Journal of the Korean Data & Information Science Society
2014, **25**(6), 1241–1251

http://dx.doi.org/10.7465/jkdi.2014.25.6.1241
한국데이터정보과학회지

국내 28개 국공립대학교의 연구성과에 대한 평가[†]

정동빈[1]

[1]강릉원주대학교 정보통계학과

접수 2014년 7월 21일, 수정 2014년 8월 7일, 게재확정 2014년 9월 5일

요 약

국내 28개 국공립대학교를 대상으로 대학정보공시제에 근거하여 제시된 교육지표 중 연구성과에 관련된 4개의 양적평가 우량성속성 (1인당 연구비 수혜금액, 1인당 연구재단등재지 (KCI) 실적, 1인당 SCI/SCOUP급 실적, 1인당 저역서 실적)에 대해, 유사한 속성을 지닌 대학들을 유사성있는 군집끼리 분류 및 세분화하고 다차원 공간상에 시각적으로 배치시켰다. 이를 통해서 국내 28개 국공립대학은 4개 군집으로 유형화가 가능하며, 각 군집간 또는 군집내에 속한 대학들 간의 유형 및 특성을 평가할 수 있다.

주요용어: 군집분석, 다차원척도법, 연구성과.

1. 서론

세계화로 국가간 경쟁이 심화되고 저출산 및 학령인구의 급격한 감소로 대학순위 상승을 위한 노력 및 연구 경쟁력 강화가 요구되고 있다. 이에 대해 대학에서는 대학정보공시제, 자체평가, 대학평가인증제 등 평가시스템을 통해 대학의 질 관리가 이루어지고 있다. 특히 대학교수의 연구력 향상을 통해 대학 경쟁력을 높이는 일은 더 이상 선택사항이 아닌 생존력의 문제로 대두되고 있다.

우리나라 2012년 총연구개발비는 전년대비 5조 5,597억 원 (11.1%) 증가한 55조 4,501억 원으로 처음으로 50조 원을 넘어섰으며, GNP 대비 연구개발비 비중은 4.36%에 해당된다. 반면에 2012년 교육부 연구개발사업의 총 연구비는 전년대비 1.5% 감소한 5,475억 원으로 기초연구는 60.4%, 학술연구는 39.1%, 기반조성은 0.5% 차지한다. 기초연구의 수준을 알아보는 2012년 NSI (National Science Indicators) DB에 수록된 우리나라 SCI 논문게재수는 49,174편으로 전년에 비해 9.9% 증가하였으며, 전세계 SCI 논문의 3.61% (세계 11위)를 점유하고 있다. 또한 2012년도 기준 최근 5년간 (2008-2012) 발표한 논문 1편당 평균 피인용도 (타 논문에 의해 인용된 횟수로 자기인용도 포함)는 4.08회로써 세계 31위 (세계평균 5.04회)이다 (Ministry of Education, 2013).

올 1월 교육부는 새로운 지식 창출과 창의적 연구자 양성을 목표로 「2014년 교육부 학술·연구지원사업 종합계획」을 발표하였다. 이 종합계획은 교육부가 추진하는 총 6천억 원 규모의 인문사회 및 이공분야 학술·연구지원, 학술기반 확충 및 성과확산 등의 사업을 총망라하여 수립되는 첫 종합계획이라는 점에서 의의가 크다. 또한 인문사회 기초연구에 2,237억 원, 과학기술 기초연구에 3,314억 원, 성과확산 및 대중화 등에 206억 원 등 2014년 학술·연구에 총 6,039억 원을 지원한다고 밝혔다. 특히 인문학 대중화 사업규모가 13년 29억 원에서 14년 60억 원으로 대폭 증가하였고, 학제간 융합연구도 13년 34억 원

[†] 이 논문은 강릉원주대학교 자연과학연구소의 지원을 받았음.
[1] (210-702) 강원도 강릉시 죽헌길 7, 강릉원주대학교 정보통계학과, 교수. dj@gwnu.ac.kr

에서 14년 50억 원으로 증가하였다. 2014년 학술·연구지원사업은 “새로운 지식 창출”과 “창의적 연구자 양성”을 통해 상상력과 창의성에 기반을 둔 창조경제 실현의 원천을 창출하는 동시에, 사회적 역할을 강화하여 국민의 삶의 질제고, 사회문제 해결에 기여하는 방향으로 추진될 예정이다. 특히 “인문정신문화 진흥”, “지역대학 육성” 등 국정과제 실현과 “공과대학 혁신”과 같은 현안 해결에 도움이 될 수 있도록 주요 과제를 선정하였다 (Ministry of Education, 2014).

본 연구에서는 2012년도 전임교원에 대한 수많은 연구성과지표 중에서 대표적으로 사용되는 4개의 우량적 속성인 1인당 연구비 수혜금액, 1인당 연구재단등재지 (KCI)실적, 1인당 SCI/SCOUP급 실적, 1인당 저역서 실적을 중심으로 군집분석 (cluster analysis)을 적용시켜 유사성이 강한 대학끼리 분류 및 유형을 파악하고, 또한 다차원척도법 (multidimensional scaling)을 이용하여 4개의 성과지표속성에 대해 각 대학을 포지셔닝맵 공간상에 시각적으로 표현한다. 이를 통해 28개 국공립대학 간 상대적인 차별적인 우위와 그 특성을 파악하여 향후 세분화된 군집에 속한 대학의 역량을 강화 및 개선 방향을 제시하는데 의의가 있을 것이다.

