

# 要素分析의 理論과 應用

韓 斗 完  
(K. I. E. T. 責任研究員)

## 1. 序 言

情報學은 圖書館學뿐만 아니라 컴퓨터공학, 通信工學, 數學, 統計學, 言語學, 心理學, 經營學 등多數의 他學問分野와 밀접한 關係를 갖고 있는 多學問的인 分野이다.<sup>1)</sup> 情報學이 이렇게 多學問的이고, 廣範圍한 原因은 그 實用性이나 社會에 대한 그 貢獻이 또한 크기 때문이라 생각된다. 情報學의 實用性은 圖書館이나 情報機關에서 情報를 合理的으로 蒐集하고 效果的으로 加工·處理하여 利用者에게 有用한 情報를 容易하게 提供할 수 있도록 하는데 있다고 볼 수 있다. 이렇게 情報學이 多學問的인 分野인만큼 業務處理나 問題解決에 있어서 그 接近方法도 多樣하고, 應用可能한 方法도 여러 가지가 있으리라 생각된다.

여기에서는 情報學에 應用하여 效率的으로 利用할 수 있는 여러 方法 가운데 하나인 要素分析(factor analysis)에 대해서 알아보기로 한다. 要素分析은 他學問分野에서 開發되고 應用되어 왔으며, 情報學分野에도 應用되고 있다. 要素分析은 특히 文献의 自動分類, 情報學의 教科課程決定, 學術雜誌의 클러스터링(clustering) 및 기타 여러 측면에서 應用될 수 있는 統計的 技法이다. 本橋에서는 要素分析의 理論에

대해서 概略的으로 알아보고, 아울러 그 應用에 대해서는 實例를 들어 살펴보기로 한다.

## 2. 要素分析의 一般的 理論

### 2.1 要素分析의 歷史的 背景

要素分析은 일반적으로 C. E. Spearman이 “General Intelligence Objectively Determined and Measured”<sup>2)</sup>라는 論文을 發表한 1904년부터 시작된 것으로 여겨지고 있다. 스피어먼은 單一假說變數가 일반적으로 一聯의 認識試驗의 相互關係를 說明한다고 주장했으며, 이 變數가 그의 유명한 要素(factor) “g”이다. 스피어먼이 생각하고 있는 試驗에 대해서는 이 假說이 合理的이었으나, 相關매트릭스가 많이 利用됨에 따라 스피어먼의 假說은 事實을 너무 單純化 하였다는 것이 나타나게 되어 多要素概念이 開發되었다. 美國의 서스토운(L.L. Thurstone)과 英國의 C. Burt 및 G. H. Thomson은 多要素概念開發의 가장 活動的인 開拓者들이었다. 그러나 心理學者인 이들이 要素를 評價하기 위해 開發한 統計方法은 다소 正確性이 결여된 점이 있었다. 要素評價問題를 다루기 위하여 數理統計學者에 의한 첫번째의 試圖는 19

1) 鄭瑛美, “情報學이란 무엇인가?”, 「情報産業」通卷 17号, 1982. 11, pp. 34-37.

2) C. E. Spearman, “General Intelligence Objectively Determined and Measured”, American Journal of psychology, Vol. 15, 1904, pp. 201-293.

40년 D. N. Lawley가 “The Estimation of Factor Loadings by the Method of Maximum Likelihood”<sup>3)</sup> 라는 論文을 發表함으로써였다. 1940년 이후 D. N. Lawley가 要素를 다루는 그의 論文들을 發表하고, 또한 W. G. Howe, T. W. Anderson, H. Rubin 및 C. R. Rao 등이 이 分野의 著作을 發表하였으며, 要素分析에 관한 現代的 텍스트북으로서 H. H. Harman의 著書 Modern Factor Analysis<sup>4)</sup> 와 D. N. Lawley와 A. E. Maxwell의 共同著書인 Factor Analysis as a Statistical Method<sup>5)</sup> 가 있다.<sup>6)</sup>

한편 應用側面에서 보면, 1960년 이전에는 要素分析이 心理學研究에 두드러지게 應用되었으나, 1960년대에 들어와서는 要素分析이 經營, 社會 및 行動科學과 情報學에도 應用되기 시작하여<sup>7)</sup> 지금은 여러 분야에서 應用되고 있다.

## 2.2 要素分析의 概念

要素分析이란 形態를 간결하게 함으로써 관찰된 데이터의 數를 줄이기 위한 一聯의 統計的 技法을 말한다. 이 要素分析은 變數에 대한 관찰된 개개의 反應을 취하여 要素(factor)라고 불리우는 적은 수의 合成變數의 線型組合을 形成하게 된다. 合成變數인 要素의 線型組合은 變數들간에 존재하는 線型依存關係에 根據가 있는 것이다. 데이터를 構造的으로 생각하여 발

견된 많은 수의 變數를 發見된 各要素의 相對的 重要性에 따라 加重值를 매긴 적은 수의 合成變數로 줄이며, 要素分析技法은 알려진 變數間의 關係를 알아내고 새로운 데이터와 一致하는 變數範圍를 決定하는 것 외에, 데이터構造의 探索, 變數間에 존재하는 關係에 관한 새로운 假說의 設定 및 科學的 parsimony (scientific parsimony)<sup>8)</sup> 를 成就하는 手段으로서도 利用될 수 있다.<sup>9)</sup>

## 2.3 要素分析의 方法

要素分析의 方法에는 여러 가지가 있으나, 이들의 歷史的인 重要性이나, 특별한 必要性이나, 一般的인 應用性 등의 問題로 사람마다 각자에게 합당한 方法을 택하여 利用하게 되는데, 이 要素分析의 첫째의 必要條件은 많은 變數間의 相互關係를 적절하게 說明해야 하는 것이다. 그러기 위해서는 훌륭한 統計的 技法이 필요하다. 統計的 技法이 만족할만 하면, 要素結果를 單純化 하는 것이 바람직하게 되며, 이들 結果는 그 対象分野에 있어서 더욱 意味있고 說明이 가능해 질 수 있게 될 것이다.

