

충청남도 시군의 성장유형 비교

- 행렬도(biplot)를 이용하여 -

최 종 후
(崔鐘厚)

고려대학교
정보통계학과
교수

양 우 성
(梁祐誠)

고려대학교
정보통계학과
석사과정

서 두 성
(徐斗星)

고려대학교
정보통계학과
석사과정

I. 서 론

II. 행렬도

III. 충청남도 시군의 성장유형 비교

1. 서 론

주지하는 바대로 이제 지방자치제의 실시에 따라 각 자치단체들은 주어진 여건 하에서 자주적이고도 창의적인 지역발전을 앞다투어 도모해 나가고 있다. 이는 충청남도 15개 시군의 경우에서도

예외가 아니다. 이러한 시대적 추이에 따라 최근 언론매체를 위시하여 여러 기관에서는 지역의 경제력 평가작업을 수행하고 있다. 이는 지역의 개발잠재력을 살펴봄으로써 지역개발의 방향과 전략을 모색하고 이에 따라 지역 특성을 살려 지역개발을 성공적으로 추진하기 위해서 당해 지역이 보유하고 있는 잠재력을 최대한으로 개발·육성하기 위한 기초적 연구라는 점에서 그 의미가 크다고 하겠다.

최근 발표된 김군수·강원철(1997)의 경기도 31개 시군에 대한 지역별 경제력 평가 역시 이러한 연구사례의 하나라고 하겠다. 김군수·강원철(1997)의 연구를 살펴보면 지역경제 성장잠재력의 근간을 이루는 대표적인 8개 부문 즉, 인구·주택 부문, 사회간접자본부문, 정보화부문, 그리고 농업, 광업, 제조업, 도소매업, 재정 등의 각 부문에 대하여 세부항목을 이용하여 경기도 31개 시군에 대한 지역별 경제력 평가를 시도하고 있다.¹⁾ 이들은 연구에서 인구·주택부문은 총인구, 생산연령인구,

1) 이들 연구의 공간적 범위는 경기도 31개 시군이며, 시간적 범위는 1989년~1994년의 5년간이다.

총세대수, 주택총수, 총인구 대비 생산연령 인구비율, 인구밀도, 주택 보급률 등의 7개 항목, 사회간접자본부문은 도로율, 운송(100인당 자동차수), 의료(100인당 병상수), 교육(100인당 고등학교 이상 학생수)의 4개 부문, 정보화부문은 100인당 교육이수자수, 전 산업에서의 서비스산업 종사자 비중, 100인당 전화대수, 조립금속 및 기계제조업의 제조업 비중의 4개 항목을 해당 부문의 성장패턴을 나타내는 자료로 이용하고 있다. 이러한 각각의 평가항목에 대한 표준화 값을 이용하여 경기도 31개 시군에 대한 항목별 순위를 매겨 지역별 비교 순위와 성장력을 평가하고 있다. 그러나 지역의 경제력은 총체적 시각으로 다루어져야 한다는 점에서 이러한 접근방식은 다소의 한계를 지니고 있다. 즉, 지역경제력을 판단하는 항목 상호간의 관계성이나 이들과 시군의 연계성을 총체적으로 살펴보기에는 미흡한 접근방식이라는 말이다.

본 연구는 이러한 점에 초점을 두고 행렬도(biplot)라는 통계분석기법을 활용하여 충남 15개 시군에 대한 성장유형을 비교해 보고자 한다. 따라서 본 연구는 지역의 경제력 평가에 대한 본격적인 연구라기 보다는 경제력 평가연구를 위한 새로운 방법론의 제기라는 점에 그 의의를 두고자 한다. 본 연구에서는 시군의 성장유형을 알아보기 위하여 인구·주택, 사회간접자본, 그리고 정보화 등의 3개 부문만을 선정하고, 이들 각 부문에 대하여 인구·주택부문은 총인구, 생산연령인구, 총세대수, 주택총수, 총인구 대비 생산연령 인구비율, 인구밀도, 주택 보급률 등의 7개 항목, 사회간접자본

부문은 도로 포장률, 운송(100인당 자동차수), 의료(100인당 병상수) 등의 3개 항목, 정보화부문은 100인당 우편물 접수 건수, 100인당 전화대수, 조립금속 및 기계제조업의 제조업 비중 등의 3개 항목을 해당 부문의 성장유형을 나타내는 항목으로 활용한다.

