

引用分析の方法と環境土木工学雑誌への応用

Some Methods of Citation Analysis and Their Application to
Journals of Environmental/Civil Engineering

宮 本 定 明

Sadaaki Miyamoto

中 山 和 彦

Kazuhiko Nakayama

Résumé

Three methods of analyzing matrices of mutual citation frequencies among scientific periodicals are proposed in order to identify substructures, to visualize information flows among elements, and to clarify the characteristic of a specific discipline of science. Two of them are agglomerative clustering and multidimensional scaling: they are well-known statistical techniques based on a similarity or dissimilarity measure among elements. Since the measure of similarity assumes symmetric interaction, these analyses need symmetrization of the citation matrix. At the same time, this transformation ignores a distinctive feature, i. e., non-symmetry of the citation matrix. Here the non-symmetry occurs when a journal frequently cites another, while the latter has not many citations of the former. Therefore, another method which is called here "hierarchical graph" is presented so that non-symmetrical relationships are represented on a directed graph whose nodes are hierarchically positioned. The last method determines two kinds of arcs of the graph. Roughly speaking, an arc of one-sided influence from A to B is drawn when journal B frequently cites journal A but the converse is not true; an arc of two-sided influence between A and B is drawn when both journals have many citations of each other. In order to portray the graph, the positions of the nodes are determined so that the arcs have upward directions. An outline of the algorithm for this configuration is given with a simple numerical example.

These methods are applied to the relationship of 43 journals of environmental/civil engineering: the data are extracted from *Science Citation Index* (1977). The citation matrix is derived

宮本定明: 筑波大学学術情報処理センター

Sadaaki Miyamoto, Scientific Information Processing Center, University of Tsukuba, Ibaraki Prefecture.

中山和彦: 筑波大学電子情報工学系教授

Kazuhiko Nakayama, Professor, Institute of Electronics and Information, University of Tsukuba, Ibaraki Prefecture.

and the above methods are applied. The clustering gives nine subgroups which consist of four groups of civil engineering, three of environmental engineering, and two of interdisciplinary studies. The hierarchical graph represents the information flows, from which the studies about water systems seem to be remarkable. Further, the MDS represents the structure of this field by a two-dimensional plot: the first axis shows the relationship of civil and environmental engineering, whereas the second axis suggests nature-human relationship. Moreover, both axes show the importance of the studies about water systems in the discipline of environmental/civil engineering.

- I. はじめに
- II. 分析の方法
 - A. クラスタ分析と多次元尺度法
 - B. 階層化法
- III. 環境土木工学関連雑誌への適用
 - A. 引用行列の作成と分析結果
 - B. クラスターの名称について
- IV. 考 察
- V. おわりに
付 録

I. はじめに

学術雑誌の相互引用関係の研究には、数理的観点からみて、2つの傾向が認められる。1つは、各々の雑誌の引用頻度の順位や、時系列の傾向をみようとするもので、方法論的には、記述統計に関連している。¹⁾

また、他方で、Narin らの研究²⁾にみられるように、学術雑誌相互の引用頻度表を作成し、これに行列演算をほどこすことによって、ある学術雑誌の集合として定義された分野の構造を調べようとするものがある。この場合、1つの要素——学術雑誌——に複数の数値が割り当てられるため、様々な統計的手法が適用できると思われる。

ここでは、引用頻度の行列や、後述の引用率の行列をまとめて、引用行列と呼ぼう。そこで、引用行列というデータに適用できる統計的方法を考えてみると、クラスタ分析や、³⁾ 多次元尺度法^{4), 5)} (Multidimensional Scaling——MDS) があげられる。これらは、標準的な多変量解析の場合と異なり、系の要素間の相互作用の測定だけが記述されていれば、適用することができるから

である。

すなわち、これら2つの方法では要素間の類似性、あるいは非類似性(距離)を表わす測度が与えられていると仮定することが多い。また、測度を S 、任意の2要素を i, j とするとき、対称性

$$S(i, j) = S(j, i)$$

が満たされなければならない。⁶⁾

このことから、引用行列の分析に、これらの方法を用いる際、何らかの対称化が必要となることは明らかである。また、これらの手段では、引用行列のある特徴が把握できないことがわかる。それは、引用行列の非対称性、すなわち、雑誌 i の論文は雑誌 j の論文をよく引用しているが、 j はほとんど i を引用していない、という性質である。このような一方的な引用傾向の重要性はNarin らによって指摘されている。⁷⁾

本稿では、クラスタ分析と多次元尺度法を引用行列に適用するとともに、上に述べたような一方的な引用関係を有向グラフによって表わすことを試みる。グラフを作成する際、後に述べるように、節点⁸⁾を階層的に配置するので、ここでは、この方法を階層化法と呼び、作成