그러면 이렇게 利用 가능한 要素分析의 方法은 일반적으로 어떠한 것들이 있나를 살펴보면 다음과 같다.<sup>10)</sup>

- 1) Principal components method.

3) D. N. Lawley, “The Estimation of Factor Loadings by the Method of Maximum Likelihood”, Royal Society of Edinburgh, Proceeding, Vol. 60, 1940, pp. 64-82.

4) H. H. Harman, Modern Factor Analysis, 3rd ed., Chicago: The University of Chicago Press, 1976.

5) D. N. Lawley & A. E. Maxwell, Factor Analysis as a statistical Method, London: Butterworth, 1963.

6) D. L. Sills ed., International Encyclopedia of Social Science, Vol. 5, New York: The Macmillan Company & The Free Press, 1968, p. 276.

7) A. Kent & H. Lancour ed., Encyclopedia of Library and Information Science, Vol. 8., New York: Marcel Dekker, Inc., 1972, p. 347.

8) H. H. Harman, op. cit. p. 280. 에서 “파시머니는 무한한 해결가능한 방법 가운데에서 선호된 해결방법을 선택하는데 있어서 기본적 기준의 하나이며, 단순구조원리의 基底인 것이다” 라고 말한다.

9) A. Kent & H. Lancour, op. cit. p. 349.

10) H. H. Harman, op. cit. p. 108.

- 2) Principal-factor method.
- 3) Centroid method.
- 4) Triangular decomposition method.
- 5) Maximum-likelihood method.
- 6) Canonical factor analysis method.
- 7) Minres method.
- 8) Image method.
- 9) Alpha method.
- 10) Multi-group method.
- 11) Unifactor method.
- 12) Two-factor method.
- 13) Bi-factor method.
- 14) Derived factor method.

上記와 같이 여러 가지 要素分析의 方法이 있으나 여기에서는 各方法의 具體的인 內容에 대한 說明은 省略하고, 要素分析의 一般的인 內容에 대해서만 알아보기로 한다.<sup>11)</sup>

要素分析에 있어서는 一般的으로 入力데이터를 가지고서 2次元에서는 표 1 과 같은 매트릭스를 形成하게 되는데, 이 內容을 調査한 다

음 데이터에 포함된 모든 情報를 담은 또다른 單純化되고 크기가 작아진 매트릭스가 필요하게 된다. 즉, 파시머니(parsimony)의 原理는 變數의 數는 적어졌지만 원래의 데이터를 충실하게 나타낼 수 있는 測定表를 要求하게 되는 것이다. 이렇게 원래보다 훨씬 줄어든 變數를 가진 새로운 測定表를 만들어야 하는데, 이것은 式 (1) 과 같은 線型組合을 形成함으로써 가능하게 되고, 이들 새로운 값(score)은 새로 形成된 매트릭스 표 2'에 나타난 바와 같다.

$$Y = w_1 X_{A1} + w_2 X_{B1} + w_3 X_{C1} + \dots \quad (1)$$

표 2.1은 표 1의 內容을 모두 포함하고 있지만 표 1'보다 훨씬 간단한 매트릭스를 形成하고 있다. 여기에서 이것을 綜合次元(Synthetic dimension)이라고 한다면, 이 綜合次元에서는 各元(individual)이 요소값(factor score)이라고 하는 값을 가지는데, 이 값을 직접 관찰할 수는 없지만, 원래 測定值의 相互關係를 함으로써 얻어질 수 있는 것이다.

要素分析에서는 예를 들어보면,  $X_{A1}, X_{B1},$

표 1.

대 상

시 험	1	2	3	...	n
A	$X_{A1}$	$X_{A2}$	$X_{A3}$	...	$X_{An}$
B	$X_{B1}$	$X_{B2}$	$X_{B3}$	...	$X_{Bn}$
C	$X_{C1}$	$X_{C2}$	$X_{C3}$	...	$X_{Cn}$
⋮	⋮	⋮	⋮	⋮	
N	$X_{N1}$	$X_{N2}$	$X_{N3}$	...	$X_{Nn}$

표 2.

대 상

요소	1	2	3	...	n
$\alpha$	$Y_{\alpha 1}$	$Y_{\alpha 2}$	$Y_{\alpha 3}$	...	$Y_{\alpha n}$
$\beta$	$Y_{\beta 1}$	$Y_{\beta 2}$	$Y_{\beta 3}$	...	$Y_{\beta n}$
$\gamma$	$Y_{\gamma 1}$	$Y_{\gamma 2}$	$Y_{\gamma 3}$	...	$Y_{\gamma n}$

11) A. Kent & H. Lancour, op. cit. pp. 351-353.

$X_{ci}$  등으로 나타나는 元(individual) 1에 대한 많은 시험값(test score)을  $X_{\alpha i}, X_{\beta i}, X_{\gamma i}$  등으로 나타나는 元(individual) 1에 대한 적은 요소값(factor score)으로 나타내고 있다. 各元은 綜合값(synthetic score)의 여러 비율을 합함으로써 원래의 값으로 再構成될 수 있는 한세트(set)의 要素값을 가질 것이다. 이들 比率는 常數이며, 이것을 팩터로딩(factor loading)이라 한다. 이 팩터로딩은  $-1$ 과  $+1$  사이의 값을 가지며, 원래 관찰된 값들과 各綜合要素(synthetic factor)와의 相互關係를 나타내고 있다.

