韓國心理學會誌: 産業 및 組織
Korean Journal of Industrial and Organizational Psychology
1994. Vol. 7, No. 1.

요인분석의 관행과 문제점

이 순 묵
성균관대학교 산업심리학과

요인분석을 사용하여 경험자료를 분석하고 그 결과에 기초하여 논문을 작성할 때 그 절차와 결과에 대해 적정한 정도로 논문에 명시해야 할 필요가 있다. 그 필요는 논문의 심사자와 논문의 독자를 위한 것이다. 어떤 논문이 공인된 학술지에 실리기 위해서는 소수의 심사위원들이 평가를 하여 그 논문에서의 연구절차, 결과 및 논의가 과연 그 학술지의 독자들에게 도움이 될 수 있는 바람직한 수준인지를 판단하게 된다. 그러므로 심사위원들에게 그러한 평가를 허용할 수 있는 정도의 정보가 논문에서 본문에 또는 부록에 제시될 필요가 있다. 또한 학술지에 제시된 논문을 통해서 학문사회에서 지식이 누적되기 위해서는 필요한 정도의 정보가 제시되어야 한다. 이 연구에서는 요인분석이 응용될 경우 현재 심리학계의 학술지에서 얼마나 그런 내용들이 충실히 보고되고 있는지를 알아보고 앞으로의 개선점을 논의하기로 한다.

요인분석의 관행과 문제점

요인분석은 심리학을 비롯한 사회과학에서 개념의 구성을 위해 가장 많이 쓰이는 기본적인 방법론이다. 하나의 구성개념이 탄생되고 측정됨을 보기 위해서 사용되는 방법론인 만큼 그 사용의 적절성 및 解의 반복가능성에 대한 검토는 이론 발전의 기초를 다진다는 의미에서 반드시 거쳐야할 작업이다. 따라서 요인분석을 사용하여 경험자료를 분석하고 그 결과에 기초하여 논문을 작성할 때 그 절차와 결과에 대해 적정한 정도로 논문에 명시해야 할 필요가 있다. 그 필요는 논문이 담고 있는 학술적 논의의 소비자인 독자를 위한 것이다. 어떤 논문이 공인된 학술지에 실리기 위해서는 소수의 심사위원들이 평가를 하여 그 논문에서의 연구절차, 결과 및 논의가 과연 그 학술지의 독자들에게 도움이 될 수 있는 바람직한 수준인지를 판단하게 된다. 따라서 심사위원들에게 그러한 평가를 허용할 수 있는 정도의 정보가 논문에서 본문에 또는 부록에 제시될 필요가 있다. 또한 학술지에 제시된 논문을 통해서 학문사회에서 지식이 누적되기 위해서는 필요한 정도의 정보가 제시되어야 한다. 예를 들어 같은 주제에 대해서 많은 연구가 진행된 후에 제 삼의 연구자가 이제껏 진행된 연구들을 계량적으로 요약해서 어느 정도 일반화가 가능한 지식을 산출하고자 해도 각각의 논문에 필요한 정보가 수록되어 있지 않은 경우 어쩔 수가 없게 된다. 그래서 이 연구에서는 심리학을 비롯한 사회과학에서 가장 많이 쓰이는 방법론중의 하나인 요인분석을 사용하여 경험연구를 한 연구자가 보고해야 할 내용들에 대해서 논의하고, 현재 심리학계의 학술지에서 얼마나 그런 내용들이 충실히 보고되고 있는지를 알아보기로 한다. 우선 요인분석의 절차에 대해서 간단히 기술한 후 보고에 필요한 내용을 기술하고자 한다. 그리고 나서 국내 학술지에 실린 요인분석을 이용한 논문들을 검토하기로 한다.

요인 분석의 단계

요인분석을 진행하는 단계를 설명하면 다음과 같다.

1단계: 원자료의 수집, 측정변수의 검토 및 상관계수 계산.

일단 원자료를 수집한 후 의미가 크게 겹치지 않는 변수들만 뽑아서 분석에 사용하도록 한다. 변수들 가운데는 다른 변수(들)에 의해 거의 완벽하게 설명되는 변수가 있다. 이 경우 다중공선성(multicollinearity)이 대단히 높다고 하며 이런 겹치는 변수는 자료분석에서 사용할 필요가 없다. 즉 겹치는 변수끼리 합치거나 하나를 제거하거나 해야 할 것이다. 얼마나 서로 겹치는지를 알기 위해서는 각 변수를 모든 다른 변수에 회귀시켜본다. 즉 X1에서 X10까지 10개의 변수가 있을 경우 X1이 나머지 9개의 변수들에 의해서 얼마나 설명되는가를 알기 위해서는 X1을 종속변수로 하고 나머지 변수들은 예측변수로 하는 회귀분석을 하여 R^2를 구해본다.