자료는「충청남도 통계연보」(충청남도, 1991·1994·1996), 「충남 주민등록인구 통계작성 결과보고서」(충청남도, 1996), 「사업체 기초통계 조사보고서」(충청남도, 1995)에서 인용하였다. 이 자료를 이용하여 인구·주택부문, 사회간접자본부문, 그리고 정보화부문의 총 13개 항목에 대하여 1990년부터 1995년까지 연평균 성장률을 산출한 결과가 <표 1>이다.

II. 행렬도

행렬도(biplot)는 Gabriel(1971)에 의하여 개발된 다변량자료에 대한 탐색적 자료분석기법의 하나이다. 자료행렬에 있어서 행을 나타내는 n개의 개체점들(여기에서는 시군) 사이에는 어떠한 기하적인 의미가 있는 지, 열을 나타내는 p개의 변수점들(여기에서는 성장유형을 알아보는 항목) 사이에는 어떠한 기하적인 의미가 있는 지, 그리고 이들 상호간에는 어떠한 의미가 있는지를 알아보기 위한 통계적 기법이다.

자료행렬의 행과 열을 동시에 플롯하는 대표적인 통계적 기법으로는 數量化 Ⅲ類(Theory of Quantification Ⅲ), 대응분석(Correspondence Analysis) 등이 있다. 數量化 Ⅲ類는 일본의 林

知己夫(하야시 치키오)에 의하여 개발된 통계분석 기법인데 이 기법은 다차원 자료가 外的 基準(하야시 박사는 從屬變數를 이렇게 부른다)을 포함하지 않는 경우의 수량화 기법이다. 개체의 여러가지 범주로의 반응양식에 의하여 개체와 범주 양쪽을 수량화하고, 그 수량을 이용하여 분류를 행하고자 하는 방법이다. 즉 行범주(개체)와 列범주를

수량화하여 低차원 상에서 行범주와 列범주의 대응관계를 살펴보고자 하는데 이용된다(최종후, 1997 ; 허명희, 1992 ; 田中 豊 & 脇本和昌, 1990).

한편, 대응분석은 프랑스에서 Jean-Paul Benzecri에 의해 주도적으로 발전된 통계기법으로 분할표 자료의 행과 열을 저차원 공간상의 점들로 동시에 나타내어 그들 상호간의 대응관계를 탐구

〈표 1〉 충남 시군의 인구·주택부문, 사회간접자본부문, 정보화부문에 대한 90~95 연평균 성장률(%)