各線型組合은 要素값으로 나타날 수 있으며, 데이터에서 發生한 各 요소의 要素값이 있을 것이며, 式 (1)은 元(individual)  $i$ 의  $m$ 要素에 대해서는 式 (2)와 같이 一般化될 수 있을 것이다.

$$Y_{mi} = w_{m1}X_{\alpha i} + w_{m2}X_{\beta i} + w_{m3}X_{\gamma i} + \dots (2)$$

그러므로 시험 A에 있어서 元(individual)  $i$ 의 원래의 값을 再構成함으로써 要素값의 加重線型組合이 形成된다. 시험 A에 있어서 元(individual) 1에 대한 원래의 값  $X_{A1}$ 은 요

소값으로부터 式 (3)과 같이 再構成될 수 있다.

$$X_{A1} = w_1X_{\alpha 1} + w_2X_{\beta 1} + w_3X_{\gamma 1} + \dots (3)$$

또한 上記 매트릭스에서 얻은 各要素와 對象間의 關係를 分明하게 나타내는 座標로 表示할 수 있으며, 이 座標上에 相互關係를 나타내어 幾何學的으로 相互關係의 分析을 하게 된다. 이 分析에 있어서 明確한 分析을 하기 위하여 座標軸을 回轉시킬 수가 있다. 座標軸을 回轉시킬 것이냐, 回轉시키지 않을 것이냐는 分析者가 그때의 경우에 따라 決定할 問題이다. 만약 座標軸을 回轉시키는 方法을 택하게 될 때에는  $-1$ 에서  $+1$ 의 값을 갖는 주어진 次元의 空間에서 試驗(test)의 로딩(loading)을 나타내는 點은 움직이지 않고 원래의 位置를 그대로 維持하게 되고, 座標軸은 原點을 中心으로 하여 回轉하게 된다. 이때 주의해야 할 점은 試驗點(test point)은 원래와 똑같이 維持되어지고 試驗(test) 相互關係가 정확히 나타나야 한다.

이러한 回轉方法에는 直角回轉方法과 斜角回轉方法의 두 가지 方法이 있다.