상관계수행렬을 보고서도 그 자료가 요인분석에 적절한지 아닌지를 검토할 수 있다. 어떤 상관계수값의 크기가 1.0에 가까우면 그 두 변수는 합치거나 어떤 조정이 필요하다. 상관계수의 값들이 대체로 0에 가까우면 그 자료에서는 어떤 요인구조를 뽑아 내기가 어렵다.

2단계: 요인분석의 모형 선정

요인분석에는 두 가지 모형이 있다. 주성분분석(Principal Component Model)과 공통요인분석(Common Factor Model)이 그것인데 이 두 가지 중 어느 것을 택해야할 지를 결정한다. 그러나 사회과학에서는 공통요인분석이 더 바람직하므로 요인분석이라고 하면 공통요인분석을 가리키는 경우가 많다. 그러나 문헌을 참고할 때 일단은 "요인분석"이란 용어가 넓은 의미(공통요인분석, 주성분분석 포함)로 쓰이고 있는지 좁은 의미(공통요인분석만을 의미)로 쓰이고 있는지 상황에 따라 연구자의 판단을 필요로 한다.

모형의 선정은 곧 분석되는 상관행렬의 모양을 결정한다. 상관계수 행렬에는 두 가지가 있다. 원상관행렬과 축소(adjusted 또는 reduced)상관행렬이 있는데, 그들이 어떻게 산출되는가를 본 후에 두 가지 행렬에 대한 설명을 하기로 한다. 측정변수 Z가 있어서 $Z = a_1F_1 + a_2F_2 + U$ 라 하자. 여기서 F는 공통요인(common factor)이고 U는 고유요인(unique factor)이다. a는 요인계수(factor loading)이다[1]. 여기서 Z의 분산(variance)을 구해 보기로 하자[2].

분산은 자신과 자신간의 공분산이므로

$$VAR(Z) = COV(Z, Z) = COV(a_1F_1 + a_2F_2 + U, a_1F_1 + a_2F_2 + U)$$

Z과 Z간의 공분산에는 Z에 있는 F_1, F_2, 그리고 U들끼리의 공분산으로 구성된다. 좀더 자세히 보면 F_1의 분산, F_2의 분산, F_1과 F_2의 공분산, U의 분산을 의미한다. 따라서 Z의 분산에는 공통요인을 반영하는 부분과 고유요인의 분산을 반영하는 부분이 함께 포함된다. 이것은 모든 측정변수의 분산에 대해서 적용된다.

VAR(Z) = 1 = 공통요인이 차지하는 분산 + 고유요인이 차지하는 분산

= 공통분산(common variance) + 고유분산(unique variance)[3]

---

1) factor loading은 요인부하값으로 번역되기도 한다. 회귀분석에서 예측변수와 결과변수간 관계를 나타내는 계수를 회귀계수로 부른다. 요인분석은 수학적으로 보면 요인이 예측변수가 되고 측정변수가 결과변수인 회귀분석의 모형이다. 따라서 요인과 측정변수간의 관계를 나타내는 계수를 요인계수로 부르는 것이 더 일관성 있게 보인다.

2) variance를 심리학, 교육학에서는 변량으로 번역하는데 통계학 사전에서는 variate를 변량으로 번역한다. 따라서 심리학도가 변량이라고 할 때와 통계학도가 변량이라고 할 때는 서로 의미가 다른 개념을 가리키고 있는 것이다. 그런데 variance는 통계학에서 온 용어이므로 통계학사전의 번역인 '분산'을 따르는 것이 학문간 의사소통에 기여할 것으로 생각한다.

3) 고유분산은 변수 특유의 성질을 나타내는 특수분산(specificity 또는 specific factor의 분산)과 랜덤오차를 나타내는 오차분산(error variance)으로

원상관행렬(original correlation matrix)은 대각선에 각 변수의 분산이 1.0이 있는 것이고, 축소상관행렬은 이 대각선을 모두 각 변수의 공통분산으로 대체하여 수정한 행렬이다. 요인분석은 원상관행렬 또는 축소상관행렬을 분해해서 小數의 잠재변수, 즉 요인을 뽑아내는 방법이다. 그런데 원상관행렬을 분해하면 측정변수의 분산을 수정없이 분해하므로 공통요인의 부분과 고유요인의 부분이 함께 뭉쳐서 요인으로 빠져나온다. 이러한 요인을 특별히 주성분(principle component)이라고 부른다. 고유분산을 제외한 공통분산만을 가지고 요인으로 뽑아냈을 때 공통요인(common factor)이 추출되었다고 한다. 즉 주성분과 공통요인의 요인추출시 고유분산까지 포함하느냐 아니냐이다. 따라서 고유분산이 영이 아닌 한 주성분이 공통요인과 같을 수는 없다.

요인분석되는 상관행렬을 분해할 때 대각선에 있는 1.0들을 각 변수의 공통분산으로 대체한 후 분해를 해야 공통요인의 구조만 빠져나온다. 이것은 "자료의 축소"인 동시에 "과학적 연구를 위한 내용있는 구조의 추출"을 하게 되는 것이다. 원상관행렬을 분해할 경우 공통요인뿐만 아니라 고유요인 까지도 섞여서 뽑혀 나오므로 과학적으로 내용있는 구조의 추출이라 말하기 어렵고 단지 "자료의 축소"에 지나지 않는다.