されたグラフを階層図と呼ぶことにする。

これら3つの手法をある分野の学術雑誌に適用するとき、互いに相補的な情報が表現される。すなわち、

クラスター分析——学術雑誌をグループ化することにより、その分野における、様々な研究傾向をまとめて示すことができる。

多次元尺度法——雑誌を平面や三次元の空間に配置することにより、その分野の構造を知ることができる。

階層化法——有向グラフによって、雑誌間の情報の流れを表現する。

ここで扱う具体例は、*Science Citation Index (SCI)* 1977年のファイルからとり出した、環境土木工学関連の52誌である。この分野は、環境工学、土木工学の2分野と、それらの境界領域を含み、データの大きさから、これら3つの手法を用いるのに適している。内容的にも、総合的で学際的な性質をもっているため、様々な研究傾向を含み、その構造は自明でないと考えられる。これら3つの分析方法によって、この分野の全体的視野、2つの分野の接点、個々の雑誌の位置付けと関連をみることができよう。

II. 分析の方法

A. クラスター分析と多次元尺度法

記述の簡単化のため、雑誌を記号 $i=1, 2, \dots, n$ で表わし、引用頻度行列を (C_{ij}) としよう。要素 C_{ij} は雑誌 j が雑誌 i を引用した回数を表わすとする。また、雑誌 i の総引用件数を \bar{C}_i で表わそう。 \bar{C}_j は $C_{1j}, C_{2j}, \dots, C_{nj}$ および記号で表わされていない、その他の雑誌への引用の合計であるので、一般に

$$\bar{C}_j \cong \sum_{i=1}^n C_{ij}$$

が成り立つ。

クラスター分析や多次元尺度法を用いるには、任意の2要素 i, j の親近性、あるいは距離を表わす測度が必要である。この測度を S_{ij} とするとき、対称性 $S_{ij}=S_{ji}$ が成り立たなければならない。引用頻度行列は、この性質をみたさないので、次の式によって、類似性の測度への変換を行う。

$$S_{ij} = \frac{(\text{雑誌 } i \text{ と } j \text{ の相互引用件数})}{(\text{雑誌 } i \text{ と } j \text{ の総引用件数})}$$

$$= \frac{C_{ij} + C_{ji}}{\bar{C}_i + \bar{C}_j}$$

単に、 C_{ij} と C_{ji} の和を考えるのではなく、総引用件数との比をとるのは、引用件数の面での雑誌の規模の影響を考慮するためである。

クラスター分析は、類似度の大きい要素どうしを結合して、全体集合の内部に、複数個のグループを形成する方法である。これには、大別して階層的方法と非階層的方法があるが、ここでは階層的方法を用いる。この方法は、比較的要素の数が少ない場合に適していて、樹形図(dendrogram)によってクラスター形成の過程が詳細に表示されるため、生成されたグループの相互関係が理解しやすいからである。

多次元尺度法は、高根によれば、“対象間の(非)類似性の測度が与えられたとき、対象を多次元空間内の点として表わし、点間の距離が観測された(非)類似性と最も良く一致するように点の布置を定める方法”と要約される。この布置から、系がどのような構造をもっているかをみようとするのが、この方法の目的である。数ある多次元尺度法のアルゴリズムの中から、ここでは、Guttman-Lingoes の MINISSA⁹⁾を選び、データを二次元平面に配置することを考える。

これら2つの手法の詳細については、文献¹⁰⁾に述べられているので、ここでは繰り返さない。

B. 階層化法

先に述べたように、一方的な依存関係を有向グラフに表わすことを考えよう。すべての引用関係を表示したグラフは複雑すぎてかえって判りにくい。そこで、次のような規則によって、グラフを簡単化することを考える。

- (i) 引用度が非常に小さい場合は無視する。
- (ii) 引用が一方的か否かは、要素対間の2つの引用度の比によって判断する。

ここで、引用度というのは、必ずしも引用頻度のこととは限らない。実際、ここでは、A節で述べたように、雑誌の規模の影響を除くため、引用頻度行列を総引用件数によって規格化する。すなわち、引用率行列 (D_{ij}) を次のように定義する。