直角回轉方法은 座標軸間의 角度를 直角으로

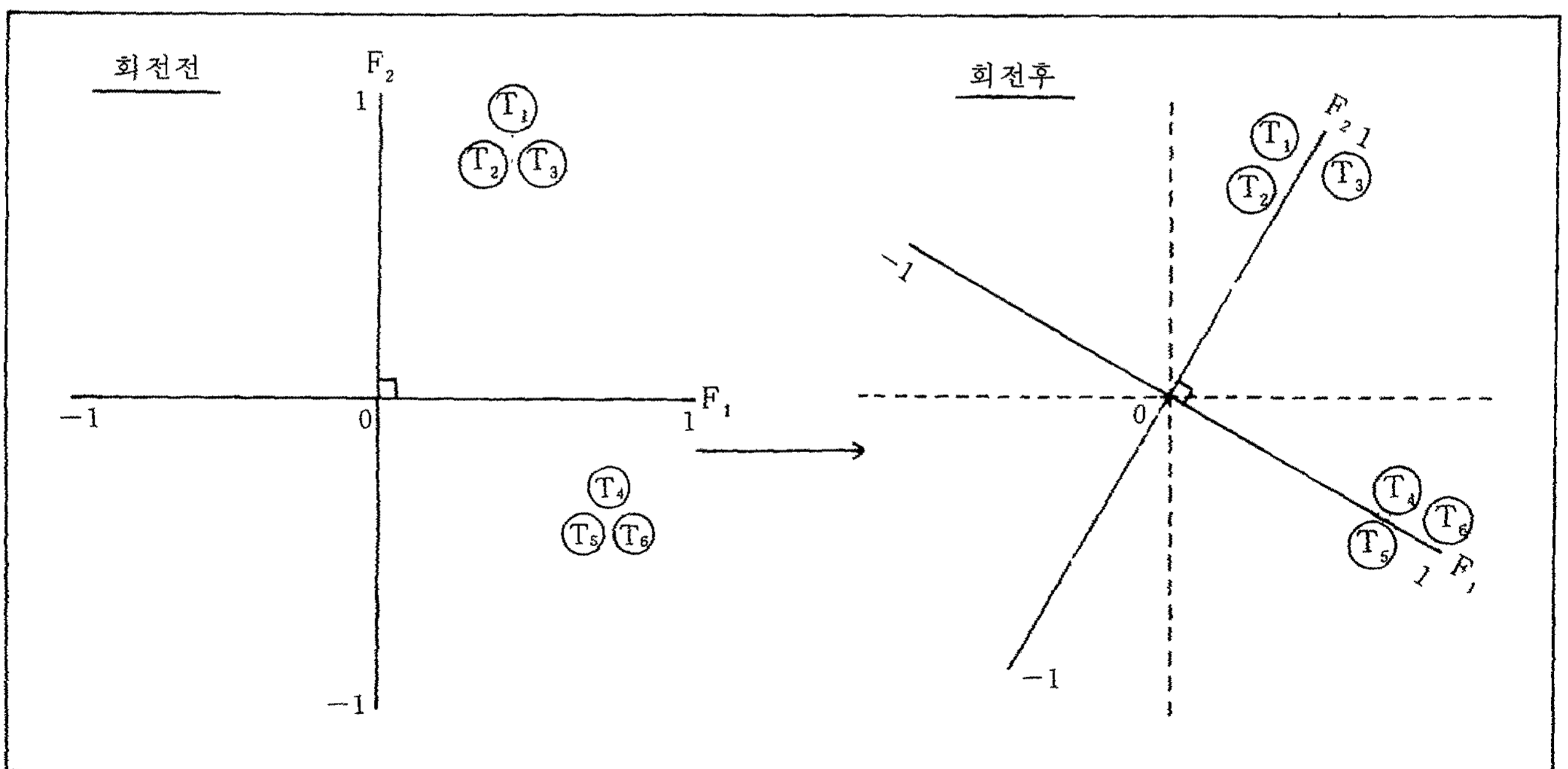


그림 1. 직각회전방법



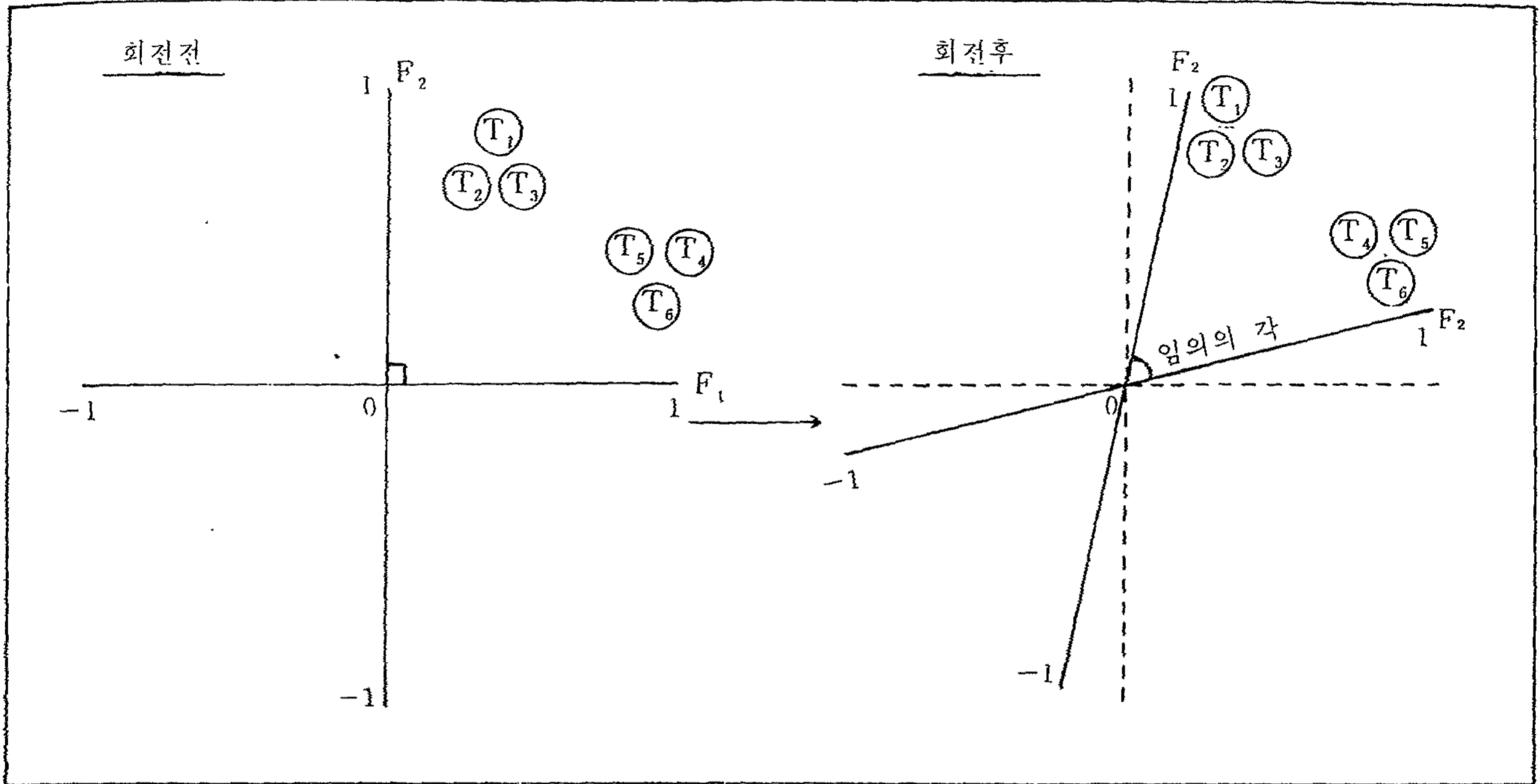


그림 2. 사각회전방법

維持하면서 回轉하는 方法이고, 斜角回轉方法은 座標軸間의 角度를 直角으로 維持하지 않고 그때의 경우에 따라 임의의 각을 취하면서 回轉하는 方法이다. 이들 方法을 그림으로 나타내어 보면 :그림 1 및 그림 2 와 같다.

이러한 要素分析은 實際에 있어서는 컴퓨터의 發達로 인하여 수 많은 컴퓨터 프로그램이 開發되어 있어 既存프로그램을 利用할 수 있다. 例로서는 SPSS<sup>12)</sup>의 利用으로 바람직한 結果를 얻을 수가 있다.

이렇게 概略적으로나마 要素分析에 대하여 알아 보았다. 그러면 이러한 要素分析은 어떤 分野에 應用될 수 있나를 살펴 보기로 한다.