공통요인이 아닌 주성분을 추출하는 요인분석의 모형은 주성분분석(Principle Component Model, Principle Component Analysis)이라 하고 축소상관행렬의 분해를 통해 공통요인만 추출하는 요인분석의 모형을 공통요인분석(Common Factor Model, Common Factor Analysis)이라고 한다. 주성분분석은 Karl Pearson을 기수로 한 통계학자들의 공헌에 의한 것이고 공통요인분석은 Charles Spearman을 기수로 한 심리학자 및 다른 사회과학자들의 공헌에 의한 것이다.

앞으로 '공통요인분석', '공통요인모형', 또는 '공통요인분석 모형'의 용어는 상호 교환적으로 쓰인다. 또한 '주성분분석', '주성분모형', 또는 '주성분분석 모형'이란 용어도 상호 교환적으로 쓰기로 한다.

원상관행렬의 대각선에 있는 원소들에서 고유분산이 영(zero)일 경우 원상관행렬과 축소상관행렬은 같게 되고 주성분분석의 결과와 공통요인분석의 결과는 같게 된다. 즉 자료에서 "고유요인=0"인 아주 극히 예외적인 경우에는 공통요인분석과 주성분분석은 다같이 공통요인을 추출한다.

그러나 사회과학에서의 측정에서는 고유요인이 차지하는 부분이 무시할 수 없을 정도로 크므로 그것을 제거하고 공통요인이라는 알맹이를 뽑아내는 작업이 필요하다. 하나의 변수에 대해서 수집된 자료에서 우리는 신뢰도를 보고하는데 이것은 측정변수의 분산가운데서 몇%만이 眞점수의 분산인지를 의미하는 것이다. 또한 眞점수가 곧 연구자가 재고자 하는 구성개념(construct)인 것은 아니다. 몇 개의 측정변수의 眞점수 부분을 기초로 하여 하나의 구성개념을 추출한다면, 하나의 측정변수의 분산에서 요인으로 뽑혀 올라가는 부분은 30%-50%에 지나지 않는 경우가 결코 드물지 않다(Widaman, 1983). 그렇다면 나머지 50% - 70%가 고유요인의 분산이 될 수 있을 것이다. 그러므로 사회과학자들이 내용적으로 의미있는 구성개념을 추출키 위해 주성분모형보다 공통요인모형을 선호하는 것은 당연하다.

그러나 주성분분석이 사회과학에서 전혀 유용하지 않다는 것은 결코 아니다. 자료가 너무 방대해서 우선은 "자료축소"를 해야 되겠다 할 경우라면 주성분분석이 필요할 것이다. 현실적으로 수백 개의 측정변수(예: 척도 개발시 수백 개의 문항을 사용할 경우)가 있을 경우, 우선은 수십 개의 주성분으로 축소하고, 그 주성분들을 측정변수처럼 사용해서 공통요인분석을 할 수도 있다. 또한 자연과학에서의 자료는 사회과학에서의 자료에 비해서 고유분산이 작다고 할 수 있다. 자연과학에서는 대부분의 측정상황에서 정확한 측정을 허용하는 물리적 도구의 이용이 가능하므로 측정변수에서 제거할 고유요인이 무시할 정도로 되는 경우가 가능할 수도 있다. 이 때 역시 주성분분석을 쓸 수 있는

나뉜다. unique factor를 고유요인으로 specific factor를 특수요인으로 번역한 것은 1995년 판 (중앙교육진흥연구소) 교육 측정·평가·연구·통계 용어사전에 따른 것임.

경우이다.

3단계: 공통분(communality)의 추정치 선정

공통요인분석을 할 경우 상관행렬을 분해하여 기초구조(initial solution, 기초解)를 구하는데는 여러 가지 방법이 있는데 그 중에는 각 변수의 분산 가운데 공통요인에 의해서 설명되는 부분, 즉 공통분(communality)의 추정치에 대해서 연구자의 결정을 필요로 하는 경우가 있다. 대체로 각 변수의 SMC(다중상관제곱치)를 공통분의 추정치로 사용한다. 그러나 원상관행렬이 逆을 가지지 않을 경우, SMC의 계산이 불가능하므로 각 변수가 다른 나머지 변수들과 가지는 상관계수의 절대치 中 가장 큰 값을 사용하는 경우가 많다. 공통분의 추정치를 경우에 따라서는 연구자가 직접 제공하기도 한다. 공통분의 값을 모두 1.0으로 주는 경우가 곧 주성분분석과 동일한 결과를 산출한다.