$$D_{ij} = \frac{C_{ij}}{\bar{C}_j}, \quad 1 \leq i, j \leq n.$$

明らかに、

引用分析の方法と環境土木工学雑誌への応用

第1表 環境土木工学の52誌

環境土木工学関連学術雑誌	省略形(検索文献数)
Agricultural Engineering	AGR ENG (104)
Ashrae Journal-American Society of Heating Refrigerating and Air-conditioning	ASHRAE J. (252)
Brennstoff-Warme-Kraft	BRENN. WARME. (83)
Concrete	CONCRETE (105)
Deep-sea Research	DEEP-SEA RE. (115)
Effluent & Water Treatment Journal	EFF. WAT. TRE. (45)
Engineering Journal of the American Institute of Steel Construction	ENG. J. (24)
Environmental Science & Technology	ENV. SCI. TEC. (246)
Ergonomics	ERGONOMICS (210)
Highway Engineer	HIGHW. ENG. (43)
Human Factors	HUMAN FACT. (61)
International Hydrographic Review	INT. HYD. REV. (24)
International Journal of Applied Radiation and Isotopes	INT. J. A. RAD. (103)
Journal of Agricultural Engineering Research	J. AGR. ENG. R. (47)
Journal of the Air Pollution Control Association	J. AIR POLLU. (181)
Journal American Water Works Association	J. AM. WATER (193)
Journal of the Construction Division-ASCE	J. CONS. ASCE (77)
Journal of the Environmental Engineering Division-ASCE	J. ENVIR. ENG. (146)
Journal of the Hydraulics Division-ASCE	J. HYDR-ASCE. (253)
Journal of the Institute of Fuel	J. I. FUEL (33)
Journal of the Power Division-ASCE	J. POWER-ASC. (24)
Journal Prestressed Concrete Institute	J. PRE. CONCR. (44)
Journal of Soil and Water Conservation	J. SOIL WAT. (67)
Journal of the Structural Division-ASCE	J. STRUCT. DI. (270)
Journal of the Surveying and Mapping Division-ASCE	J. SURV. MAPP. (13)
Journal Water Pollution Control Federation	J. WATER P. C. (252)
Journal of the Waterways Harbors and Coastal Engineering Division-ASCE	J. WATERWAY (59)
Magazine of Concrete Research	MAG. CONCR. R. (26)
Marine Technology Society Journal	MAR. TECH. SJ. (43)
Ocean Engineering	OCEAN ENG. (22)
Proceedings of the Institution of Civil Engineers	P. I. CIV. E. (150)
Photogrammetric Engineering and Remote Sensing	PHOTOGR. ENG. (144)
Public Roads	PUBL. ROADS (28)
Transactions of the ASAE	T. ASAE (254)
Traffic Quarterly	TRAFFIC Q. (32)
Transportation Engineering Journal of ASCE	TRANSP. EN. J. (71)
Transportation Research	TRANSP. RES. (39)
Travail Humain	TRAV. HUMAIN (90)
Water Research	WATER RES. (157)
Water Resources Research	WATER RES. R. (247)
Water Services	WATER SERV. (66)

環境土木工学関連学術雑誌	省略形(検索文献数)
International Biodeterioration Bulletin Water and Wastes Engineering	INT. BIOD. B. (28) WATER WASTE. (131)
*American Gas Association Monthly *Automotive Engineering *Civil Engineering *IEEE Transactions on Electromagnetic Compatibility *Concrete Quarterly *Journal of Environmental Sciences *Mechanik Miesiecznik Naukowo-techniczny *Power Engineering *Reclamation Era	AM. GAS AS. M. (81) AUT. ENG. (93) CIVIL ENG. (120) IEEE ELMAGN. (39) CONCRETE Q. (36) J. ENVIR. SCI. (24) MECHANIK (207) POWER. ENG. (113) RECLAM. ERA (20)

(* 印の9誌は引用頻度行列の要素が自己引用以外すべて0となるので以後、分析対象としない)

$$0 \leq \sum_{i=1}^n p_{ij} \leq 1, \quad 1 \leq i \leq n$$

が成り立つ。

上に述べた規則に従ってグラフの弧を決定するため、2つの閾値パラメータ $\alpha, \beta (0 < \alpha < 1, \beta > 1)$ を導入する。これによって、雑誌 i, j の関連を次の2通りに要約する。

- (1) i が j に一方的に影響する ($i \rightarrow j$) というのは、引用率 p_{ij} が α 以上で、しかも、比 p_{ij}/p_{ji} が β 以上となる場合である。すなわち、 $i \rightarrow j \Leftrightarrow p_{ij} \geq \alpha, p_{ij} \geq \beta p_{ji}$ 。
- (2) i と j の間に両方向の影響がある ($i \leftrightarrow j$) というのは、引用率 p_{ij}, p_{ji} がともに α 以上で、かつ、 p_{ij} と p_{ji} の比が β と $(1/\beta)$ の間にはいる場合である。すなわち、

$$i \leftrightarrow j \Leftrightarrow p_{ij} \geq \alpha, p_{ji} \geq \alpha, \frac{1}{\beta} < \frac{p_{ij}}{p_{ji}} < \beta.$$

わかりやすくいえば、 $i \rightarrow j$ は、 j が i をよく引用し、 i は j をあまり引用していないことを表わし、 $i \leftrightarrow j$ は、 i と j は互によく引用しあっていることを表わす。

隣接関係を定めた後、このグラフの節点をどのように配置すれば、理解しやすい表示が得られるかを考える。ここでは、グラフの弧ができるだけ上方向に向かうような、階層的表現をとる。このようにすれば、他によく引用されている雑誌は、図の下部に位置し、他を引用する傾向の強いものは、上部におかれることになる。