### 3. 要素分析의 應用

要素分析은 처음에는 單一要素概念으로 시작

하여, 多要素技法으로 發達한 科學的 方法인데, 지금은 自然科學 및 社會科學의 여러 分野에 걸쳐 널리 應用되고 있다 특히 情報學分野에서는 自動分類, 教科課程決定, 學術雜誌의 클러스터링 (clustering), 利用者研究 및 圖書館이나 情報機關의 職務分析 등에 應用될 수 있으리라 생각된다. 그러면 여기에서는 몇 가지의 例를 들어서 說明해 보기로 한다.

#### 3.1 自動分類

H. P. Luhn은 文獻의 單語를 統計적으로 分析함으로써 그 內容을 把握할 수 있다고 하였으며,<sup>13)</sup> M. E. Maron은 自動索引의 問題에 公式化된 統計方法으로 접근하였다.<sup>14)</sup> 즉, M. E. Maron은 文獻內容의 各單語와 主題카테고리 (subject category)와의 確率關係를 決定하고 그 單語를 포함하는 文獻이 속할 카테고리

12) Statistical Package for the Social Sciences.

13) H. P. Luhn, "A Statistical Approach to Mechanized Encoding and Searching of Literary Information", IBM Journal of Research and Development, Vol. 1, 1957, pp.309-317

14) M. E. Maron, "Automatic Indexing: An Experimental Inquiry", Journal of Association for Computing Machinery, Vol. 8, 1961, 98 pp.407-417.

를 豫測하는 關係를 利用하였다. 또한 H. Borko는 文獻의 分類카테고리를 抽出해내기 위하여 要素分析技法을 應用하여 自動分類에 대한 研究를 하였다.<sup>15)</sup>

H. Borko와 M. Bernick는 上記와 같은 研究들을 基礎로 하여 分類시스템과 自動文獻分類 節次와의 關係를 나타내는 아래와 같은 實驗을 하였다.<sup>16)</sup> 文獻의 領域을 가장 잘 記述하는 카테고리를 抽出할 수 있다면 分類業務는 간단해지고, 실제로 分類를 自動化 할 수 있을 것이라고 假定하였는데, 이러한 假定을 試驗하기 위해서는 다음과 같은 것이 필요할 것이다.

- 1) 實驗에 根據를 두고 數學的으로 抽出된 分類시스템의 構成.
- 2) 文獻을 自動的으로 이들 카테고리로 分類할 수 있는 節次的 考案.
- 3) 어떤 基準에 比較함으로써 分類의 正確性 決定.

또한 自動分類에 要素分析을 應用함에 있어서 文獻에 포함된 單語에 의해서 文獻을 分類할 수 있고, 유사한 單語群을 포함하고 있는 文獻은 같은 카테고리에 속한다고 假定할 수 있다. H. Borko와 M. Bernick는 260개의 抄錄을 선택한 다음 키펀치(key punch)하여 컴퓨터 프로그램<sup>17)</sup>에 의해 각 文獻에 나타난 90개 索引語의 發生數를 計算하였다. 이 頻度數에 따라 水平軸에는 索引語를, 水直軸에는 確認된 文獻數를 나타내는 데이터매트릭스를 形成하였는데 이 매트릭스의 各元(cell)은 주어진 文獻에서의 特定用語의 發生數를 나타내는 것이다. 이 매트릭스에 나타난 頻度數에 따라 90개 개개의 索引語와 다른 개개의 用語와의 相關係數가 計算되었으며, 그 結果는 90 × 90 相關매트릭스로 나타나게 되었다. 이렇게 要素分析을 한

結果, 21개 要素가 分類카테고리로 가장 적합하고, 가장 의미있다고 決定되었다. 이들 要素들은 各要素에 나타난 로딩이 높은 單語에 根據를 두었다고 說明하고 있다.

分類表가 設定되었고, 自動分類節次가 決定되었다면, 그 方法은 精確한 分類를 할 수 있고, 컴퓨터에 의한 데이터 分析으로 自動的으로 遂行할 수 있도록 간단하고 機械的이어야 한다. 文獻에 포함된 單語에 根據를 둔 豫測函數와 그 要素로딩이 開發되었는데, 이렇게 開發된 豫測公式은 式 (4)와 같다.

$$P=f(L_1 \times T_1+L_2 \times T_2+\dots+L_n \times T_n) \dots (4)$$

여기에서 P는 豫측된 分類, Ln은 주어진 카테고리에 대한 用語 n의 수정된 要素로딩, Tn은 n번째 用語의 發生數이다. 여기서 컴퓨터 처리에 利用할 수 있는 데이터는 다음과 같다.

- 1) 文獻의 完全抄錄.
- 2) 선택된 90개의 색인어.
- 3) 21개의 주제카테고리.
- 4) 각 카테고리의 각 주제어의 要素로딩.
- 5) 각 카테고리의 각 주제어의 수정된 要素로딩.

그리고 文獻의 自動分類의 節次는 다음과 같다.

- 1) 機械可讀型的 各文獻은 컴퓨터에 의해 分析되어 各文獻에서의 索引語와 그 發生數의 리스트가 作成된다.
- 2) 索引語를 포함한 카테고리는 抄錄에서 單語의 發生係數와 카테고리에서 수정된 要素로딩과 같은 값이 할당된다.
- 3) 各索引語가 고려된 후에 最高數值를 가진 카테고리가 그 文獻의 가장 有力한 主題領域으로 선택된다.

15) H. Borko, "The Construction of an Empirically Based Mathematically Derived Classification System", Proceedings of Spring Joint Computer Conference, Vol. 21, 1962, pp. 279-289.

16) H. Borko & M. Bernick, "Automatic Document Classification", Journal of Association for Computing Machinery, Vol. 10, 1963, pp. 151-162.