4단계: 분석상관행렬의 분해 및 요인의 수효결정

모형의 선정에 따라 분석될 상관행렬(분석상관행렬로 略함)이 원상관행렬 또는 축소상관행렬로 결정된다. 분석상관행렬의 분해(factoring)결과로 기초解(기초구조)가 얻어지는데, 기초解의 계산법으로서 주로 주축분해(principal axis factoring)[4]나 최대우도법(maximum likelihood method)을 사용한다. 기초解를 구하는 과정에서 분석상관행렬의 분해결과로 그 행렬의 고유치(eigenvalue)들이 얻어진다. 연구자는 고유치들 및 기타 다른 사항을 참조하여 몇 개의 요인을 뽑을 것인지 결정한다. 요인의 수효를 결정하는 방식은 분석상관행렬의 분해를 주축분해로 하느냐 최대우도법으로 하느냐에 따라서 다르다. 일단 요인의 수효가 결정되면 그에 따라 계산된 요인구조, 즉 기초구조를 나타내는 요인계수행렬을 얻을 수 있다. 이 행렬에서 2개요인 씩 짝지어 만들어진 2차원 평면상에 각 변수를 점으로 표시하면 기초해를 요인공간에 나타내는 것이다. 즉 기초해는 행렬로 얻을 수도 있고 요인공간의 그림으로 얻을 수도 있다.

그러면 요인수효의 결정에 대해서 주축분해법의 경우와 최대우도법의 경우에 대해서 알아본다.

1) 주축분해법으로 기초解를 구할 경우

보통은 주축분해법이 공통요인분석에 관련되어서 언급된다. 그러나 이 방법은 원래 1901년 Karl Pearson에 의해 "principal axis"란 이름으로 발표된 것으로서 주성분석에 먼저 쓰이기 시작하였다. 이 방식이 공통요인분석에 응용을 보게 된 것은 1930년대에 이르러 Hotelling의 공헌에 의한 것이다. 따라서 주축분해법은 기초구조를 구하는데 있어서 분석상관행렬이 원상관행렬인지, 축소상관행렬인지를 가리지 않고 사용될 수 있는 것이다. 주축분해법이 사용될 때 요인분석의 모형이 공통요인모형인지, 주성분모형인지에 따라서 요인의 수효결정 방식이 달라진다.

(1) 공통요인모형에 주축분해 사용시 요인의 수효결정

① Cattell(1966)의 스크리검사(scree test)

분석상관행렬의 고유치가 현저하게 작아지면서 평준화되기 前까지의 요인의 수효를 취하는 방식이다. 이 검사는 고유치의 값을 보고서 할 수도 있고 2차원 평면상에 고유치를 점으로 표시해서 검토할 수도 있다. 비교적 주관적 방식이다.

② Montanelli와 Humphreys(1976)의 평행성분석(parallel analysis)

자료內에 의미 있는 요인구조가 들어있다면 거기서 나온 고유치는 무선자료(random data)에서 구한 고유치보다 커야 할 것이다. 스크리검사를 위해 2차원 평면상에 수평축에 요인의 번호, 수직축에 경험자료의 고유치를 표시하듯이, 무선자료의 고유치도 같이 표시해 나갈 수 있다. 이때 경험자료의 고유치가 무선자료의 고유치보다 큰 상태(parallelism, 평행성)가 유지되는 범위 내에서의 요인의 수효를 경험자료내의 의미있는 요인수효의 상한값으로 생각하면 된다.

③ 누적분산비율(또는 누적분산퍼센트)[5]

4) factoring은 '분해', '추출'등으로 번역될 수 있다. 그러나 일단은 분석상관행렬을 분해해야 기초구조가 추출되므로 '분해'로 번역하였다.

요인의 고유치는 그 요인이 설명하는 분산이다. 따라서 추출되는 요인들이 누적해서 설명하는 분산의 퍼센트는 100%를 넘을 수가 없다. 공통요인분석에서는 누적분산의 비율이 전체공통분산의 75-85퍼센트가 되면 그 다음에 추가되는 요인이 설명하는 분산 퍼센트가 지극히 크지 않는 한 더 이상의 요인을 추가하지 않는다(Gorsuch, 1983). 그러나 저자의 경험에 의하면 누적분산비율이 100%가까이 또는 약간 넘게 나오는 경우도 많다. 기초구조 계산시 순서로 볼 때 처음에 나오는 요인의 고유치가 큰 양수가 되고 나중에 나오는 요인의 고유치가 음수가 되면 이들이 모두 합해서 100%인데 처음 몇 개의 요인만 추출할 경우 수학적으로는 100%가 넘을 수 있는 것이다.

④ 해석가능성

아무리 탐색적 분석이라 해도 연구주제의 문맥상 요인이 몇 개 정도라는 생각은 가지고 있는 것이 건강하다. 또한 1개의 요인을 뽑으려면 적어도 3개 이상의 변수를 기초로 해야 의미있는 해석이 된다(Thurstone, 1947; Zwick & Velicer, 1986). 또는 일단 몇 개의 요인수준을 설정하여 기초구조를 구한 후 그것을 회전해서 최종구조를 구한다. 이 최종구조가 해석가능한지를 보고서 앞서 설정한 '요인의 수효'가 적절한지를 판단할 수도 있다.