Jeong (2013)은 국내 14개 국립대학교를 대상으로 교육과학기술부에서 제시한 주요 8개의 교육지표를 사용하여 다차원척도법을 분석한 반면에, 본 연구는 28개 국내 국공립대학교의 연구성과 지표에 대해 군집분석과 다차원척도법을 동시에 적용하였다. Lee와 Kim (2013)은 비계층적 군집분석을 사용하여 계획된 소비습관에 따른 저축률의 차이 및 소비습관에 대한 교육의 저축률에 대한 영향을 알아보았다. Jeong (2012)는 대학조직을 대상으로 미국의 말콤볼드리지 (MBNQA) 국가품질상 교육분야의 기준을 적용하여 리더십, 전략기획, 고객중시, 인적자원중시, 프로세스관리, 측정분석 및 지식경영과 교육성과간의 인과관계 및 매개효과를 검증하였다.

2절에서는 연구에서 적용할 군집분석과 다차원척도법에 대한 정의와 이론적 개념을 정리하며, 또한 이 다변량분석들에 적용할 대학정보공시제의 연구성과부문에 관련된 4개 평가지표 속성을 제시한다. 3절에서는 연구성과부문에 연관된 평가항목변수들을 중심으로 두 가지 다변량분석을 이용하여 세분화 (segmentation) 및 시각적인 포지셔닝맵을 표현함으로써 집단 내적으로는 균일하고 외적으로는 타 집단과 이질적인 대학들의 특성과 향후 각 군집에 속한 대학들의 개선방향을 제시한다. 마지막으로 4절에서는 연구결과와 결론을 정리한다.

2. 군집분석, 다차원척도법 및 연구성과 우량적 평가항목

교육부가 발표한 대학평가활용 구조조정 방안의 일환으로 대학정보공시에서 발췌한 연구성과부문 중 우량적 속성이라 판단되는 4가지 평가항목 속성만을 고려하여 군집분석과 다차원척도법을 실시하고자 한다.

2.1. 군집분석

개인 또는 개체 중에서 유사한 것들을 몇 개의 집단으로 그룹화하여, 각 집단의 성격을 파악함으로써 데이터 전체의 구조에 대한 이해를 돕고자 하는 탐색적 다변량 분석방법이다. 다음과 같이 유사성거리를 측정하는 방법이 있다.

대상 i의 관측값을 $x_i = (x_{i1}, x_{i2}, \cdots, x_{ip})', i = 1, 2, \cdots, n$이라고 할 때, 개체간 비유사성을 정의하는 제곱 유클리드거리는 $\sum_{j=1}^{p}(x_{ij} - x_{kj})^2$이며, Euclidean 거리는 $\sqrt{\sum_{j=1}^{p}(x_{ij} - x_{kj})^2}$, Manhattan (또는 Block) 거리는 $\sum_{j=1}^{p}|x_{ij} - x_{kj}|$, Chebyshev 거리는 $\max_j |x_{ij} - x_{kj}|$, Minkowski 거리는 $\left(\sum_{j=1}^{p}|x_{ij} - x_{kj}|^r\right)^{1/r}$ 등이다 (Yang, 2013).

계층적 군집방법은 군집이 다른 군집에 포함되나 군집간의 중복이 허용되지 않고 계보형식이 나뭇

가지와 같은 구조를 취하며, 군집 내의 유사성은 가깝게, 군집간의 유사성을 멀게 하는 개념에서 다음과 같이 대표적으로 사용되는 대상들을 군집화하는 방법이 있다 (Mercury Learning and Information, 2014).

1) 최단연결법 (single linkage method) : 두 군집 간의 거리를 각 군집에 속하는 임의의 두 개체들 사이의 거리 중 최단거리로 정의하여 가장 유사성이 큰 군집을 묶어 나가는 방법이다.
2) 최장연결법 (complete linkage method) : 기존 군집에 포함된 모든 대상에 대해 특정 거리 내에 포함되면 동일한 군집에 포함시키는 벙법으로, 군집간의 거리를 각 군집에 속하는 임의의 두 개체들 사이의 거리 중 최장거리로 정의하여 가장 유사성이 큰 군집을 묶어 진행한다.
3) 중심연결법 (centroid linkage method): 군집간의 거리는 각 군집의 중심들 사이의 거리를 사용하며, 새로운 대상을 포함시킬 때 군집의 중심과의 거리가 가장 가까운 대상을 포함시키는 방법이다.
4) 평균연결법 (between- or within-groups linkage) : 각 군집에서 하나씩의 개체를 선택해 연결한 모든 조합가능한 거리의 평균을 정의한 후, 두 군집사이의 거리를 각 군집에 속하는 모든 개체들의 평균거리로 정의하여 가장 유사성이 큰 군집을 묶어 나가는 방법이다.

반면에 K-평균 군집분석법은 사전에 결정된 군집 수 (K)에 근거하여 각 관측값을 군집의 중심들 중에서 가장 가까운 군집을 할당하는 방법이며, 여기에서 군집 수는 계층적 방법 또는 여러 개의 K를 사용하여 군집간 평균거리나 군집 내 개체 평균거리를 활용하여 최적의 군집수를 결정한다 (Borg와 Groenen, 2005; Savaresi와 Boley, 2004).

2.2. 다차원척도법

개체들을 군집화하는 군집분석과는 달리, 다차원척도법은 n개 개체를 개체간 유사성 또는 비유사성을 근거로 하여 2차원의 가시적공간에 위치시키는 시각화 분석기법이다. 다시 말해서, 다차원척도법은 개체들의 상대적 거리에 근거하여 2차원평면 또는 다차원공간상에 개체들을 위치화 시켜주는 분석법으로, 물리적 거리뿐만 아니라 개체들의 유사성과 선호도를 나타내는 자료를 근거로 응답자들이 지각하고 있는 심리적 거리를 계산하여 공간상에 도식화시킬 수 있다 (Jeong, 2013).