17) J. C. Olney, FEAT, An Inventory Program for Information Retrieval, FN-4018, Santa Monica, California, System Development Corp., 1960.

### 3.2 教科課程決定

要素分析은 情報學의 教科課程의 통일된 概念을 파악하는 데에도 利用될 수 있다. Georgia Institute of Technology에서는 1968년 57개 機關에서 가르치는 811개 科目을 分析하였으며, 美國情報學會 教科課程委員會는 教科課程의 開發 및 科目의 提案이라는 觀點에서 美國과 캐나다의 情報學프로그램의 範圍 및 特性을 決定하기로 하였다.<sup>18)</sup> 이 課業을 遂行하기 위하여 情報學關係科目에 관한 믿을만한 情報를 얻어내도록 設問이 開發되었으며, 여기에서 要求된 데이터는 科目의 水準, 必須科目, 使用된 教材, 포함된 主題, 開設된 頻度 등이 포함되어 있다. 45개 學校에서 應答을 하였는데, 185개 科目 및 242개 主題에 관한 情報를 提供하였다. 이들 데이터를 클러스터링 (clustering)하는 몇 가지의 方法을 使用하였지만, 이 分野에 포함된 主題의 多樣性과 散漫性 때문에 확실한 結果에 이르는 것은 매우 어려웠다. 그래서 設問 結果의 妥當性을 調査하기 위해 專門家の 워크 샵을 가졌는데, 여기에서 情報學 教科課程에 貢獻이 있는 9개의 要素 및 碩士課程의 核心을 이루고 있는 7개 科目을 確認하게 되었다.

9개 要素는 心理學 및 行動科學, 語學 및 言語學, 經營學, 統計學, 圖書館學, 시스템論, 數學, 情報 및 커뮤니케이션理論, 電算學 및 自動化論이고, 7개 科目은 情報學概論, 시스템理論 및 應用, 情報學에서의 數學的 方法, 컴퓨터 組織 및 프로그래밍시스템論, 抄録·索引·目錄, 情報커뮤니케이션理論과 研究方法論이다.

그러나 여기에서는 要素分析方法의 應用이 成功的이 아니었는데, 傳統的 要素分析 및 클러스터링技法은 分析에서 모든 主題를 포함하는 큰 要素를 形成하게 되었다.

### 3.3 學術雜誌의 클러스터링 (clustering)

A. Zavala와 H. Van Cott는 要素分析을 利用하여 다른 群의 雜誌보다 共通性이 많은 雜誌들을 한 群으로 클러스터링할 수 있었다.<sup>19)</sup> 이들의 研究는 要素分析技法을 利用하여 어떤 學問의 定期刊行物의 構造를 抽出할 수 있다는 것을 나타내 주었다. 지금 이 研究에 대해서 간단하게나마 알아 보기로 한다.

이 研究에 利用된 데이터는 S. Keenan과 P. Atherton이 蒐集하였다. 그런데 S. Keenan과 P. Atherton은 Physical Abstracts의 1961년도 발행분의 주제카테고리에 의해 抄録總數表를 作成했는데, 39개국에서 發行된 총 405종의 雜誌에 대한 20,287건의 抄録이 表에 실렸으며, 주제카테고리는 127개였고, 이 研究에는 78개의 카테고리가 採擇되었다. 405종의 雜誌 가운데 1961년도 雜誌抄録의 85% (17,244건)를 나타내고, 記事抄録의 數에 있어서 上位에 記錄되는 74종의 雜誌가 있는데, 이들 雜誌는 美國을 포함한 17개국에서 發行되는 것이었다.

이 研究에서는 4가지의 分析技法을 利用하여 컴퓨터 CDC-3600에 의해 수행되었다. 그 4가지의 分析技法은 雜誌分析, Physical Abstracts分析, 키난-아데르톤 (Keenan-Atherton)分析 및 頻度數 데이터分析인데 여기에서는 例로서 키난-아데르톤分析에 대해서만 알아보기로 한다.

키난-아데르톤의 22개 카테고리에 대한 相關係數는 表 3에 나타난 바와 같다. 한 쌍의 카테고리간의 陽의 높은 相關係數는 그 카테고리들이 같은 잡지에 나타남을 보여주고, 陰의 높은 相關係數는 같이 나타나지 않음을 나타내어 주고 있다. 表 3의 매트릭스는 對角線의

18) J. Belzer, et al, "Curricula in Information Science: Analysis and Development", Journal of the American Society for Information Science, Vol. 22, 1971, pp. 193-223.

19) A. Zavala & H. Van Cott, "Extracting the Basic Structure of Scientific Literature", American Documentation, Vol. 19, 1968, pp. 247-258.