이상의 4가지 방식은 어느 하나가 가장 우수하다고 할 수 없고 모두 종합적으로 사용하여 가장 적절한 의사결정에 이르도록 한다.

(2) 주성분모형에서 주축분해 사용시 요인수효 결정

① 스크리검사

공통요인모형에서와 같은 방식으로 사용된다. 단지 여기서의 고유치는 원상관행렬의 고유치임에 유의한다.

② 평행성분석

Allen 과 Hubbard(1986)가 주성분모형에 대한 평행성분석을 개발하였다. 평행성분석을 하려면 무선자료의 고유치를 구하는 것이 핵심이다. 공통요인모형인 경우와 주성분모형인 경우에 계산되는 무선자료의 고유치를 비상업용의 프로그램 PARALLEL(Hays, 1987)을 가지고 계산해 볼 수 있다. Zwick과 Velicer(1986)는 평행성분석이 주성분을 과대추출하는 경향이 약간 있으나 다음에 설명하는 MAP방법과 함께 일반적으로 가장 정확했다고 보고하고 있다.

③ Velicer(1976)의 MAP(Minimum Average Partial)방법

m개의 주성분이 있다고 할 때, 처음부터 하나씩의 주성분이 통제된(partialled out) 상태에서 주성분간에 계산된 부분상관제곱들의 평균값이 아주 사소한(trivial) 값이 되기 前까지의 주성분을 추출한다. 이 값들이 사소하다는 것은 그만큼 중요한 주성분들이 이미 구해졌음을 의미한다. Reddon(1985)의 비상업용프로그램을 쓰면 이 방식을 쉽게 사용할 수 있다.

④ 카이저 방식

Kaiser rule 또는 Eigenvalues-greater-than-one-rule로 불리는 것으로서 Kaiser(1960)가 다음의 세 가지 논리를 기초로 개발한 것이다. 첫째로, Guttman(1954)은 공통요인분석에 대한 연구에서 모집단에서라면 이 방식이 공통요인수효의 하한선을 나타낸다고 증명하였다. 둘째로, 어떤 주성분이 설명하는 분산의 양이 어느 한 변수의 분산(1.0)보다도 적다면 그런 주성분은 별가치가 없는 것이라는 논리이다. 셋째로, 만일에 주성분분석에서 고유치가 1보다 작다면 주성분의 신뢰도는 음수가 되어 의미가 없어진다는 것이 Kaiser자신의 주장이다.

여기서 첫번째의 Guttman의 증명은 공통요인분석에 대한 것이므로 주성분분석에 이용하기에는 제한적이다. 공통분이 1.0인 특수한 경우의 공통요인분석이 곧 주성분분석이므로 쓸 수는 있다. 그러나 요즈음 공통요인분석에서도 카이저방식은 쓰이지 않는다. 두번째의 논리는 공식화된 논리는 아니고 非공식적 논리(informal reasoning)라고 Cliff(1988)는 비평하고 있다. 세번째의 논리는 Cliff(1988)에 의해 오류임이 증명되었다. 즉 주성분의 고유치가 1보다 작아도 그 신뢰도는 음수가 되지 않으므로 고유치의 값이 1보다 작은 주성분도 추출될 가치가 있다는 것이다.

Cliff(1988)에 의하면 카이저 방식에 의해서 주성분의 수효를 결정할 경우 원칙적으로는 과

5) 공통요인분석이므로 전체공통분산 중의 퍼센트를 의미한다.

소추출이 되지만, 현실적으로 표본에서는 표집 효과로 인하여 "1보다 큰 고유치의 수효가 증가"되는 경향이 있으므로 과대추출에 기여한다(p.279). 이미 Cliff(1988) 이전에 Zwick과 Velicer(1986)는 현재의 카이저 방식을 "절대적인 증거로 사용하는 것을 지지할 증거가 없으며... 이 방법이 더 이상 사용되어서는 안된다"(p. 441)고 까지 주장하고 있다. 이제 심리학계의 학술지에서 주성분의 수효결정을 위해 카이저 방식에 지나친 비중으로 의존하는 것은 탈피해야 할 것이다.

여기서도 역시 공통요인분석의 경우와 마찬가지로 여러 가지 방식을 종합적으로 사용해서 주성분의 수효결정을 내리도록 한다.

2) 최대우도법으로 기초解를 구할 경우

최대우도법은 경험자료의 모든 변수가 다변량정규분포[6]를 따른다는 가정하에 사용된다. 하나의 측정변수가 여러 개의 문항점수의 合 또는 평균일 경우 통계학의 중앙극한정리(central limit theorem)에 의해서 그 변수는 정규분포를 따르기가 쉬워지고, 변수간에 다변량정규분포가 성립할 가능성이 커진다. 그러나 검사를 개발할 때와 같이 하나의 문항이 하나의 측정변수로 사용될 때는 우선 그 변수 자체가 정규분포보다는 편포(skewed)된 분포를 따르기가 쉬우므로 최대우도법을 사용하기에 적합치 않을 수가 많다. 최대우도법이 사용될 때 요인분석의 모형에 따라서 요인수효결정의 방식이 다르다.