다차원척도법에서 사용되는 비유사성 (dissimilarity) 측도는 두 개체가 얼마나 상이한가를 나타내는 것으로, 숫자가 클수록 상이함이 강하며, 대표적으로 유클리드 거리, Chebychev 거리, Minkowski 거리 등을 들 수 있다.

개체간 비유사성 거리 (d_{ij})를 나타내는 일반화된 유클리드 거리는 각 개체에 대한 다변량 관측 $(x_{i1}, x_{i2}, \cdots, x_{ip})$, $i, j = 1, \cdots, n$으로부터 개체 i와 j사이의 거리로 아래와 같이 측정된다.

$$d(x_i, x_j) = \left\{ \sum_{k=1}^{p} (x_{ik} - x_{jk})^2 \right\}^{1/2}, \ i, \ j = 1, \cdots, n \tag{2.1}$$

이러한 비유사성거리의 개념을 착안하여, 다차원척도법에서 부적합도 (badness-of-fit)를 나타내는 측도로 Kruskal은 스트레스 (STRESS; standardized residual sum of squares)를 다음과 같이 제안하였다 (Kruskal, 1964; Kruskal과 Wish, 1978).

$$STRESS = \sqrt{\sum_{i<j} \left\{ d_{ij} - \widehat{d}_{ij} \right\}^2 / \sum_{i<j} d_{ij}^2} \tag{2.2}$$

여기에서 $\widehat{d}_{ij}$는 개체 i와 j간 유사성의 값 (즉, 실제거리)인 d_{ij}의 적합값 (또는 추정된 거리)이다. STRESS값은 평가시에 고려하는 기준이 되는 차원을 결정하거나 출력자료의 정확도를 평가하기 위

해 사용되는 도구로써 개체들간의 실제거리와 적합된 상대적 거리간의 오차의 정도를 나타내준다. 만일 입력자료의 일관성이 높고 정확도가 높을수록 두 거리가 일치하게 되므로 0에 가깝게 될 것이다. STRESS 값이 작으면 작을수록 좋으며, 일반적으로 5% 이내이면 대체로 좋은 것으로 20% 이상이면 나쁘다고 판단한다.

2.3. 연구성과 우량적 평가항목

다음은 국공립대학교 연구성과 우량적 평가항목과 각 항목에 관련된 구체적인 측정방법을 나열한 것이다.

1) 1인당 연구비 수혜금액 : 2013년 4월 1일 기준 소속대학 전임교원이 연구책임자로서 2012년도에 수행한 과제의 계약 (협약) 건수에 해당되며, 교내와 교외 연구비는 다음과 같이 측정한다.
   - (교내연구비+교내대응자금)/전임교원수 (단위: 천원)
   - (교외연구비+교외대응자금)/전임교원수 (단위: 천원)

   또한 연구책임자 소속대학 및 학과 기준으로, 두 기관 이상에서 지원받은 연구의 경우 지원규모가 더 큰 기관 기준으로 계산하였다.
2) 1인당 SCI/SCOUP 실적 : 2013년 4월 1일 기준 소속대학 전임교원이 2012년 1월 1일 ∼2012년 12월 31일 기간에 국제전문학술지 (SCI, SCIE, SSCI, A&HCI, SCOPUS)에 게재한 논문수와 국내에서 발행되는 SCI급/SCOPUS 학술지 게재 논문을 포함하여 아래와 같이 계산한다.
   - (SCI급 논문수+SCOPUS 논문수)/ 전임교원 수
3) 1인당 한국연구재단 KCI(후보)실적 :2013년 4월 1일 기준 소속대학 전임교원이 2012년 1월 1일 ∼2012년 12월 31일 기간에 한국연구재단의 등재 또는 등재후보학술지(국내에서 발행되는 국제학술지는 제외)에 게재한 논문 수를 의미한다.
4) 1인당 저역서 실적2013년 학과별 전임교원의 전년도 (2012년) 저·역서 실적에 한하며, ISBN이 수록된 출판물만 인정하며 아래와 같이 산출한다.
   - (저역서 수/전임교원 수)

그리고, 4개의 항목변수는 모두 전임교원을 대상으로 작성되었으며, 여기에서 전임교원은 2013년 4월 1일 현재 재직 중인 전임교원 (휴직교원 포함)을 말하며, 교수, 부교수, 조교수로 분류, 외국인 전임교원 포함되며, 총장은 전임교원에 포함되지 않는다. 연구실적 및 저역서 실적은 2012년 1월 1일 ∼2012년 12월 31일 기간 동안의 등재 학술지 기준으로 게재된 논문 수를 산정한다. 참고로 논문게재 실적 산정 기준은 단독저자인 경우 실적은 1건이며, 논문저자 수가 2명 이상일 경우 주저자 실적 $= 2/(n+2)$, 교신저자 실적 $= 2/(n+2)$, 공동저자 실적 $= 1/(n+2)$이며, 논문저자 수가 15명 이상일 경우 $n = 15$로 처리한다. 반면에 저·역서 수 산정 기준은 단독저술 실적은 1건이며, 공동저술의 경우 $1/n$, 여기에서 n은 저술에 참여한 인원수이다.

연구성과에 관련된 각각의 4개 평가항목 속성들의 관측척도가 다르기 때문에 평균이 0, 표준편차가 1인 Z-스코어로 표준화시켰다.