표 3. 키난 - 아테르톤 카테고리 의 상관계수

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII
1. Physics, General	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
2. Mathematics	22	05	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
3. Astrophysics	26	00	32	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
4. Mathematical physics	35	45	14	-24	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
5. Mechanics	-15	07	00	00	00	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
6. Fluids & Gases	06	24	-01	07	39	09	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
7. Acoustics	17	22	09	11	30	34	20	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..
8. Optics	27	11	11	37	06	23	26	16	..	..	..	..	..	..	..	..	..	..	..	..	..	..
9. Heat & Thermal phenomena	17	01	56	17	23	04	25	17	19	..	..	..	..	..	..	..	..	..	..	..	..	..
10. Geophysics	10	22	00	-31	46	19	15	25	-24	09	..	..	..	..	..	..	..	..	..	..	..	..
11. Biophysics	03	29	-12	-19	28	15	18	25	23	03	07	..	..	..	..	..	..	..	..	..	..	..
12. Techniques & Materials	22	12	30	42	10	-06	-01	19	25	30	-07	..	..	..	..	..	..	..	..	..	..	..
13. Nuclear physics	20	18	35	37	06	06	08	16	16	37	-09	-06	73	..	..	..	..	..	..	..	..	..
14. Elementary particles	22	00	29	25	-02	05	13	14	13	37	-03	-12	35	49	..	..	..	..	..	..	..	..
15. Cosmic rays	26	14	21	32	-04	07	05	28	33	20	-12	07	34	41	13	..	..	..	..	..	..	..
16. Atomic & Molecular physics	-01	14	02	02	15	02	12	18	15	17	-03	22	05	03	-14	-03	..	..	..	..	..	..
17. Physical chemistry	03	25	30	39	08	04	17	26	23	27	-01	15	47	48	27	36	..	..	..	..	..	..
18. Electricity & Magnetism	01	16	-03	21	09	31	21	26	18	00	12	09	10	06	21	-11	..	..	..	..	..	..
19. Solid State physics	10	22	-01	30	02	26	29	40	40	-01	-12	11	08	20	07	30	19	..	..	..	..	..
20. Electrical properties, Solids	01	08	11	37	-13	36	12	36	37	07	-06	06	15	25	12	37	28	32	..	..	..	..
21. Optical properties, Solids	13	06	09	59	-18	24	13	29	49	06	-31	17	29	31	13	39	31	36	62	..	..	..
22. Magnetic properties, Solids	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..	..

표 4. 키난 - 아테르톤 카테고리 의 주축요소분석

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX	XXI	XXII
1. Physics, General	34	01	-37	36	28	-31	-46	-32	09	-12	02	-04	-06	13	-04	09	-07	-19	-11	11	-02	07
2. Mathematics	30	47	-28	21	18	27	-14	09	-19	38	41	01	-15	02	09	-13	17	05	-01	03	02	04
3. Astrophysics	40	-29	-43	-13	-35	-25	00	-22	-05	43	-09	06	-11	-09	-03	03	-27	15	09	13	02	03
4. Mathematical physics	63	-39	21	14	-22	-06	-32	00	-03	-03	-03	-24	-14	03	12	-28	-07	-06	09	-11	-10	-14
5. Mechanics	14	60	-61	13	06	-07	-14	03	01	02	-06	15	09	-18	-06	02	-17	06	06	-32	-03	-06
6. Fluids & Gases	28	37	28	-57	00	-08	17	-14	36	06	34	-08	-04	-09	01	-03	-17	-19	02	-06	-06	06
7. Acoustics	32	44	-14	10	-06	-31	05	68	01	-12	34	-05	-01	-01	17	-07	-17	-01	-02	11	01	07
8. Optics	51	44	-04	-27	13	02	05	-28	-19	-29	-23	00	-23	-31	07	-10	17	02	-06	00	07	-02
9. Heat & Thermal phenomena	59	08	21	36	03	-29	13	-01	39	04	11	-15	24	-20	-15	05	17	12	03	07	08	-06
10. Geophysics	43	-10	-51	-12	-38	-33	27	00	04	06	-07	12	01	13	12	07	17	-16	-03	-09	03	00
11. Biophysics	-10	52	-44	-46	00	00	-22	-08	-03	-03	-10	-18	27	25	-06	-12	04	03	21	05	06	-03
12. Techniques & materials	18	51	-03	32	12	29	45	-14	29	10	-25	00	-19	25	01	-01	-07	00	02	05	-05	-00
13. Nuclear physics	60	-30	-31	04	-06	45	-08	-03	27	-18	-04	15	14	-07	05	-12	04	09	07	02	-13	18
14. Elementary particles	65	-33	-30	-13	-02	32	-03	12	06	-12	19	29	10	01	02	03	-11	-13	-02	11	13	-18
15. Cosmic rays	45	-39	-34	-23	27	-16	19	18	00	-25	17	-07	-27	18	-24	00	01	22	-02	-07	-02	00
16. Atomic & Molecular physics	61	-23	-06	02	44	-03	22	-14	-18	00	06	-22	14	-01	35	23	-07	06	06	-05	-03	-01
17. Physical chemistry	13	34	01	33	-50	-03	31	-26	-35	-32	28	00	11	08	-06	-06	-11	01	04	01	-01	03
18. Electricity & Magnetism	63	-06	-16	-03	-17	40	09	17	-18	13	-15	-39	04	-08	-26	12	-07	-16	-09	-08	02	04
19. Solid State physics	33	43	30	-13	-45	14	-40	02	17	-11	05	-07	-08	09	09	31	17	19	-07	02	-02	-03
20. Electrical properties, Solids	62	25	44	00	18	-10	-08	14	-23	05	-05	26	-01	-05	-19	15	06	-07	25	04	-15	-03
21. Optical properties, Solids	64	08	46	-24	10	-10	01	-06	-16	16	-09	13	23	17	-04	-16	-03	08	-32	03	-08	-01
22. Magnetic properties, Solids	73	-03	52	10	02	-01	-05	-03	01	03	-12	12	06	16	03	-02	-17	-01	09	-18	24	13