(1)공통요인모형에서 최대우도법 사용시 요인수효 결정

① 카이제곱검증

공통요인모형에서 사용되는 χ^2검증은 Lawley와 그의 동료(SAS에서는 Lawley & Maxwell, 1971, SPSS에서는 Joreskog & Lawley, 1968)에 의한 것이다. 여기서의 영가설과 대립가설은 아래와 같다.

영가설:모집단내에 m개의 공통요인이 존재한다.

대립가설:모집단내에 m개보다 많은 공통요인이 존재한다.

이 검증은 표본크기에 민감해서 표본이 크면 대부분의 영가설이 기각되는 경향이 있다. 따라서 카이제곱치를 해석할 때 단지 표본자료와, 가정된 요인구조 사이에 서로 합치되는 정도(degree of fit)로서 보고자 하는 의미에서 표본크기의 영향을 배제하도록 개발된 터커-루이스 계수가 더 중요한 참고로 사용된다.

② 터커-루이스계수

Tucker 와 Lewis(1973)가 개발한 이 계수의 값이 0.9이상일 때의 요인의 수효를 적절한 수효로 결정한다. 그 미만일 때는 가정된 요인의 수효가 너무 적음을 의미한다.

③ 해석가능성

'계량적인 방법'에 우선해서 고려되어야 하는 것이 바로 이 해석가능성이다. 따라서 주축분해시와 같은 요령으로 해석가능성을 검토한다.

(2) 주성분모형에서 최대우도법 사용시 요인수효 결정

주성분모형에 사용되는 χ^2검증은 Bartlett (1950, 1951)에 의한 것인데 공통요인모형에서의 χ^2와 같이 표본크기에 민감하다는 단점이 있다(Zwick과 Velicer, 1986). 이 방식은 통계패키지인 SPSS와 SAS에는 들어있지 않다.

5단계: 요인구조의 회전

측정변수의 수효가 n개일 때 요인의 수효가 m개로 결정되면 분석상관행렬의 분해결과로 제공되는 기초요인행렬은 n × m 의 행렬이 된다. 이 행렬은 기초구조를 행렬로 표현한 것이다.

기초구조는 항상 요인축사이의 각도가 90도

---

6) 다변량정규분포는 단일변수일 때의 정규분포개념을 다변수일 때의 상황으로 확대시킨 개념이다. 여기서 '多變量'이란 곧 여러개 변수를 의미한다. 다변수 상황에서 변수간에 상관관계가 전혀 없다면 다변량정규분포는 '각각의 변수가 모두 정규분포'라는 진술과 동일하다. 그러나 변수간 상관관계가 있을 경우 다변량정규분포를 설명하는데 이 진술만으로는 충분치 않다. 즉 변수간의 공변하는 정도가 설명되면서 정규분포의 개념이 단일변수에서 다변수의 상황으로 확장된 것이 다변량정규분포이다.

인 상태, 즉 요인간 상관이 영인 解이다. 이것을 보다 해석에 편리한 解로 만들기 위해서, 즉 요인 축들이 요인 공간상에 있는 변수들의 모임을 보다 잘 반영하도록 요인축을 회전한다. 이 때 직각을 유지하면서 회전을 하면 직각회전이고 직각을 깨뜨리면서 회전하면 사각회전(斜角回轉)이다.

직각회전에는 대체로 VARIMAX방식을 취하지만 사각회전은 통계패키지에 따라서 다르다. SAS에서는 Harris와 Kaiser(1964)의 Orthoblique를 쓰고 있고 SPSS에서는 Jennrich와 Sampson(1966)의 Direct Oblimin을 쓰고 있다. 사각회전에서는 파라미터의 값에 따라 요인축들을 최소한에서 최대한으로 사각회전을 한다.

회전 결과가 경험자료를 설명하는 정도는 회전 후의 요인행렬을 그림으로 나타낸 요인공간을 볼 때 새로운 요인축들이 변수들의 모임에 충분히 근접해서 지나가고 있는가를 보고 판단한다. 기초解가 되는 요인계수행렬(factor loading matrix) 즉, 기초요인행렬을 요인공간으로 옮긴 것(일종의 PREPLOT)을 볼 때 요인들 간에 큰 상관이 있는 것으로 보이는데도 불구하고 직각회전을 하거나, 사각회전을 하되 사각회전이 거의 되지 않도록 파라미터의 값을 주면 회전 후의 요인구조의 그림(요인공간의 PLOT)에서 경험자료를 충분히 설명치 못할 것이다.

직각회전과 사각회전 : 어느 것을 선택하는가?