3. 실증분석결과

3.1. 상관분석과 상이성 거리행렬

아래에 주어진 Table 3.1은 각 4개의 연구성과 지표들간의 Pearson 상관계수와 유의성 검정결과를 나타낸 것이다. 특히, 유의한 쌍 중 (1인당 연구비 수혜금액, 1인당 SCI/SCOUP급 실적)과 (1인당 연

구등재지 실적, 1인당 저역서 실적)은 양의 상관관계를 나타내고 있음을 확인할 수 있다 (각각 p값은 <0.001과 =0.017).

추가로 비모수적인 접근방법으로 고려한 두 개의 상관계수인 Spearman의 ρ와 Kendall의 τ의 분석 결과를 살펴보면, (1인당 연구비 수혜금액, 1인당 SCI/SCOUP급 실적)만이 유의하다 (p값<0.01). 즉, 이 두 개의 평가지표의 상관계수가 0이 아니라는 대립가설을 매우 강하게 지지하는 것을 알 수 있다 (Table 3.2, Table 3.3 참고).

Table 3.4에 주어진 값은 8개의 교육지표에 근거하여 대응하는 대학 간에 상이성 정도를 거리로 계산하여 행렬로 표현한 것으로, 값이 커질수록 두 대학 간의 상이성이 크다는 것을 의미한다.

Table 3.1 Parametric correlation analysis

variables to be evaluated		research funds *p.c.	SCI/SCOUP achievements *p.c.	KCI achievements *p.c.	book (or translated) achievements *p.c.
research funds*p.c.	corrl. coeff.	1	-.090	.659	-.147
	p-value		.649	.000	.454
SCI/SCOUP achievements *p.c.	corrl. coeff.		1	-.169	.446
	p-value			.389	.017
KCI achievements*p.c.	corrl. coeff.			1	-.198
	p-value				.312
book (or translated)achievements *p.c.	corrl. coeff.				1
	p-value				

*p.c. stands for per capita

Table 3.2 Nonparametric correlation analysis (Spearman's ρ)

variables to be evaluated		research funds *p.c.	SCI/SCOUP achievements *p.c.	KCI achievements *p.c.	book (or translated) achievements *p.c.
research funds*p.c.	corrl. coeff.p-value	1	.012	.696	-.096
			.952	.000	.626
SCI/SCOUP achievements *p.c.	corrl. coeff.p-value		1	-.069	.303
				.727	.117
KCI achievements *p.c.	corrl. coeff.p-value			1	-.211
					.282
book (or translated) achievements *p.c.	corrl. coeff.p-value				1

*p.c. stands for per capita

Table 3.3 Nonparametric correlation analysis (Kendall' τ)

variables to be evaluated		research funds *p.c.	SCI/SCOUP achievements *p.c.	KCI achievements *p.c.	book (ortranslated) achievements *p.c.
research funds *p.c.	corrl. coeff.p-value	1	.016	.582	-.058
			.906	.000	.664
SCI/SCOUP achievements *p.c.	corrl. coeff.p-value		1	-.053	.175
				.693	.192
KCI achievements *p.c.	corrl. coeff.p-value			1	-.138
					.304
book (or translated) achievements *p.c.	corrl. coeff.p-value				1

*p.c. stands for per capita

우선 한국교원대에 해당하는 모든 값이 전반적으로 크기 때문에 다른 대학과 상이하게 구분된다고 할 수 있다. 몇몇의 지역거점대학 (전북대, 경북대, 부산대, 충남대, 충북대)끼리 상대적으로 작은 값을 갖기 때문에 상대적으로 동일한 성격을 가진 집단으로 간주할 수 있으며, 특히 한국교원대는 이 지역거점대학들과 가장 거리가 멀리 떨어져 있다. 지역거점대학 중 강원대는 대체로 지역거점대학이 아닌 다른 대학과, 서울과기대는 반대로 대부분의 지역거점대학과 거리가 가까운 특징을 발견할 수 있다. 한 예로, 강릉원주대는 한국교원대와 대부분의 지역거점대학과 멀리 떨어져 있으며, 창원대와 한경대, 경남과기대, 목포해양대, 군산대 순으로 가까운 거리에 있어 이 대학들과는 상대적인 상이성이 작다고 판단된다.

