

『脱ダム』宣言」に対する学生の意識構造特性

信州大学工学部
信州大学大学院工学系研究科
信州大学工学部
信州大学工学部

○小林 雅彦
清水谷勇介
正会員 富所 五郎
正会員 豊田 政史

1. はじめに

田中康夫長野県知事から『脱ダム』宣言」が発せられ、全国的な議論が巻き起こっている。ダムにとどまらず公共事業の進め方や環境への配慮に対する国民の関心は高まりを見せている。この宣言に対する賛否両論は財政政策の問題をはじめ、政治的な経緯や技術面、安全面の観点、環境への配慮など多岐にわたる問題が挙げられ、まさに論点が絞り切れない様相を呈している。本研究では学生がこのような情勢の中で記述した自由記述式のレポートに対し、クラスター分析を用いて科学的に解析する。『脱ダム』宣言に対する意識の全体構造を客観的なデータ解析を通して理解するために、膨大な自由記述文のデータをエクセル・マクロの文字列操作関数などを用いて処理する。これにより、自由記述文のデータベース化の省力化と客観的評価が可能となる。

2. 回答者の構成とデータの構造

平成 13 年 4 月から 11 月にかけて、信州大学の学生に対し『脱ダム』宣言に対する意見を 1000～1200 字にまとめよ。」とレポートを課した。なお、『脱ダム』宣言が発表されたのは平成 13 年 2 月である。回答者の構成は表 1 の通りである。回答者の所属によってなんらかの傾向が現われると考え、A(土木について専門的な知識を持つ)、B(土木の知識を身につける過程にある)、C(土木の知識を持たない)の 3 つのグループに分けて分析を行うことにした。それぞれのグループに属する回答者が記述した文章をまとめて一つの長文と見なし、クラスター分析は回答者が書いた文章に表れる単語について行う。助詞などを除いた意味を持つ単語に対して頻度分析を行い、ある程度の頻度以上記述された単語をクラスター分析の対象とする。

3. 単語の使用頻度分析

表 1 の 163 名の記述に対して単語の使用頻度を調べた。表 2 に頻度の上位 20 語を示す。A グループでは「治水」が 3 位に現れ治水の重要性を示し、他と比べて「自然」に対する意識は低い。B グループでは「県」「知事」「脱ダム宣言」が上位に位置し、田中知事に対する意識の高さを伺い知ることができる。C グループでは「自然」が目立って上位に位置し「ダム」「建設」よりも自然を大切にしたい意向がわかる。

4. クラスター分析のための近傍法

クラスター分析はたくさんの要素の中から似たもの同士をまとめるように分類する作業であるが、クラスター分析を行うためには、それぞれの個体間についてある尺度で定義された属性の距離を求めておく必要がある。単語間の関連性を定義する

表 1 回答者の構成

グループ	対象	人数
A	河川工学研究室	12
	土木系大学院1年生	11
B	土木系3年生	63
C	1年生	77
		計163

表 2 単語の頻度分析

順位	A		B		C	
	単語	頻度	単語	頻度	単語	頻度
1	ダム	317	ダム	787	ダム	796
2	水	119	県	238	自然	343
3	治水	87	知事	222	水	326
4	建設	82	考え	213	考え	267
5	県	81	水	200	県	198
6	考え	80	建設	185	建設	186
7	環境	75	思う	182	川	163
8	河川	73	自然	167	治水	161
9	思う	62	脱ダム宣言	152	環境	154
10	脱ダム宣言	56	川	146	知事	149
11	問題	53	治水	139	必要	145
12	知事	50	河川	138	思う	143
13	川	44	環境	135	河川	140
14	自然	41	必要	94	脱ダム宣言	137
15	洪水	40	利水	90	日本	97
16	必要	39	問題	85	問題	94
17	利水	36	日本	66	利水	80
18	日本	34	住民	62	多く	61
19	対策	31	対策	55	洪水	57
20	住民	30	改修	54	影響	51

類似度の生成にあたっては、宮本ら¹⁾が近傍法を用いた手法によって成果を上げている。近傍の概念を用いた類似度は、互いに接近して用いられることの多い単語同士ほど関連性が高いという考え方に基づいている。宮本ら¹⁾は助詞などを除いた意味を持つ単語のみの文字列に対して近傍の概念を適用しているが、本研究では助詞や句読点なども含めた文字列に対して近傍法を用いている。