一般に、このような階層的表現ができるためには、グ

ラフに半順序¹¹⁾が導入されなければならない。もし、グラフの弧が、推移律

$$i \rightarrow j, j \rightarrow k \quad \text{ならば} \quad i \rightarrow k$$

を満たすならば、階層的表現ができるが、実際には、これが成り立つとは限らない。そこで、次の手順により、部分的な階層的配置を行うことを考える。

- (1) 弧を通して互いに到達可能な要素をまとめ、クラス G_1, G_2, \dots, G_m を形成する。
- (2) クラス $\{G_1, G_2, \dots, G_m\}$ に半順序 \leq を導入する、すなわち、 G_i から弧を通して G_j に到達可能 なとき $G_i \leq G_j$ とする。

この関係が、半順序の定義をみたすことは容易に確認される。この方法によって節点の配置を行うアルゴリズムの詳細は省略し¹²⁾ 付録に簡単な例を示すとどめる。

III. 環境土木工学関連雑誌への適用

A. 引用行列の作成と分析結果

筑波大学学術情報処理センターで保有している *Science Citation Index* 1977年のファイルから、環境土木工学関連52誌(第1表参照)に属する文献をすべて検索した。なお、雑誌の分類は、SCIの発行元である Institute for Scientific Information (ISI) の作成した *Journal File* (1978年)によった。

検索された文献数は4892件である。これらの引用文献のうち、学術雑誌と、それ以外の単行本、学位論文等を

J I F U E L	J P O W E R A S C	J P R E C O N C R	J S O I L W A T	J S T R U C T I D I	J S U R V M A P P	J W A T E R P C	J W A T E R W A Y	M A G C O N C R	M A R T E C H S J	O C E A N E N G	P I C I V E	P H O T O G R E N G	P U B L R O A D S	T A S A E	T R A F F I C Q	T R A N S P E N J	T R A N S P R E S	T R A V H U M A I N	W A T E R R E S	W A T E R R E S R	W A T E R S E R V	I N T B I O D	W A T E R W A S T E
0	0	0	1	0	0	0	0	0	0	0	0	0	0	74	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	2	0	0	0	20	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	6	2	0	5	2	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	61	0	0	4	0	0	0	0	5	0	0	0	0	22	6	0	2	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	23	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	0	0	5	0	0	0	0	0	0	0	25	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	6	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	75	0	0	0	0	0	0	0	1	0	0	0	0	39	2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	9	5	0	0	0
0	0	0	7	0	0	35	22	0	2	0	24	1	0	8	0	0	0	0	0	90	0	0	0
37	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	22	0	0	4	0	2	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
0	0	6	0	2	0	0	0	3	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	22	0	0	1	0	0	0	0	0	2	0	10	0	0	0	0	1	2	0	0	0
0	0	1	0	293	0	0	0	1	0	0	50	0	0	0	0	3	0	0	0	0	0	0	0
0	0	0	0	0	9	0	0	0	0	3	0	3	0	0	0	0	0	0	0	0	0	0	0
0	0	0	1	0	0	496	0	0	0	0	0	0	0	10	0	0	0	0	121	13	0	0	1
0	0	0	0	4	0	0	35	0	1	0	3	0	0	0	0	0	0	0	0	2	0	0	0
0	0	0	0	11	0	0	0	25	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	3	0	0	18	2	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0
0	0	0	0	12	0	3	0	9	0	0	82	0	1	0	0	1	6	0	0	12	0	0	0
0	0	0	0	0	3	1	0	0	0	0	0	65	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	0	0	0	0	20	0	0	2	0	0	0	0	0	0	0
0	0	0	14	0	0	16	0	0	0	0	0	0	0	379	0	0	0	0	2	15	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	10	0	1	1	0	0	0	0	0
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	6	53	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	25	0	0	0	0	0
0	0	0	0	0	0	232	0	0	0	0	0	0	0	0	0	0	0	0	138	1	2	0	0
0	0	0	5	0	0	46	0	0	1	0	1	1	0	35	0	0	0	0	8	588	2	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	13	0
0	0	0	0	0	0	59	0	0	0	0	0	0	0	0	0	0	0	4	0	0	0	0	0
174	51	70	251	852	20	4332	198	156	207	95	533	300	43	1457	67	82	195	295	1793	2019	23	167	3

引用分析の方法と環境土木工学雑誌への応用

区別し,¹³⁾ 52誌の各々について、学術雑誌の総引用件数を求めて \bar{C}_i とした。学術雑誌以外のものを総引用件数に含ませなかったのは、学術雑誌と、その他のものでは引用の傾向、内容が異なる可能性があると考えたからである。

さらに、引用のリストを調べて、引用頻度表(第2表参照)を作成した。¹⁴⁾ 52誌のうち、引用のないものや自己引用以外に引用を52誌のうちにもたないものが9誌あった。これらは、第1表において星印をつけて示してある。従って以後の分析は残りの43誌を対象に行われる。