Table 3.4 Dissimilarity distance matrix between national universities

	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	26	27	28
1	0.0																											
2	1.9	.0																										
3	1.2	1.9	.0																									
4	3.4	1.7	3.5	.0																								
5	2.2	.9	2.7	1.4	.0																							
6	2.3]	2.3	2.7	3.2	2.4	.0																						
7	1.7	1.5	1.0	3.0	2.4	2.2	.0																					
8	2.5	1.6	2.7	2.0	1.6	3.5	2.6	.0																				
9	2.7	2.3	3.1	2.6	2.5	2.8	2.6	2.0	.0																			
10	1.4	1.9	1.1	3.4	2.6	3.4	1.7	2.1	3.0	.0																		
11	3.5	2.6	3.5	2.8	2.7	1.6	2.7	3.9	3.4	4.2	.0																	
12	4.1	2.4	4.0	1.4	2.5	3.2	3.1	3.1	3.0	4.2	2.3	.0																
13	3.8	2.2	3.7	1.6	2.0	4.1	3.4	2.5	3.9	3.5	3.7	2.4	.0															
14	3.6	2.4	3.3	2.2	2.8	4.2	2.9	1.8	2.4	2.9	4.1	2.6	2.7	.0														
15	1.9	1.6	1.8	2.8	2.1	3.6	2.1	1.3	2.8	0.9	4.2	3.7	2.8	2.4	.0													
16	1.8	2.1	2.3	3.2	2.2	.6	2.1	3.2	2.7	3.0	2.1	3.5	4.0	4.1	3.2	.0												
17	2.5	1.6	3.2	1.7	.8	2.4	2.9	2.0	2.4	3.1	2.8	2.7	2.6	3.2	2.7	2.2	.0											
18	3.7	2.2	4.0	.8	1.7	3.3	3.4	2.2	2.4	3.9	3.0	1.7	2.2	2.6	3.2	3.3	1.6	.0										
19	2.0	.7	2.2	1.7	1.0	1.7	1.6	1.9	1.9	2.4	2.1	2.2	2.6	2.6	2.2	1.6	1.3	2.0	.0									
20	0.9	1.0	1.1	2.6	1.6	2.0	1.1	2.0	2.3	1.4	2.9	3.3	3.1	2.9	1.6	1.6	2.1	3.0	1.2	.0								
21	3.2	1.5	3.2	1.1	1.5	3.6	2.8	1.3	2.7	2.8	3.6	2.2	1.3	1.7	2.0	3.5	2.0	1.6	1.9	2.4	.0							
22	3.6	1.8	3.5	.7	1.8	3.5	2.9	1.9	2.6	3.3	3.1	1.4	1.6	1.7	2.7	3.5	2.2	1.2	1.9	2.8	.90	.0						
23	1.0	1.6	1.3	3.1	2.1	3.0	1.8	1.8	2.6	.7	3.9	4.0	3.3	3.0	.9	2.5	2.5	3.5	2.0	1.0	2.6	3.2	.0					
24	4.7	4.6	4.0	5.5	5.3	3.7	3.4	5.9	5.2	5.0	3.4	4.7	5.9	5.6	5.5	4.0	5.6	5.9	4.4	4.4	5.8	5.5	5.2	.0				
25	1.7	1.0	1.2	2.5	1.8	2.8	1.2	1.9	2.8	1.3	3.2	3.0	2.5	2.5	1.2	2.5	2.5	3.1	1.6	1.1	2.0	2.4	1.3	4.4	.0			
26	1.9	3.0	1.3	4.5	3.6	2.8	1.7	4.0	3.9	2.3	3.7	4.7	4.8	4.4	3.1	2.5	4.1	5.0	3.0	2.1	4.4	4.6	2.5	3.3	2.4	.0		
27	3.0	2.4	3.1	2.8	2.7	1.6	2.2	3.3	2.0	3.6	1.6	2.4	4.0	3.3	3.7	2.0	2.7	2.8	1.8	2.5	3.3	2.9	3.4	3.5	2.9	3.4	.0	
28	1.9	.3	1.9	1.6	1.1	2.2	1.5	1.5	2.0	2.0	2.5	2.3	2.3	2.2	1.7	2.0	1.6	2.1	0.6	1.0	1.6	1.7	1.7	4.5	1.1	3.0	2.2	.0

The numerical values above stand for the following:
1: Gangneung-Wonju National Univ. 2: Kangwon National Univ.
3: Gyungnam National Univ. of Science & Technology
4: Kyungpook National Univ. 5: Gyungsang National Univ. 6: Kongju National Univ.
7: Kunsan National Univ. 8: Kumho National Institute of Technology
9: Mokpo National Univ. 10: Mokpo National Maritime Univ.
11: Pukyong National Univ. 12: Pusan National Univ.
13: Seoul National Univ. of Science & Technology
14: Univ. of Seoul 15: Sunchon National Univ. 16: Andong National Univ.
17: Chonnam National Univ. 18: Chonbuk National Univ.
19: Jeju Mokpo National Univ. 20: Changwon National Univ.
21: Chungnam National Univ. 22: Chungbuk National Univ.
23: Hankyung National Univ. 24: Korea National Univ. of Education
25: Korea National Univ. of Transportation 26: Korea National Sport Univ.

3.2. 군집분석

3.2.1. 계층적 군집방법

계층적 군집방법을 사용하여 병합되는 과정에서 거리가 상대적으로 큰 변화가 보이는 즉, 군집화 일정표에서 계수가 급속하게 증가하는 시점에서 적절한 군집의 수를 결정할 수 있다. 아래 주어진 Table 3.5를 통하여 이러한 변화가 감지되는 시점이 4군데 (즉, 단계 21에서 22, 단계 22에서 23, 단계 24에서 25, 단계 26에서 27)에서 발견할 수 있기 때문에 4개 군집으로 유형화할 수 있다.

Table 3.5 Agglomeration schedule

stage	cluster combined		coefficients	stage cluster first appears		nextstage
	cluster 1	cluster 2		cluster 1	cluster 2	
1	2	28	0.099	0	0	4
2	6	16	0.346	0	0	19
3	4	22	0.428	0	0	10
4	2	19	0.451	1	0	13
5	10	23	0.527	0	0	8
6	5	17	0.639	0	0	13
7	1	20	0.763	0	0	14
8	10	15	0.871	5	0	15
9	3	7	0.976	0	0	11
10	4	21	0.985	3	0	12
11	3	25	1.411	9	0	14
12	4	18	1.492	10	0	17
13	2	5	1.681	4	6	18
14	1	3	1.834	7	11	15
15	1	10	2.271	14	8	22
16	11	27	2.668	0	0	19
17	4	13	2.950	12	0	20
18	2	8	2.993	13	0	21
19	6	11	3.367	2	16	26
20	4	12	3.542	17	0	21
21	2	4	4.196	18	20	24
22	1	26	4.985	15	0	25
23	9	14	5.792	0	0	24
24	2	9	6.238	21	23	25
25	1	2	8.551	22	24	26
26	1	6	9.078	25	19	27
27	1	24	23.462	26	0	0