次のような文章を例にとる。「安全な生活にはダムは欠かせない。しかし、ダムを造れば河川環境や生態系に影響を及ぼす。環境への影響を考慮すれば、やはりダムを造るべきではないと思う。」文頭から注目する文字が出現するまでの文字数を「文字番地」と呼ぶことにし {環境₂₉} {環境₄₃} のように添え字で表す。分析の対象となる単語とその文字番地を表3に示す。{環境₂₉} を中心とした前後10文字ずつの集合 {し、ダムを造れば河川環境や生態系に影響を及ぼす。環境₂₉} を中心とする大きさ10の近傍と定義し、 U_{10} {環境₂₉} と表す。類似度を生成する前の段階として、接近度Pを定義しておく必要がある。例えば環境と影響に注目すると、{環境₂₉} と影響の接近度を $P(\text{環境}_{29}, \text{影響}) = \{U_{10} \text{ {環境}_{29}} \text{ に属する影響の個数} \}$ と定義する。すると $P(\text{環境}_{29}, \text{影響})=1$ 。すなわち文字番地 $29-10=19$ から文字番地 $29+10=39$ の間に {影響} の文字番地が含まれていればその個数が数えられる。同様に {環境₄₃} について考えると $P(\text{環境}_{43}, \text{影響})=2$ 、環境と影響の接近度は $P(\text{環境}, \text{影響}) = \sum P(\text{環境}_i, \text{影響}) = 1+2=3$ となる。さらに、類似度 $S(\text{環境}, \text{影響}) = P(\text{環境}, \text{影響}) / \{ \text{環境の個数} + \text{影響の個数} \}$ と定義し、 $S(\text{環境}, \text{影響}) = 3 / (2+2) = 0.75$ となる。

表3 単語と文字番地

単語	文字番地		
安全	1		
生活	4		
ダム	8	21	59
欠かせない	11		
造る	24	62	
河川	27		
環境	29	43	
生態系	32		
影響	36	47	
及ぼす	39		
考慮	50		
思う	71		

先にも述べたように、類似度Sは文章中で接近して出現する単語同士ほど大きくなる。 $S(\text{環境}, \text{影響}) = S(\text{影響}, \text{環境})$ であり、類似度Sには対象性がある。また、類似度Sは $0 \leq S \leq 1$ 、 $S(\text{環境}, \text{環境}) = 1$ を満たす。

実際の文章に適用する際は近傍の大きさを20～40の範囲で検討し、最適な近傍を適用する。近傍の大きさを小さく取ると、例えば {環境} と {欠かせない} とでは文字番地を占有している値に差があるため接近度が公平に得られない可能性がある。また、異なる回答者の記述した単語が同じ近傍に属さないよう、それぞれの回答者の間には十分な空白を挿入する。近傍の大きさの変化に対する類似度やクラスターの挙動は興味深いものであり、十分な検討を重ねる必要があると思われる。

5. 群平均法を用いたクラスター分析

近傍法による類似度に基づいて、記述された単語の群平均法を用いた階層的クラスター分析²⁾を行う。グループA、B、Cのそれぞれについて個別に分析を行い、それらの間の傾向を比較する。得られたクラスターに含まれる単語の構成について考察することで、グループ全体としての意向を割り出すことができる。また、「脱ダム」宣言に対する意識を形作る要因とその構造の形成過程をつかむことが可能となる。^{3), 4)}

クラスター分析の結果と、それに対する考察については本学会の発表の場で述べることにする。

参考文献

1) 宮本定明ら：自由連想法による環境意識調査データのクラスター分析のための近傍法，システムと制御，Vol. 32, No. 4, pp. 255～263, 1988.
2) 宮本定明：階層的クラスター分析の方法－I，システム/制御/情報，Vol. 33, No. 8, pp. 424～430, 1989.
3) 小浜明ら：河川流域の特徴が住民の河川に対する意識に与える影響～近傍法によるアンケート自由記述分の分析～，環境システム研究，Vol. 22, pp. 9～14, 1994.
4) 大井紘ら：生活環境に関する住民の認知の広がり構造，土木学会論文集，Vol. 389/IV-8, pp. 83～92, 1988.