사각회전이 직각회전에 비해서 복잡하고, 요인간 상관은 물론 최종요인행렬도 두 가지(형태계수행렬, 구조계수행렬)[7]나 되므로 해석할 때 조심해야 한다. 따라서 사회과학도들 사이에서 사각회전이 채택되는 속도가 느린 것은 하나도 놀랄 일이 안된다. 그러나 요인분석의 선구자인 Thurstone이나 Cattell(R. B. Cattell을 말함)은 그들의 후반기 연구에서는 모두 사각회전을 사용하였다. 직각회전은 사각회전의 특수한 경우이므로 Thurstone(1947)은 사각회전이야말로 과학도들의 목표에 부응하는 방법이라고 하였다. 이들 사각회전의 주창자들이 말하는 제일의 논리는 '표본에서 요인들 간의 상관이 영이라는 것은 비현실적'이라는 것이다. 이와 같은 논리의 입증을 위해서 Thurstone(1947), Cattell과 Dickman(1962), 그리고 Cattell과 Gorsuch(1963)등은 독창적으로 잘 조성된 실험을 통해서, 사각회전이 자료 속의 요인들 간의 관계를 보여주는데 더 유용함을 제시하고 있다. Cattell과 Dickman(1962)은, "직각회전과 사각회전에 대한 논쟁은 이미 [자신들이 글을 쓰기] 20여년 전에 끝난 것이지만"(p.390), 다시 한번 실험을 통해서 사각회전의 필요성("Necessity of Oblique Simple Structure", p.389)을 보인다는 것을 분명히 하고 있다.

"요인"들이란 곧 이론변수인데 이론구조의 구축에 사용되는 변수이다. 그런데 이 이론변수들간에 상관이 없는 가운데 이론구조가 구축되는 경우는 드물다. 이론변수들간의 관계는 서로간의 관계가 있을 때에 비로소 논의가 되는 것이다. 따라서 어떤 이론적 맥락 속에서 신중히 고려되어 추출되는 요인들이라면 서로간에 상관관계가 영인 경우가 흔하지 않은 것이다. 이론변수들간의 상관관계가 영이 아닐 때 우리는 비로소 그들 간의 관계구조를 여러 가지로 가설화해서(예: 직접관계, 간접관계, 제삼변수에 의한 관계, ...) 검증해 보는 의미가 있다. 특히 응용연구에서 이론적으로 '하나의 맥락'에 있는 요인들 간에 직각구조를 가정하면 그 요인들이 '하나의 맥락'으로 수렴함을 보일 수도 없는 모순을 경험하게 된다.

Loo(1979)는 임상심리분야의 요인분석연구에서 '회전'에 대해 검토한 결과 요인간 직각구조를 가정하고 있는 대부분의 연구가 정당하지 못함을 지적하였다. 오랫동안 '직각회전'의 주창자였던 Guilford(1981)도 그의 과거 연구들을 '사각회전'을 통해서 재분석하고 있다.

자료내 요인들 간의 구조가 직각구조인지 사각구조인지 알아보기 위해서는 기초구조의 요인들을 두개씩 짝지어 만든 2차원 요인공간상에 변수들의 점찍기를 해보는 것이 편리하다(Tabachnik & Fidell, 1989). 기초요인행렬의

7) 형태계수 행렬은 pattern loading matrix, 구조계수행렬은 structure loading matrix 의 번역이다.

가로줄의 값들은 각 변수가 각 요인의 축에서 가지는 좌표값이다. 점찍기는 회전후의 최종 요인행렬을 사용해서도 가능하다. 따라서 회전後의 점찍기(PLOT)에 대비해서 회전前의 점찍기를 통계패키지에서는 PREPLOT으로 부르는 경우가 있다. 여기서도 그 관행을 따른다. PREPLOT결과에서 점들이(즉, 변수들이) 모여 있으면 원점으로부터 그 모임(요인을 나타냄)에 이르는 직선을 그어 그 직선간 각도에 cosine을 씌워서 두 요인간의 대략적인 상관계수를 알아볼 수 있다. 그러나 요인간 상관을 알기 위해 이 정보에만 의존해서는 안된다. 요인의 수효가 3개 이상일 경우 2차원 평면에서의 이 방법은 요인간 상관을 잘못 추정할 우려가 있다. 요인간 상관을 알아보는 또 하나의 방법은, 일단 직각회전을 해서 얻은 최종 요인행렬을 검토하여 각 요인에서의 대표변수들을 (여러개일 수 있음) 알 수 있고, 두 요인을 나타내는 대표변수들간에 상관계수(변수간의 상관행렬 참조)를 요인간의 대략적인 상관관계로 간주하는 것이다(Gorsuch, 1983). 각 요인간 상관계수는 적어도 이러한 대표변수간의 상관계수 값만큼은 된다. 왜냐하면 측정변수간의 상관은 각 변수의 신뢰도가 완벽하지 않으므로 진점수(true score)간 상관에 비해 작은 값이 된다.