3.2.2. K-평균 군집방법

1) 소속군집과 최종군집중심

집단 4에 속한 대학들 (경북대, 부산대, 서울과기대, 전북대, 충남대, 충북대)은 전남대와 강원대를 제외한 모든 지역거점대학들이 포함된 것을 발견할 수 있고, 각 집단이 고루 나눠 이뤄졌음을 확인할 수 있다. 최종적으로 각 4개의 군집에 대한 4개의 평가항목변수의 평균값을 비교하면 다음과 같은 결과를 얻을 수 있다 (Table 3.6 참고).

다른 집단에 비해 집단 4는 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'의 값이 월등히 큰 반면에 (각각 약 1.04, 1.40), '1인당 저역서 실적'은 저조한 편이다. 이는 수혜 받은 연구비의 지원으로 외국유명저널에 기여했음을 시사한다. 또한 '1인당 연구재단등재지 실적' 평가항목도 우수한 위치를 차지하고 있다.

집단 4와는 반대로 집단 1에 속한 대학들 (공주대, 부경대, 안동대, 한국교원대, 한국해양대)은 나머지 두 개의 평가항목인 '1인당 연구재단등재지 실적'과'1인당 저역서 실적'에 있어 다른 집단에 비해 우월한 위치를 차지하고 있음을 확인할 수 있다 (각각 약 1.24, 1.64).

집단 3에 속한 대학들 (강릉원주대, 경남과기대, 군산대, 목포해양대, 순천대, 창원대, 한경대, 한국교통대, 한국체육대)은 모든 평가항목에서 음의 값을 가지며, 특히 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'은 다른 집단에 비해 열등하다고 판단된다. 이러한 현상은 대부분 후발국립대학으로 이뤄진 집단 3은 지역거점대학으로 이뤄진 집단 4와 극단적인 대조를 보여준다.

집단 2에 속한 대학들 (강원대, 경상대, 금호공과대, 목포대, 서울시립대, 전남대, 제주대, 한밭대)은 4개의 평가항목 중 집단 1에 속한 대학들과 상반되는 결과를 발견할 수 있다. 즉, 집단 1에 속한 대학들은 '1인당 연구재단등재지 실적'과 '1인당 저역서 실적'에 있어 가장 큰 양의 값을 갖는 반면에, 집단 2에 속한 대학들은 이 두 항목에 대해 음의 값을 가져 하위의 위치를 차지한다. 반면에 나머지 두 개의 평가항목인 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'의 경우, 집단 4와 같이 다소 큰 양의 값을 가지며, 이는 양호한 점수로 평가할 수 있다.

Table 3.6 Final clustering centers

variables to be evaluated	cluster			
	1	2	3	4
research funds *p.c.	-.1819	.4982	-1.0342	1.0387
SCI/SCOUP achievements *p.c.	1.2427	-.4937	-.2773	.0385
KCI achievements *p.c.	-.3738	.1603	-.8657	1.3963
book (or translated) achievements *p.c.	1.6449	-.1869	-.3542	-.5903

*p.c. stands for per capita

2) 군집중심간 거리와 분산분석

각 집단간 평균차를 살펴보면, (집단 1과 집단 4)와 (집단 3과 집단 4)가 크게 차이가 있고, 상대적으로 (집단 2와 집단 4)는 다른 집단들에 비해서 상대적으로 가장 가까운 집단임을 알 수 있다. 전반적으로 양의 점수가 크고 많은 집단 4는 다른 집단보다 거리가 훨씬 멀리 떨어져 있고, 반면에 고려한 대학의 약 32%를 차지하며 점수가 가장 저조한 집단 3은 다른 집단에 비해 집단 2와 거리가 가깝게 놓여있음을 알 수 있다 (Table 3.7 참고).

4개 평가항목변수에 대해 4개의 집단별 평균차 검증을 한 것으로 4개 평가항목 변수 모두 4개 집단 분류에 유의한 변수임을 확인할 수 있다 (Table 3.8 참고).

특히 '1인당 연구비 수혜금액','1인당 SCI/SCOUP급 실적'과 '1인당 저역서 실적'은 큰 F-값을 가지므로, 4개 집단에 대한 차이가 크게 식별되고 있다.

Table 3.7 Distances between final cluster centers

cluster	1	2	3	4
1		2.668	2.697	3.327
2			1.864	1.505
3				3.093
4				

Table 3.8 Analysis of variance

variables to be evaluated	cluster		error		F	Sig.
	MS	df	MS	df		
research funds p.c.	6.083	3	.365	24	16.686	.000
SCI/SCOUP achievements p.c.	3.457	3	.693	24	4.990	.008
KCI achievements p.c.	6.449	3	.319	24	20.227	.000
book (or translated) achievements p.c.	5.676	3	.416	24	13.659	.000

3.3. 다차원척도법

Figure 3.1은 대학 간 근접성 변환값과 2차원 다차원척도법에 타점된 대학 간의 거리를 플롯한 것으로, 대응점들이 대략 일직선상에 놓여 있는 것을 알 수 있으며, 이는 고려한 다차원척도 맵이 대학교들을 어느 정도 정확히 표현하고 있음을 확인할 수 있다.