そこで、この行列に、第II章の手法を適用した結果、第1図、第2図、第3図を得た。

第1図はクラスター分析による樹形図である。クラスターの結合は群平均法によって行った。¹⁵⁾ 図において、記号が付けられた、類似度0.009以上のグループについては、階層図、多次元尺度法の結果とあわせて表示する。¹⁶⁾

第2図は階層化法により得られた有向グラフである。

閾値パラメータは $(\alpha, \beta)=(0.01, 2.0)$ を用い、特に $\max(p_{ij}, p_{ji}) \geq 0.05$ のところを太線で示してある。図の左上の3誌は弧を持たない孤立誌である。また、破線で囲まれた部分は、上に述べた9つのグループである。

第3図は、MINISSAによる2次元プロットである。図に示された9つの領域は、先に述べたグループである。

B. クラスターの名称について

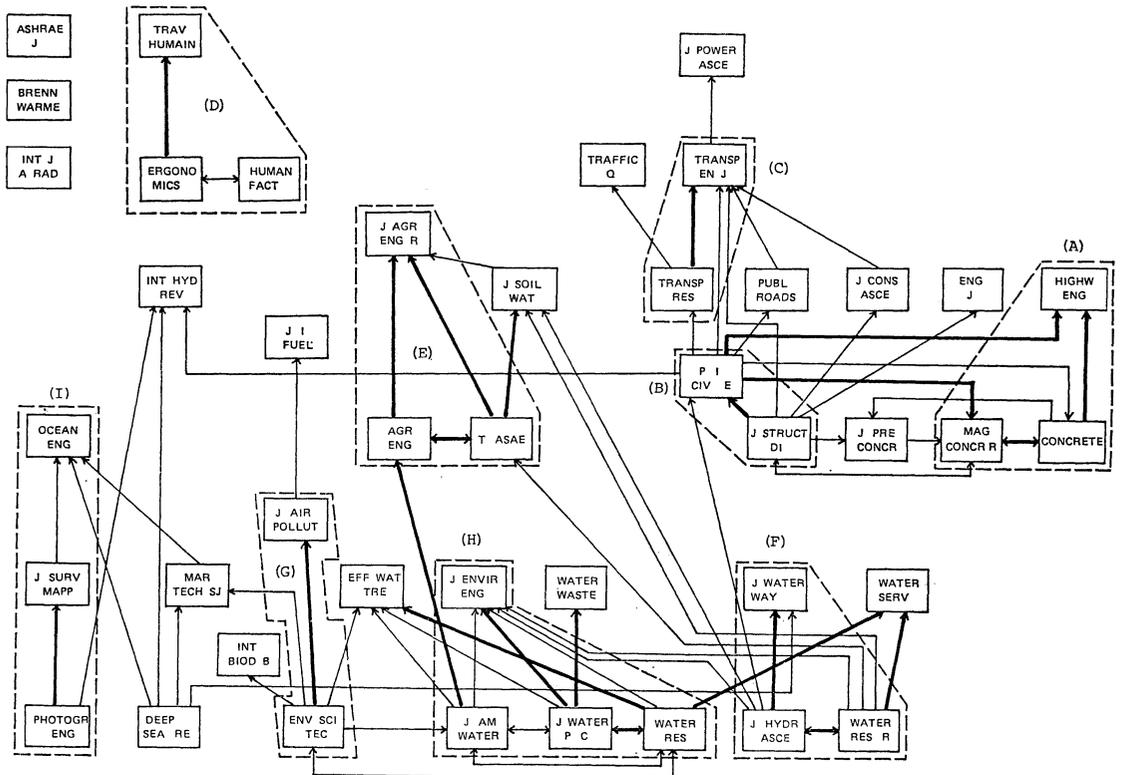
先にとり出した9つのクラスターは、いずれも明確な概念を表わしている。これらは次のようにまとめることができる。

- (A) コンクリート (concrete)
- (B) 構造工学 (structural engineering)
- (C) 輸送 (transportation)
- (D) 人間工学 (human engineering)
- (E) 農業 (agriculture)
- (F) 水系の物理 (physics in water systems)
- (G) 大気汚染 (air pollution)