시 군	총인구	생산 연령 인구	총 세대수	주택 총수	총인구 대 비 생산연령 인구비율	인구 밀도	주택 보급률	도로 포장률	운송 (100인당 자동차 수)	의료 (100인당 병상수)	100인당 총 우편물 접수 (천통)	100인당 전화 대수	조립금속 및 기계 제조업의 제조업 비중
천 안	1.2	8.1	5.9	6.7	7.31	-0.83	1.1	1.3	16.1	7.3	6.0	8.4	3.0
공 주	-2.9	2.3	2.1	1.8	4.49	-1.12	-0.4	-0.4	13.5	6.3	7.4	4.9	10.1
보 령	-3.8	-2.5	1.6	0.7	1.10	-1.08	-1.0	3.4	13.3	-0.5	1.0	4.2	14.3
아 산	-1.5	0.6	4.1	3.6	1.92	-0.97	-0.5	1.8	13.8	1.3	-0.2	6.2	8.0
서 산	0.9	-0.1	3.3	3.6	0.78	-1.07	0.4	3.3	13.2	3.1	2.4	5.2	-27.1
논 산	-3.4	-3.0	1.9	-0.7	0.31	-0.04	-1.2	3.7	13.1	1.0	1.9	-5.2	10.3
금 산	-3.7	-2.3	1.2	-0.5	1.17	-0.04	-1.1	0.3	12.8	0.8	-0.8	3.4	12.3
연 기	-2.3	-1.7	1.2	0.7	0.60	-0.03	-1.1	3.0	13.8	4.2	3.3	4.2	8.9
부 여	-2.6	-2.0	0.8	-0.1	0.53	-0.01	-1.0	5.5	13.3	1.7	0.0	2.6	0.7
서 천	-3.6	-3.4	0.4	1.0	0.16	-0.04	0.7	4.5	13.0	8.4	0.9	-3.4	12.2
청 양	-3.4	-2.9	0.2	0.2	0.50	-0.04	-0.3	6.8	13.7	.	0.6	2.4	-4.3
홍 성	-0.8	-0.1	1.5	0.9	0.63	-0.02	-0.4	6.6	12.2	2.7	-2.9	-4.2	10.6
예 산	-4.1	-3.2	1.4	0.2	0.74	-0.04	-1.1	6.2	12.6	15.2	-0.5	3.4	11.2
태 안	-3.1	-2.1	0.5	0.2	0.91	-0.05	-0.6	-0.2	13.5	7.0	1.6	3.1	13.6
당 진	-2.2	-0.6	2.6	1.0	1.46	-0.03	-1.8	10.2	13.5	1.4	-119.6	-0.5	-4.7

하려는 자료분석기법이다. 이 기법은 'reciprocal averaging', 'appropriate scoring', 'dual scaling', 'scalogram' 등의 이름으로 불리우기도 하지만 대응 분석이 가장 통용되는 명칭이다(최용석, 1993).

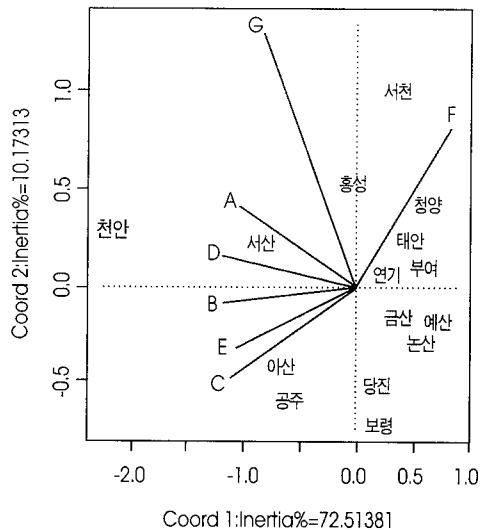
행렬도가 연속형 자료에 대한 기법인 반면 數量化 Ⅲ類, 대응분석은 이산형 자료를 다룬다는 점에서 차이가 있다. 본 연구에서 작성한 행렬도는 통계소프트웨어의 하나인 S-PLUS를 이용하여 구현하였다.

Ⅲ. 충청남도 시군의 성장유형 비교

1. 인구·주택부문

〈그림 1〉로 부터 인구·주택부문 성장유형의 지역적 특성은 주택보급률(G), 인구밀도(F), 총세대수(C) 성장에서 주로 나타남을 알 수 있다. 그 중 주택보급률(G)은 총세대수(C)에 무상관관계이며, 인구밀도(F)와는 약한 양의 상관관계를 보인다. 한편, 인구밀도(F)와 총세대수(C)는 음의 상관관계에 있음을 알 수 있다. 즉, 〈표 1〉에 의하면 인구밀도 성장은 공주(-1.12), 보령(-1.08), 서산(-1.07), 아산(-0.97), 천안(-0.83) 등에서 타 자치단체에 비

하여 감소추세가 큰 반면, 총세대수 성장은 공주(21), 보령(16), 서산(33), 아산(41), 천안(59) 등에서 타 자치단체에 비하여 높은 증가추세를 보여주고 있다.