Figure 3.2는 2차원 다차원척도에서 각 28개의 대학들이 위치할 좌표를 계산한 후 2차원 산점도로 표현한 것이다. 다차원 공간상의 대학교별 좌표를 구한 후 계산된 STRESS는 '0.15'로 Kruskal 기준에 따라 채택할 수 있다고 판정할 수 있다.각 속성들은 원점에서 화살표 방향으로 멀어질수록 각 속성의 성향이 강하게 가지고 있다. 즉, 원점을 중심으로 가로축 (제1축)의 좌측으로 갈수록 4개의 연구성과 평가속성 중 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'이 강한 양의 값으로, 반면에 원점을 중심으로 세로축 (제2축)의 위로 갈수록 '1인당 연구재단등재지 실적'과 '1인당 저역서 실적'이 강한 음의 값을 갖는 성향이 나타낸다. 또한 연구성과 평가항목변수와 대학교간의 관계는 선에 수직으로 내렸을 때 가깝게 접근하는 대상 대학교일수록 연구성과 평가속성의 성격을 많이 가진 대학교로 판정할 수 있을 것이다.

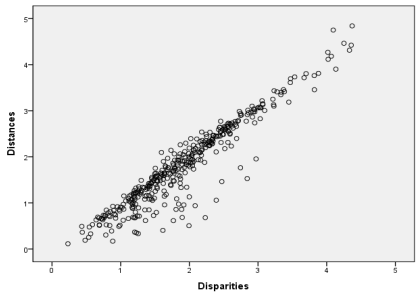


Figure 3.1 Scatter plot of linear fit

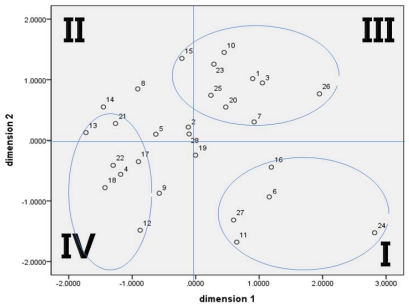
Figure 3.2의 4개의 사분면에 각각 표시된 I, II, III, IV는 각각 3.2절에서 분류한 집단 1, 집단 2, 집단 3, 집단 4을 의미하며, 3개의 타원은 그 중 3개 집단에 묶어 시각적으로 표현한 것이다.

이러한 맥락에서 보면, 우선 집단 4 (IV)에 속한 대학교 중 전북대와 집단 2 (II)에 속한 대학교 중 서

울과기대, 서울시립대 등이 연구성과 평가속성 중 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'이 매우 유사하며 다른 대학들에 비해 우위에 있음을 알 수 있다. 반면에 집단 1 (I)에 속하는 대학 중 한국체육대와 한국교원대는 두개의 평가속성에 대해 가장 열등한 위치에 있음을 알 수 있다.

반면에 집단 1 (I)에 속한 대학 중 목포해양대와 순천대는 나머지 대학에 비해서 연구성과 평가 속성 중 '1인당 연구재단등재지 실적'과 '1인당 저역서 실적'은 낮은 점수를 가지며, 집단 1 (I)에 속한 대학 중 부경대와 한국교원대, 집단 4 (IV)에 속한 대학 중 부산대는 다른 대학들에 비해 월등히 높은 점수를 갖는다는 것을 알 수 있다.

4개의 평가항목변수에 대해 가장 열세한 위치에 있는 집단 3 (III)에 속한 대학 중 순천대는 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'에, 군산대는 '1인당 연구재단등재지 실적'과'1인당 저역서 실적'에 보다 주력을 하면 각각 집단 2 (II)와 집단 1 (I)의 방향으로 전환할 수 있을 것이다. 특히 집단 2에 속한 전남대, 금호공과대는 '1인당 연구비 수혜금액'과'1인당 SCI/SCOUP급 실적'에 더 많은 점수를 가지면 전체적으로 총점이 높은 집단 4 (IV)로 이동할 수 있을 것이다. 또한 집단 1 (I)에 속한 한국해양대와 부경대는 향후 '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'에 향상을 보인다면 전반적으로 연구성과의 발전가능성이 크다고 할 것이다.



* Note that numerical values above are same as those of Table 3.4

Figure 3.2 Derived stimulus configuration

4. 결론

본 연구는 대학정보공시제에 근거하여 교육부가 실시하려는 대학구조개혁에 사용하는 주요 교육지표 중 특히 연구성과부문을 기준으로 두 종류의 군집분석 적용시켜, 각기 국내 28개의 국공립대학교를 유사한 속성이 있는 4개의 군집으로 분류하여 유형화가 가능함을 확인하였다. 그리고, 다차원척도법을 이용하여 유형화된 군집들의 위치와 또한 각 군집에 속한 28개 국공립대학교의 포지셔닝을 실시하였다. 상이성 거리행렬을 사용하여 각 대학 간의 유사성과 상이성을 분석할 수 있었으며, 각 대학에 대해 연구성과부문 중 어느 평가속성에 더 주안점을 두고 정책을 세워야 할지를 예상할 수 있었다.

연구성과부문을 대변하는 4개의 우량적 속성 (1인당 연구비 수혜금액, 1인당 연구재단등재지 실적, 1인당 SCI/SCOUP급 실적, 1인당 저역서 실적)에 근거하여, 집단 4에 속한 대학 (경북대,부산대,서울과기대,전북대,충남대,충북대)은'1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'에, 집단 1에 속한 대학 (공주대, 부경대, 안동대, 한국교원대, 한국해양대)은 '1인당 연구재단등재지 실적'과 '1인당 저역서 실적'에 우등한 우위에 있다. 반면에 집단 3에 속한 대학 (강릉원주대, 경남과기대, 군산대, 목포해양대, 순천대, 창원대, 한경대, 한국교통대, 한국체육대)은 전반적인 평가속성에 대해 열약한 위치

에 있는 것으로 판단된다. 대부분의 연구비를 수혜한 대상자에게는 SCI/SCOUP급의 논문수록을 요구되는 상관성이 이를 반영한다.