similarity = 0.009

ITEM NAME	ID NO	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
CONCRETE	4	---	I																							
MAG CONCR R	28	---	I																						I	(A)
HIGHW ENG	10	---	I																						I	
J STRUCT DI	24	---	I																						I	
P I CIV E	31	---	I																						I	
J PRE CONCR	22	---	I																						I	
TRANSP EN J	36	---	I																						I	
TRANSP RES	37	---	I																						I	
PUEL ROADS	33	---	I																						I	
J CONS ASCE	17	---	I																						I	
J POWER ASCE	21	---	I																						I	
ENG J	7	---	I																						I	
ERGONOMICS	9	---	I																						I	
HUMAN FACT	11	---	I																						I	
TRAV HUNAIN	38	---	I																						I	
TRAFFIC Q	35	---	I																						I	
AGR ENG	1	---	I																						I	
J AGR ENG R	14	---	I																						I	
T ASAE	34	---	I																						I	
J SOIL WAT	23	---	I																						I	
J HYDR-ASCE	19	---	I																						I	
WATER RES R	40	---	I																						I	
J WATERWAY	27	---	I																						I	
ENV SCI TEC	8	---	I																						I	
J AIR POLLU	15	---	I																						I	
J WATER P C	26	---	I																						I	
WATER RES	39	---	I																						I	
J AM WATER	16	---	I																						I	
J ENVIR ENG	18	---	I																						I	
WATER WASTE	43	---	I																						I	
EFF WAT TRE	6	---	I																						I	
DEEP SEA RE	5	---	I																						I	
HAR TECH S J	29	---	I																						I	
J SURV MAPP	25	---	I																						I	
OCEAN ENG	30	---	I																						I	
PHOTOGR ENG	32	---	I																						I	
INT HYD REV	12	---	I																						I	
ASHRAE J	2	---	I																						I	
WATER SERV	41	---	I																						I	
BURN WASTE	3	---	I																						I	
J I PUEL	20	---	I																						I	
INT J A RAD	13	---	I																						I	
INT BIOD B	42	---	I																						I	

第1図 群平均法による樹形図



第2図 階層化法による有向グラフ

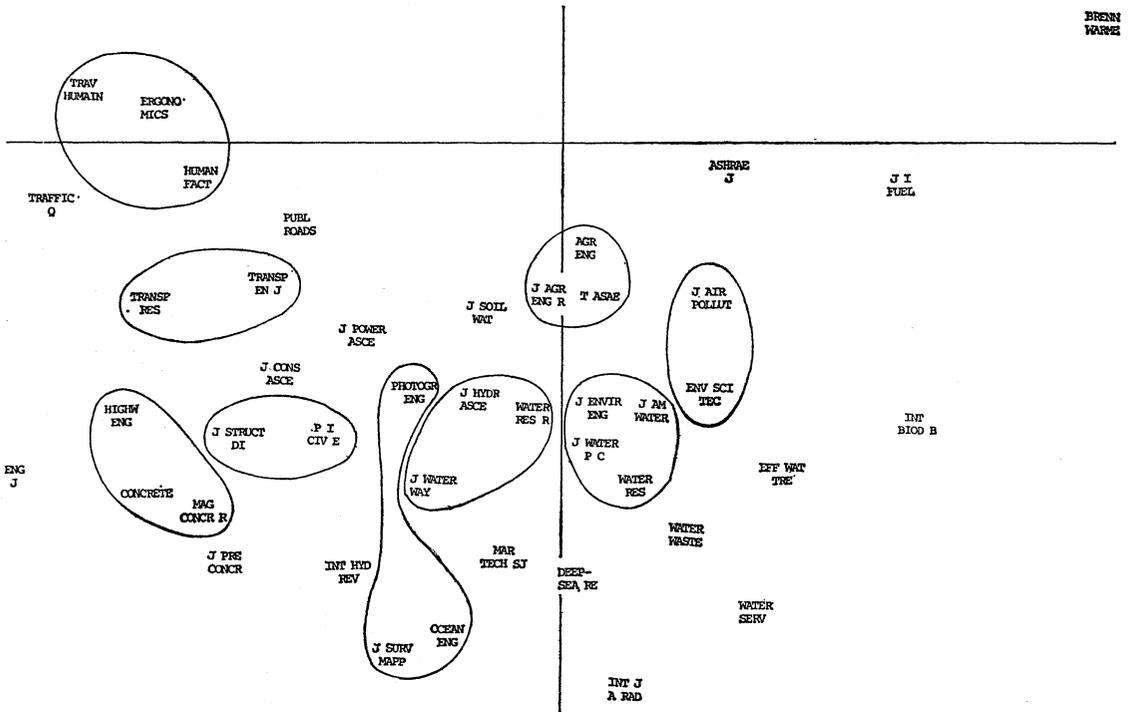
- (H) 水系の生物・化学
(biology and chemistry in water systems)
- (I) 写真測量学 (photogrammetry)

上記の名称は、雑誌のタイトルと、その雑誌に含まれている文献の表題により決定した。そのほとんどは、雑誌のタイトルから容易に理解できると思われる。ただ、グ

ループFとHの相異については自明でない。これは、2つのグループの引用傾向をみることで明らかとなる。第3表はF、Hに属する雑誌の引用している分野をISIの分類に基づいて調べた結果で、2つのグループの引用傾向に、はっきりと差が認められる。