A : 총 인구, B : 생산연령인구, C : 총세대수,
D : 주택총수, E : 총인구대비 생산연령 인구비율,
F : 인구밀도, G : 주택보급률

〈그림 1〉 인구·주택부문의 행렬도

〈표 2〉 인구·주택부문 성장유형에 따른 충남 15개 시군의 군집화

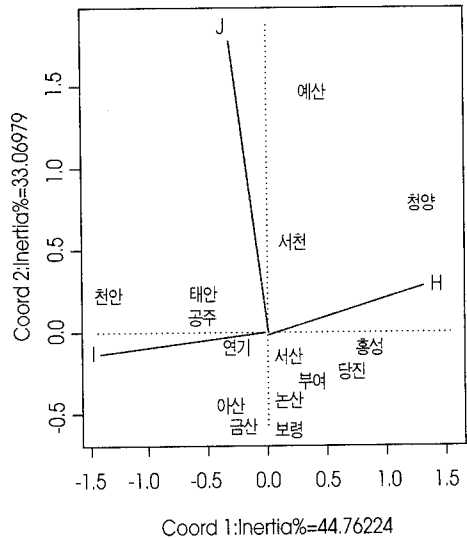
그룹	특성	시군
그룹 1	총인구(A), 주택총수(D) 성장	천안, 서산
그룹 2	총세대수(C) 성장	아산, 공주
그룹 3	인구밀도(F) 성장	청양, 태안, 부여, 연기
그룹 4	인구밀도(F), 주택보급률(G) 중간 성장	서천, 홍성
그룹 5	인구·주택부문 성장의 특성 없음	금산, 예산, 논산, 당진, 보령

또, 총인구(A) 성장은 주택총수(D), 생산연령인구(B), 총인구대비 생산연령 인구비율(E), 총세대수(C)의 성장순으로 양의 상관관계가 있으며, 인구밀도(F) 성장은 주택 보급률(G) 성장과의 사이에 약간의 양의 상관이 존재하고 있는 것으로 나타났다.

인구·주택부문 성장유형에 따라 충청남도 15개 시군을 군집화해 보면 <표 2>와 같이 5개 군집체(clustering)로 나누어 볼 수 있다. 특히, 천안은 주택총수(D) 성장, 아산은 총세대수(C) 성장, 그리고 청양은 인구밀도(F) 성장에서 깊은 특성을 나타내고 있다.

2. 사회간접자본부문

<그림 2>로부터 사회간접자본부문 성장유형의 지역적 특성은 의료(100인당 병상수 : J), 도로포장률(H), 그리고 운송(100인당 자동차수 : I) 성장에 두루 나타난다. 의료(J)성장은 도로포장률(H) 성장과 운송(I) 성장과 각각 무상관의 관계를 나타내고 있다. 그런데 도로포장률(H) 성장이 운송(I) 성장과 음의 상관관계를 보여주는 것은 이채로운 결과이다.



H : 도로포장률, I : 운송(100인당 자동차수),

J : 의료(100인당 병상수)

▶ 청양은 의료(100인당 병상수:J)가 결측치이므로 행렬도에 나타나지 않음.

<그림 2> 사회간접자본부문의 행렬도

즉, <표 1>에 의하면 도로포장률(H) 성장은 당진(10.2), 청양(6.8), 홍성(6.6), 예산(6.2), 부여(5.5), 서천(4.5) 등에서 타 자치단체에 비하여 증

<표 3> 사회간접자본부문 성장유형에 따른 충남 15개 시군의 군집화

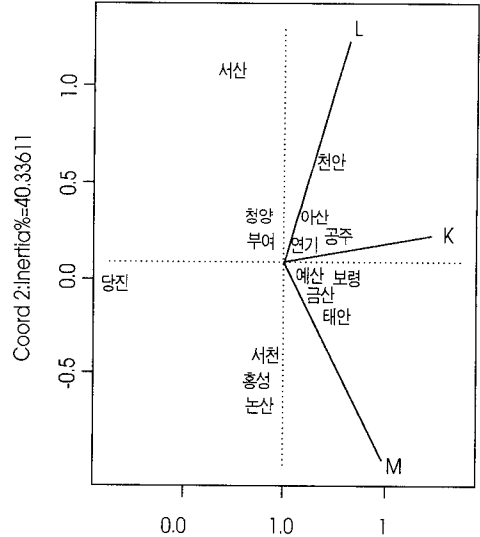
그룹	특 성	시 군
그룹 1	운송(100인당 자동차수:I) 성장	천안, 공주, 태안, 연기
그룹 2	도로포장률(H) 성장	홍성, 당진
그룹 3	의료(100인당 병상수:J) 성장	예산, 서천
그룹 4	사회간접자본부문 성장의 특성 없음	서산, 부여, 논산, 아산, 금산, 보령