표 5. 7개 요소에 대한 배리맥스 로딩

	요 소						
	I	II	III	IV	V	VI	VII
1 Physics, General	15	65	-36	29	15	06	-27
2 Mathematics	-18	63	06	05	-05	28	24
3 Astrophysics	72	05	-04	03	04	30	-13
4 Mathematical physics	24	-11	-21	63	-17	41	-25
5 Mechanical	16	83	07	-18	-04	01	20
6 Fluids & Gases	01	05	77	21	-06	-06	04
7 Acoustics	28	43	17	24	-04	-13	24
8 Optics	05	36	55	26	03	18	12
9 Heat & Thermal phenomena	19	13	-07	70	12	02	27
10 Geophysics	83	11	09	-01	12	21	14
11 Biophysics	01	54	19	-11	-11	-05	-16
12 Techniques & Materials	-23	26	13	06	11	09	74
13 Nuclear physics	13	09	-10	12	07	84	-02
14 Elementary particles	22	05	05	15	17	79	-10
15 Cosmic rays	37	03	11	11	58	34	-21
16 Atomic & Molecular physics	06	08	08	13	59	36	00
17 Physical chemistry	27	04	-06	10	31	-05	63
18 Electricity & Magnetism	15	05	13	19	-03	71	20
19 Solid State physics	-01	15	30	32	74	15	01
20 Electrical properties, Solids	-11	11	31	75	02	07	05
21 Optical properties, Solids	00	-08	15	69	02	15	-05
22 Magnetic properties, Solids	-04	-10	13	85	-06	26	05

한쪽만을 利用하고, 構造를 單純化 하기 위해 배리맥스方法 (Varimax method)으로 回轉시켜 표 4에 나타난 바와 같이 主軸解에 따랐으며, 그 結果는 7개의 要素로서 표 5에 나타난 바와 같다. 표 5에서 各主題에 대한 가장 높은 값을 가진 要素를 찾아서 確認하여 보면 各要素에 속하는 主題를 알 수 있다. 표 5에서 各要素에 속하는 主題를 例를 들어보면 다음과 같다.

例 : 要素 I

Geophysics	0.83
Astrophysics	0.72

要素 II

Mechanics	0.83
Physics, General	0.65
Mathematics	0.63
Biophysic	0.54
Acoustics	0.43

4. 結 論

要素分析은 20세기초부터 시작되어 현재에 이

르고 있다. 要素分析의 바람직한 應用 및 活用은 學問的 理論뿐만 아니라, 業務現場에도 可能하여야 할 것이며, 앞으로는 지금보다 더 널리 應用되리라 믿는다. 이렇게 널리 應用될 수 있는 要素分析은 관찰된 데이터를 간소하게 하여 變數相互間의 關係를 測定하는 一種의 統計的 技法인 것이다. 초기에는 복잡한 統計的인 處理 때문에 어려움이 많았으나, 그후 컴퓨터의 出現으로 統計的인 計算處理는 解決이 되었으며, 지금에 와서는 要素分析에 관한 컴퓨터 프로그램이 開發되어 實務에 있어서 既存프로그램을 利用하여 데이터를 處理할 수 있게 되었다.

지금까지 要素分析의 理論에 대해서 概略적으로나마 알아보았으며, 그 應用에 관하여는 實例를 들어 살펴보았는 바와 같이 이러한 要素分析의 技法을 情報學에 應用하여 實務에 適用해 나아갈 수 있으리라 믿는다. 情報學에 있어서 같은 問題를 解決하는데에 여러 가지 方法이 있을 수 있다. 그 모든 方法을 最適의 경우에 活用하여 效果的인 運營이 되어야 함은 의문의 여지가 없는 것이다. 그리고 情報學의 理論이나 應用possible한 方法을 研究하고 發展시

켜 現實性있는 모델로 開發시켜 나아가는 일 또한 무엇보다도 중요하리라 믿는다. 要素分析의 情報學에의 應用도 더욱 研究되고 發展되어야 하리라 생각된다.

〈参考文献〉

1. 鄭瑛美, "情報學이란 무엇인가?", 「情報産業」, 通卷17号, 1982. 11, pp. 34-37.
2. Spearman, C. E., "General Intelligence Objectively Determined and Measured", American Journal of Psychology, Vol.15, 1904, pp.201-293.
3. Lawley, D. N., "The Estimation of Factor Loadings by the Method of Maximum Likelihood", Royal Society of Edinburgh, Proceeding, Vol.60, 1940, pp.64-82.
4. Harman, H. H., Modern Factor Analysis, 3rd ed., Chicago: The University of Chicago Press, 1976.
5. Lawley, D. N., and Maxwell, A. E., Factor Analysis as a Statistical Method, London: Butterworth, 1963.
6. Sills, D. L. ed., International Encyclopedia of Social Science, Vol.5, New York: The Macmillan Company & The Free Press, 1968.
7. Kent, A., and Lancour, H. ed., Encyclopedia of Library and Information Science, Vol.8, New York: Marcel Dekker, 1968.
8. Luhn, H. P., "A Statistical Approach to Mechanized Encoding and Searching of Literary Information", IBM Journal of Research and Development, Vol.1, 1957, pp.309-317.
9. Maron, M. E., "Automatic Indexing: An Experimental Inquiry", Journal of Association for Computing Machinery, Vol.8, 1961, pp.407-417.
10. Borko, H., "The Construction of an Empirically Based Mathematically Derived Classification System", Proceedings of Spring Joint Computer conference, Vol.21, 1962, pp.279-289.
11. Borko, H., and Bernick, M., "Automatic Documentation Classification", Journal of Association for Computing Machinery, Vol.10, 1963, pp.151-162.
12. Belzer, J., "Curricula in Information Science: Analysis and Development", Journal of the American Society for Information Science, Vol.22, 1971, pp.193-223.
13. Zavala, A., and Van Cott, H., "Extracting the Basic Structure of Scientific Literature", American Documentation, Vol.19, 1968, pp.247-258.