第3表 水系に関する2つのグループの引用傾向

JOURNAL CITATION	J AM WATER				J ENVIR ENG			J WATER P C		WATER RES	J HYDR- ASCE		WATER RES R	J WATER WAY
	ENVIRONMENTAL & CIVIL	146	129	1027	210	133	728	30						
EARTH SCIENCE	12	19	172	81	34	252	11							
CHEMISTRY	24	11	158	142	2	26	1							
PHYSICS	0	3	13	12	30	34	6							
MICROBIOLOGY	6	6	109	72	0	0	0							



第3図 MINISSA による2次元プロット

IV. 考 察

上に述べた9つのグループを土木工学と環境工学に分けてみると、次のようになるであろう。

- (1) 土木工学
A, B, C, I
- (2) 環境工学
E, G, H
- (3) どちらともいえないもの
D, F

ここで、グループFに着目してみよう。階層図をみると、土木工学のグループは右側に多くみられ、グループBはグループFを介して、水系の研究とつながっている。また、グラフの弧は水系の研究F, Hの付近に最も多くみられる。これから、次のことがいえよう。すなわち、この分野では、水系に関する研究が盛んで、その中に、生物・化学的傾向と物理学的傾向がみられる。さら

に、後者は構造工学と関連している。

グループFは流体力学を中心としていると考えてさしつかえないので、構造工学と流体力学の共通点を考えてみると、方法論としての偏微分方程式論とその数値解法があげられる。実際、構造工学での基本的な解法である有限要素法¹⁷⁾は、最近、流体の解析に盛んに用いられている。¹⁸⁾

さらに第3図をみよう。図の左側には土木工学に関するものが位置し、右側には、環境工学に関するものが多い。よって、水平軸は土木工学—環境工学の変化を表わすものといえる。また、これら2つの分野の境界領域に、流体解析のグループが位置している。

一方、垂直方向については、上方に、人間工学、農業、燃料に関するものがみられ、下方には、写真測量、海洋等に関するものが位置している。これから、この方向は人間活動—自然の構造を表わしていると考えられることができる。また、水平、垂直両方の軸に関し、水に関連するものが、中央に位置しているのが顕著である。このよ

うに、環境土木工学においては、水系が中心的主題となっていると考えられる。

V. おわりに

ここにとりあげた3つの手法は、互いに相補的な情報を表現し、ある学術分野の大域的な構造をみるのに適している。これらの方法のうち、クラスター分析、多次元尺度法については、多くの解法があるので、それらを比較検討する必要がある。また、階層化法は、非対称な関係を表現するのに適していて、学術雑誌以外にも様々な応用の可能性があると思われるので、さらに研究を進めていきたい。

ここで行った計算は、すべて筑波大学学術情報処理センターの ACOS 800/II による。また、ここで用いた階層化のプログラムは、要求により、公表する。

最後に、本稿作成にあたって、数々の有益な御助言を頂いた筑波大学電子・情報工学系上田修一講師、また、とりまとめに際し、御協力頂いた慶応義塾大学文学研究科図書館・情報学科博士課程三輪真木子氏、および大分工業高等専門学校青木照子助手に厚く感謝の意を表する。

付 録

簡単な例にもとづいて階層化の手順を示そう。まず、第4表の引用頻度表を考える。相互引用頻度だけを観察すると、A、Bの間の引用が多く、C、Dは孤立しやすいようにみえる。しかしながら、先に述べたような規格化を行った第5表をみると、引用率からは、DはCに依存し、CはAに大きく依存している。

そこで、この引用率行列に対して、閾値 $(\alpha, \beta) = (0.05, 2.0)$ で隣接行列を作成すると第6表のようにな

第4表 引用頻度行列の一例

CITATION \ SOURCE	SOURCE			
	A	B	C	D
A	200	70	20	0
B	100	300	3	0
C	5	0	30	10
D	2	0	0	60
TOTAL	1000	500	100	100

第5表 表4より得られた引用率行列

CITATION \ SOURCE	SOURCE			
	A	B	C	D
A	0.20	0.14	0.20	0.0
B	0.10	0.60	0.03	0.0
C	.005	0.0	0.30	0.10
D	.002	0.0	0.0	0.60

第6表 表5より得られた隣接行列

	A	B	C	D
A	0	1	1	0
B	1	0	0	0
C	0	0	0	1
D	0	0	0	0

る。ここで隣接行列とは、グラフにおいて i から j への弧が存在するとき、またそのときに限り $a_{ij}=1$ となるような 0-1 行列 $A=(a_{ij})$ のことである。なお、無条件に $a_{ii}=0$ としておく。

第2章Bで述べた、互いに可到達な要素で構成されるクラスは、明らかに

$$G_1 = \{A, B\}, G_2 = \{C\}, G_3 = \{D\}$$

である。また、半順序関係は次のようになる。

$$G_1 \leq G_2, G_2 \leq G_3, G_1 \leq G_3$$

これを可到達行列として表わしたのが第7表である。

第7表 可到達行列

	G_1	G_2	G_3
G_1	0	1	1
G_2	0	0	1
G_3	0	0	0

この行列から、まず孤立クラスを探す。このクラスに対応する行、列の要素はすべて0である。第7表では孤立クラスは存在しない。孤立クラスがある場合は、それらを可到達行列から削除する。