가추세가 큰 반면, 운송(I) 성장은 천안(16.1), 아산(13.8), 연기(13.8), 청양(13.7), 태안(13.5), 당진(13.5), 공주(13.5) 등에서 타 자치단체에 비하여 높은 증가추세를 보여주고 있다. 특이한 점은 공주, 태안의 경우 운송(I) 성장이 연평균 135%로 기초자치단체 중에서 비교적 높은 편이나 도로포장률(H) 성장은 공주가 -0.4, 태안이 -0.2로 최하위에 자리하고 있다.

사회간접자본부문 성장유형에 따라 충청남도 15개 시군을 군집화해 보면 <표 3>과 같이 4개 군집체(clustering)로 나누어 볼 수 있다. 특히, 천안은 운송(I) 성장, 예산은 의료(J) 성장에서 깊은 특성을 나타내고 있다.

3. 정보화부문

<그림 3>으로 부터 정보화부문 성장유형의 지역적 특성은 조립금속 및 기계제조업의 제조업 비중(M), 100인당 총우편물접수(K), 그리고 100인당 전화대수(L) 성장에 두루 나타난다. 100인당 총우편



Coord 1:Inertia%=41.63653

K : 100인당 총우편물접수,

L : 100인당 전화대수,

M : 조립금속 및 기계제조업의 제조업 비중

<그림 3> 정보화부문의 행렬도

<표 4> 정보화부문 성장유형에 따른 충남 15개 시군의 군집화

그룹	특 성	시 군
그룹 1	100인당 전화대수(L) 성장	천안, 서산, 청양, 부여
그룹 2	조립금속 및 기계제조업의 제조업 비중(M) 성장	논산, 홍성, 서천
그룹 3	100인당 전화대수(L), 100인당 총우편물접수(K) 성장의 중간	아산, 공주, 연기
그룹 4	100인당 총우편물접수(K), 조립금속 및 기계제조업의 제조업 비중(M) 성장의 중간	태안, 보령, 금산, 예산
그룹 5	정보화부문 성장의 특성 없음	당진

물접수(K)와 100인당 전화대수(L) 간에는 약한 양의 상관관계가 있으며, 100인당 전화대수(L)과 조립금속 및 기계제조업의 제조업 비중(M)은 음의 상관관계가 있다. 100인당 총우편물접수(K) 성장은 조립금속 및 기계제조업의 제조업 비중(M)과는 약한 양의 상관관계가 있다.

정보화부문 성장유형에 따라 충청남도 15개 시군을 군집화해 보면 <표 4>와 같이 5개 군집체(clustering)로 나누어 볼 수 있다.

4. 인구·주택부문, 사회간접자본부문, 정보화부문의 총괄

<그림 4>로부터 충청남도 15개 시군의 인구·주택부문, 사회간접자본부문, 정보화부문의 총괄적 성장유형에 대한 지역적 특성을 살펴보고자 한다.

총인구(A), 생산연령인구(B), 총세대수(C), 주택총수(D), 총인구대비 생산연령 인구 비율(E), 운송(100인당 자동차수 : I), 그리고 주택보급률(G),

100인당 전화대수(L) 성장은 상호간 강한 양의 상관관계를 나타내고 있다. 특히, 총인구(A)와 총세대수(C), 생산연령인구(B), 주택총수(D)와 운송(100인당 자동차수 : I), 그리고 주택보급률(G)과 100인당 전화대수(L) 성장은 매우 강한 양의 상관관계를 나타내고 있다.