집단 2에 속한 대학 (강원대, 경상대, 금호공과대, 목포대, 서울시립대, 전남대, 제주대, 한밭대)은 집단 4와는 유사하고, 집단 1에 속한 대학들과 상반되는 결과를 발견할 수 있다. 특히 2차원 공간에서 살펴본 포지셔닝을 평가하면 집단 2에 속한 대부분의 대학들은 대체로 집단 4와 근접해 있기 때문에, '1인당 연구비 수혜금액'과 '1인당 SCI/SCOUP급 실적'에 좀 더 노력을 기울인다면 집단 4로 합류하여 연구력에서 경쟁력을 향상시킬 수 있을 것으로 기대된다. 그러나, 집단 4에 속한 한국교원대와 집단 1에 속한 한국체육대는 이 두 가지 평가속성에 대해 집중적으로 노력해야 할 것이다.

반면에 '1인당 연구재단등재지 실적'과 '1인당 저역서 실적'에 대해서는 각각 집단 4 (IV)에 속한 부산대와 집단 1 (I)에 속한 부경대는 강세를 보이지만, 집단 3 (III)에 속한 목포해양대와 순천대는 열등한 위치에 있다고 평가할 수 있다.

본 논문에서는 각각 동일한 특성을 갖는 대학들끼리 집단화시켜 그 특성을 파악하였고, 대학의 연구력 향상을 위해서 우위에 있는 속성은 지속적으로 파악하고 유지시키는데 힘쓰고 열성에 속성은 보다 많은 관심을 기울임으로써 국가경쟁력에 이바지 할 것으로 믿어 의심치 않을 것이다.

본 연구에서는 양적지표인 4개의 우량적 속성을 근거로 하여 두 개의 다변량분석을 실시하였다. 논문성과 관련 계량지표로서 질적지표인 세계3대 저널 (NSP) 게재논문수, 1,000억 당 NSP 논문수, 평균 IF (impact factor), IF 상위 논문수, 주요과학저널 (Essential Science Indicators 활용한 22개 분야별 피인용횟수를 기준으로 한 Top5 저널선정) 논문수 등을 양적지표와 여타의 정성적 평가를 포함시켜, 국공립대학교 또는 국내 4년제 대학 간의 군집분석 및 포지셔닝을 차후 연구로 실시할 수 있을 것이다.

References

Aldenderfer, M. S. and Blashfield, R. K. (1985). *Cluster analysis*, Sage Publications, Los Angeles.

Borg, I. and Groenen, P. J. F. (2005). *Modern multidimensional scaling*, 2nd Ed., Springer-Verag, New York.

Jeong, D. B. (2013). A study on universities positioning using multidimensional scaling. *Korean Business Education Review*, **28**, 1-15.

Jeong, I. H. (2012). A study on the causal relationship of educational performance factors for university educational competence - Focus on Malcolm Baldrige national quality award model-. *Korean Business Education Review*, **27**, 207-239.

Kruskal. J. B. (1964). Major MDS based on a firm numerical analysis foundation. *Psychometrika*, **29**, 1-27.

Kruskal, J. B., and Wish. M. (1978). *Multidimensional scaling*, Sage Publications, Beverly Hills, CA.

Lee, G. H. and Kim, P. S. (2013). Effect of savings rate by adolescent consumption habit and business/economic education. *Korean Business Education Review*, **28**, 363-377.

Mercury Learning and Information. (2014). *Cluster analysis and data mining*, Mercury Learning & Information, Dulles, VA.

Ministry of Education (2013). *Main research development report*, National Research Federation of Korea, Sejong.

Ministry of Education (2014). *Reported data in Ministry of Education*, Department of Academic Promotion, Sejong.

Savaresi, S. M. and Boley, D. (2004). A comparative analysis on the bisecting k-means and PDDP clustering algorithm. *Intelligent Data Analysis*, **8**, 345-362.

Yang, B. H. (2013). *Understanding multivariate analysis*, Communication Books, Seoul.

Journal of the Korean Data & Information Science Society
2014, **25**(6), 1241–1251

http://dx.doi.org/10.7465/jkdi.2014.25.6.1241
한국데이터정보과학회지

Evaluation of research performances for 28 national universities

Dong Bin Jeong[1]

[1]Department of Information Statistics, Gangnueing-Wonju National University

Received 21 July 2014, revised 7 August 2014, accepted 5 September 2014

Abstract

Based on the 4 principal research-performance criteria in 28 national universities in Korea, both cluster analysis and multidimensional scaling are performed in this paper. We can classify and/or specialize the initially unknown groups into a group of relatively homogeneous universities and then create new groupings without any preconceived notion of what clusters may arise. Furthermore, the level of similarity of individual universities can be visualized on the multidimensional space so that each university is then assigned coordinates in each of the 2 dimensions. Both types and characteristics of each university can be relatively evaluated and be practically exploited for the policy of the university authority through these results.

Keywords: Cluster analysis, multidimensional scaling, research performances.

[1] Professor, Department of Information Statistics, Gangneung-Wonju National University, Gangneung 210-702, Korea. E-mail: dj@gwnu.ac.kr