つぎに、階層の第1レベルのクラスを探す。このクラ

スは、可到達行列において、どこからも到達できない。従って、対応する列の要素がすべて0である。第7表では G_1 がこれにあたる。このクラスを第1レベルとし、可到達行列から削除する。(第8表参照)

第8表 レベル1を除いた可到達行列

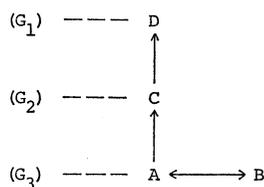
	G_2	G_3
G_2	0	1
G_3	0	0

つぎに、第2レベルのクラスを探す。このクラスは、第1レベルを除いた後では、どこからも到達できないので、この段階の可到達行列において、列の要素がすべて0となっている。第8表では G_2 がこれにあたる。このクラスを第2レベルとし、行列から削除する。(第9表参照)

第9表 レベル1, 2を除いた可到達行列

	G_3
G_3	0

以下同様にして、行列の要素がなくなるまで続ける。従って、この例では、 G_1, G_2, G_3 が各レベルに属し、第4図のグラフを得ることができる。



第4図 階層図の一例

- 1) Garfield, E. "Citation analysis as a tool in journal evaluation," *Science*, vol. 178, 1972, p. 472-9.
- 2) Narin, F., Pinski, G. and Gee, H. H. "Structure of the biomedical literature," *Journal of the American Society for Information Science*, vol. 27, Jan.-Feb. 1976, p. 25-45.
- 3) Anderberg, M. R. *Cluster analysis for applications*. New York, Academic Press, 1973. 359 p.

- 4) 高根芳雄. 多次元尺度法. 東京大学出版会, 1980. 332 p.
- 5) Shepard et al. *Multidimensional scaling, theory and applications in the behavioral sciences*, vol. 1, theory, New York, Seminar Press, 1972. 261 p. 邦訳: 多次元尺度構成法 I—理論編. 岡太彬訓, 渡辺恵子共訳. 東京, 共立出版, 1976. 278 p.
- 6) 非対称な測度に対する多次元尺度法も考案されている. 次の文献を参照されたい. Coombs, C. H., Dawes, R. H. and Tversky, A. *Mathematical psychology—an elementary introduction*. Englewood Cliffs, New Jersey, Prentice-Hall, 1970. 419 p. 邦訳: 数理心理学序説. 小野茂監訳. 東京, 新曜社, 1974. 452 p.
- 7) Narin, F. et al. *op. cit.*, p. 29.
- 8) グラフとは、節点とそれをつなぐ弧(矢印で表わされる)の集合で、ここでは、節点が雑誌を、弧が引用関係を表わす。
- 9) Lingoes, J. C. *The Guttman-Lingoes nonmetric program series*. Ann Arbor, Michigan, Mathesis Press, 1973. 354 p. また、高根, *op. cit.*, p. 107-16, Shepard et al. *op. cit.*, p. 51-75. にも説明がある。
- 10) クラスター分析については, Anderberg, *op. cit.*, 多次元尺度法については, 高根, *op. cit.*, を参照されたい。
- 11) 関係 R が半順序であるというのは、任意の要素 i, j, k に対して次の3つが成り立つ場合をいう。
 - (i) iRi ,
 - (ii) iRj, jRi ならば $i=j$,
 - (iii) iRj, jRk ならば iRk .
 ただし、本稿の場合、(i)は無条件に成り立つものとする。
- 12) 実際には、作図の完全自動化はむずかしい。現状では、このアルゴリズムを用いていったん節点の位置を決めた後、人手によって配置を修正し、より見やすいグラフを作成している。
- 13) *SCI*では、学術雑誌が引用されている場合、巻、ページが明記されているので、これを利用して区別することができる。ただ、少数ではあるが、個人全集を引用している場合に、やはり、巻、ページが記されていて、誤分類されたものがあつた。
- 14) この表の値は、*Journal Citation Report, 1977*, Philadelphia, Institute for Scientific Information. によるものとは若干異なっている。この理由は、*JCR*作成の後、更新がなされたためであろうと考えている。また、総引用件数は大きく異なるが、これは、*JCR*では、上述のような区別をしていないためであろう。
- 15) Anderberg, *op. cit.*, p. 140.
- 16) 他に、最短距離法, 最長距離法を用いた場合も、ここにあげたグループとはほぼ同様のものが形成された。

- 17) Connor, J. J. and Brebbia, C. A. *Finite element techniques for fluid flow*. London, Butterworth & Co. Ltd., 1976. 310 p.
邦訳：流体解析への有限要素法の応用．奥村敏恵監
- 18) *Ibid*, 邦訳．まえがき，および訳者あとがき等を参照されたい．
訳．東京，サイエンス社，1978．232 p.