608	-.582	-.718		-.400
9. 飽和度	-.219	-.012	-.675	-.317	-.177	-.176	-.371	-.435	

表-2 相関行列

水温は、大腸菌群数と腸球菌数とは正の相関をもっており、他の水質とは負の相関がある。透明度は、BOD、COD、大腸菌群数、腸球菌数とは負の相関をもっており、明らかに、透明度が高くなれば、水質がよくなることを示している。pHと Cl^- も透明度とほぼ同様である。BODは、CODと正の相関を示し、すべての水質項目間の相関のうちもっとも高い相関を示した。また、大腸菌群数と腸球菌

水温は、大腸菌群数と腸球菌数とは正の相関をもっており、他の水質とは負の相関がある。透明度は、BOD、COD、大腸菌群数、腸球菌数とは負の相関をもっており、明らかに、透明度が高くなれば、水質がよくなることを示している。pHと Cl^- も透明度とほぼ同様である。BODは、CODと正の相関を示し、すべての水質項目間の相関のうちもっとも高い相関を示した。また、大腸菌群数と腸球菌

主成分	(A)		(B)	
	固有値	寄与率	固有値	寄与率
1	3.913	43	4.241	47
2	1.683	62	1.810	67
3	1.384	77	1.179	80
4	2.727	86	0.527	86
5	0.349	90	0.361	90
6	0.288	93	0.285	93
7	0.270	96	0.268	96
8	0.194	98	0.170	98
9	0.148	100	0.149	100

表-3 固有値と寄与率

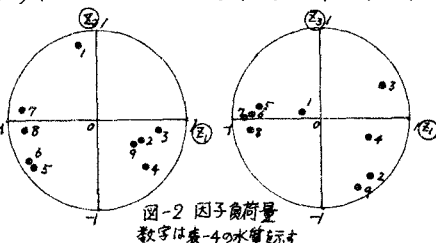
水質	1	2	3	4
水温	-.228	.832	.121	-.346
透明度	-.574	-.202	-.687	-.053
pH	-.742	-.107	-.411	-.344
Cl^-	-.518	-.525	-.203	-.533
BOD	-.732	-.559	-.162	-.020
COD	-.207	-.483	-.101	-.062
大腸菌群	-.835	-.122	-.063	-.319
腸球菌	-.807	-.149	-.181	-.374
飽和度	-.427	-.252	-.782	-.024

表-4 因子負荷量

表の相関は、正の相関を示し、Bの0とCの0との相関のつぎに高かった。

(4) 主成分分析結果および考察 表-3に、相関行列より求めた固有値と変動の累積寄与率を示す。(A)欄は、すべてのデータより求めた値であり、(B)欄は、異常値をのぞいたデータから求めた値である。両方とも、第1主成分まで変動の60%を表わすことができる。しかし、異常値を含めない場合には、第1主成分の寄与率が大きくなり、第1主成分でより大きな変動を表わすことができる。

これより以下の結果と考察は、異常値を含めたすべてのデータから求めたものについて行なう。表-4に、第1主成分までの因子負荷量を示す。因子負荷量は、各水質と各主成分との相関を示すもので、それぞれ主成分の物理的な意味を明確にすることができる。第1主成分は、水温、透明度、Ca、飽和度を除く水質と高い相関を示し、とくに大腸菌群数と腸球菌数との負の相関が顕著である。したがって、第1主成分は人為的汚染の総合汚染指標と考えられる。すなわち、第1主成分の得点が、小さいほど人為的な汚染が進んでいることを示すことになる。



第2主成分は、水温と高い相関を示し、気候的な因子と考えられる。第3主成分は、透明度と飽和度と高い相関を示し、第4主成分はCaと若干の相関を示しているけれども、両者ともあまりとした意味づけはできない。

図-2は、因子負荷量を図示したものであり、上に述べた関係をよく理解することができる。

図-3に、調査地ごとに第1主成分から第3主成分の因子得点を時系列として図示した。これより、第1主成分は、A点、B点で一年間を通じてほとんど負の値を示しており、人為的汚染が相当進んでいるものと考えられる。C点はすべて正の値であり、A点、B点にくらべてはるかに人為的汚染が少ないものと考えられる。B点は、その中間の汚染状況を示している。第2主成分は、水温と高い相関を示しているため、A点、B点のすべてにおいて、よりとした季節変化がみられる。

(5) まとめ (a) 女川湖の水質データに対して、主成分分析を行った所、第1主成分として、人為的な総合汚染因子と、第2主成分として、気候因子が抽出された。しかし、主成分分析をする場合には、水質の異常値の取扱により、分析結果が異なるので、異常値の取扱や調査地点の選定を事前に調査目的と見極めた上で決定する必要がある。この点に関しては、今後さらに検討する必要があると思われる。

参考文献

- 1) 南部祥一他 湖沼水質の評価方法に関する一考察 第2回環境問題シンポジウム講演集 pp61~66, 1974
- 2) 寺島泰 河川水質日変動特性の解析ならびに重回帰分析, Kalman filterによる日変動の予測-淀川下流部水質変動を対象として- 第11回化生工学研究討論会論文集 pp130~135, 1975
- 3) 安田正志他 河川における水質環境の評価に関する研究 第13回化生工学研究討論会論文集 pp13~18, 1977

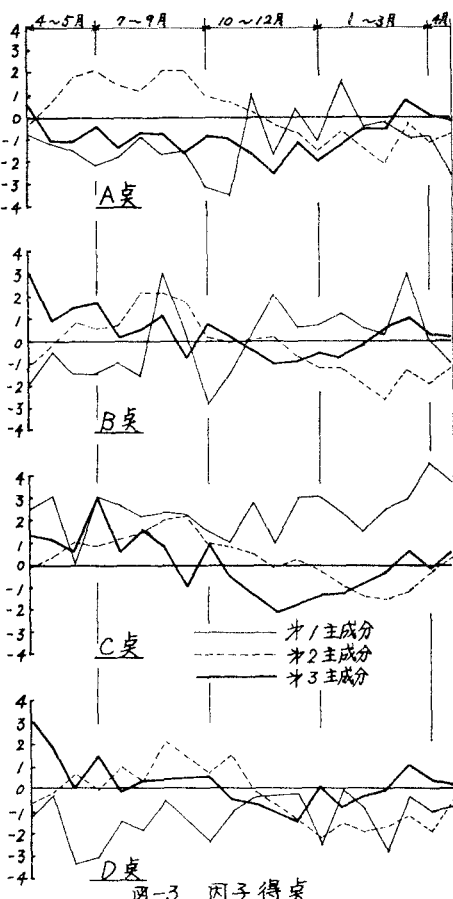


図-3 因子得点