한편, 도로포장률(H) 성장과 100인당 총우편물접수(K), 조립금속 및 기계제조업의 제조업 비중(M) 성장, 그리고 인구밀도(F) 성장과 주택보급률(G), 100인당 전화대수(L), 총인구(A), 생산연령인구(B), 총세대수(C), 주택총수(D), 총인구대비 생산연령 인구비율(E), 도로포장률(I) 성장은 강한 음의 상관관계를 나타내고 있다.

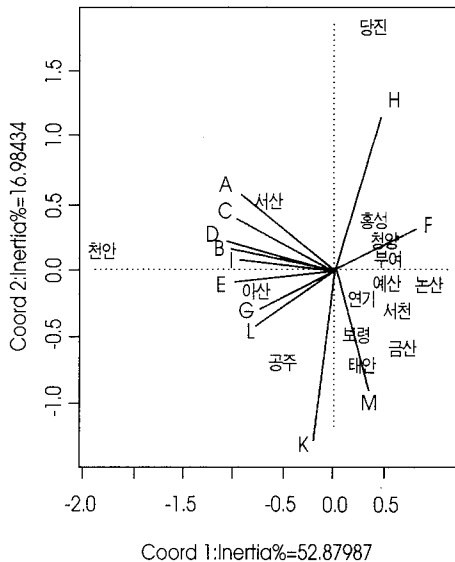
인구·주택부문, 사회간접자본부문, 정보화부문의 총괄적 성장유형에 따라 충청남도 15개 기초자치단체를 군집화해 보면 <표 5>와 같이 6개 군집체(clustering)로 나누어 볼 수 있다. 특히, 천안은 생산연령인구(B), 주택총수(D), 운송(100인당 자동차

<표 5> 인구·주택부문, 사회간접자본부문, 정보화부문의 총괄적 성장유형에 따른 충남 15개 시군의 군집화

그룹	특 성	시 군
그룹 1	생산연령인구(B), 주택총수(D), 운송(100인당 자동차수:I), 총세대수(C), 총인구대비 생산연령 인구비율(E), 총인구(A), 주택 보급률(G), 100인당 전화대수(L)의 성장	천안, 서산, 아산
그룹 2	주택 보급률(G), 100인당 전화대수(L)과 100인당 총우편물접수(K) 중간 성장	공주
그룹 3	도로포장률(H) 성장	당진
그룹 4	인구밀도(F) 성장	부여, 홍성, 청양
그룹 5	조립금속 및 기계제조업의 제조업 비중(M) 성장	태안, 보령
그룹 6	성장의 특성 없음	논산, 금산, 서천, 예산, 연기

차수 : I), 총인구대비 생산연령 인구비율(E) 성장에 깊은 특성을 나타내고 있으며, 당진은 도로포장률(H) 성장에 깊은 특성을 나타내고 있다.

열린충남



- A : 총인구, B : 생산연령인구, C : 총세대수,
D : 주택총수, E : 총인구대비 생산연령 인구 비율,
F : 인구밀도, G : 주택보급률 H : 도로포장률,
I : 운송(100인당 자동차수), K : 100인당 총우편물접수,
L : 100인당 전화대수,
M : 조립금속 및 기계제조업의 제조업 비중

〈그림 4〉 인구·주택부문, 사회간접자본부문,
정보화부문 총괄의 행렬도

참 고 문 헌

김군수·강원철, “경기도의 지역별 경제력 평가-비교우위 및 성장력 평가를 중심으로”, 통계분석연구, 제2권 제1호, 1997, pp.75~112.

최용석, SAS 대응분석, 서울:자유아카데미, 1993.
최종후, “15대 국회의원 총선거에 대한 수량화 분석”, 한국분류학회지, 제1권, 1997, pp.15~23.

허명희, 數量化 方法論의 이해, 서울:자유아카데미, 1992.

충청남도, 충청남도 통계연보, 1991·1994·1996.
_____, 충남 주민등록인구 통계작성 결과보고서, 1996.

_____, 사업체 기초 통계 조사보고서, 1995.
田中 豊 & 脇本和昌, 多變量統計解析法, 京都:現代數學社, 1990.

Gabriel K. R., “The biplot graphics display of multivariate matrices with applications to principal component analysis”, *Biometrika*, 58, pp.453~467.