요인은 측정의 오차가 배제된 순수한 개념이므로 이들 간의 상관은 대표변수간의 상관보다 작지 않을 것으로 기대된다. 만일에 각 요인을 나타내는 대표변수들로 구성되는 소척도의 신뢰도가 구해질 수 있다면 고전 검사이론의 지식을 이용하여 요인간 상관의 값을 보다 정확하게 추정할 수 있다. 즉 한 요인에 대한 대표 변수들이 다른 요인에 대한 대표 변수들과 가지는 상관계수의 평균값을 각 요인에 대한 소척도들의 신뢰도에 루트를 씌운 값으로 나누어주면, 각 소척도들의 신뢰도가 완벽하게 개선될 때 그 소척도들의 진점수간 상관(이 경우는 두 요인간 상관이 됨)의 추정치가 된다.

만일에 PREPLOT에서 요인간 상관이 영(변수모임간 각도가 90도)으로 보이지 않는데도 직각회전(예: VARIMAX)을 해서 그 최종요인행렬을 기초로 2차원 공간상에 변수들을 점찍기할 수 있다. 즉 PLOT을 해보면 요인의 축들이 변수들의 모임에 근접해서 지나지 않음을 알 수 있다.

또한 사각회전시에 연구자가 제공하는 파라메터(SAS의 경우 orthoblique 회전에서 "HKP", SPSS의 경우 Direct Oblimin의 "DELTA")의 값에 따라서 요인간의 상관이 달라지긴 하지만 그 범위는 자료의 성질에 따라 제한된다. 예를 들어 사각회전을 최대한으로 시키는 파라메터 값을(예: SAS에서는 HKP=0) 줄 때, 요인간 상관의 최대값이 나오겠지만 그 값은 자료의 내용에 무관하게 움직이지는 않는다. 또한 자료내에 요인간 상관이 없는데 "사각회전을 한다고 요인간 상관이 반드시 발생하지는 않는다"(Tabachnik & Fidell, 1983, p.630). 한편 요인간 상관이 너무 크면(예: 중간정도 크기를 상회) 요인의 수효를 하나 줄여서 다시 분석해볼 필요가 있다(Gorsuch, 1983). 즉 상관이 큰 두개의 요인은 하나로 합쳐도 될 수 있는지 검토해 보는 것이다.

그런데 다른 통계방법과 마찬가지로 요인분석에서도 우발성의 이론화(capitalization on chance)가 발생한다(Gorsuch, 1983; Humphreys, Ilgen, McGrath, & Montanelli, 1969). 자료에 충실하게 분석을 하다보면 그 자료에 고유한 성질, 일반화될 수 없는 우발적인 부분까지 요인추출에 포함시킬 수가 있다. 요인분석의 각 절차(예: 요인수효결정, 기초구조계산, 요인회전)에서 우발성의 이론화를 피하기 위해 각 단계별로 한가지가 아닌 여러가지 방식으로 검토·분석할 것을 권한다. 그런 의미에서, 요인회전의 경우 사각회전이라 해도 Harris와 Kaiser의 orthoblique뿐 아니라 Jennrich와 Sampson의 Direct Oblimin 방식을 다같이 실시해 보는 것이 좋다. 또한 지나친 사각회전을 하지 않도록 항상 조심해야할 것이다. 그리고 PLOT의 결과에서 각 요인축상에 변수모임들이 깔끔하게 놓이지 않는다고 해서 무리하게 사각회전의 파라메터(orthoblique는 HKP, Direct Oblimin은 Delta)를 조정하여 요인간 상관을 지나치게 뽑을 경우 요인회전에서 '우발

성의 이론화'를 범할 수가 있다.

6단계: 요인점수의 계산

요인분석에서 주요 관심사는 요인구조를 뽑아내서 해석하는 것이지만 관찰된 응답자들이 각 요인에서는 어느 정도의 점수를 보이는지 계산해 보는 것이 때로는 쓸모 있는 일이다. 즉 연구자는 응답자들이 각 요인의 척도에서 어느 위치에 있는지 그 순수한 점수를 알고 싶어할 것이다. 왜냐하면 그 점수들이야말로 이론변수라고 하는 오염되지 않은 개념의 척도(尺度)로서 각 응답자를 잴 때의 값이기 때문이다.

요인분석후 보고사항

요인분석을 사용한 연구자가 심사자와 독자를 위해서 논문에서 보고해야할 사항은 다음과 같다.

원상관행렬, 표본크기

해외 학술지에서는 요인분석이 응용될 경우 원상관행렬이 제공되는 경우가 원칙이고 그렇지 않은 경우가 예외인데 국내의 학술지에서는 현재 그 반대이다.

분석에 사용되는 컴퓨터 소프트웨어

이것이 보고되지 않고는 심사자가 연구자의 논문을 재분석해서 계산결과의 정확도를 가려낼 수가 없다. 또는 일반적으로 인정되고 있는 소프트웨어인지 비공인의 소프트웨어인지 알 수가 없다.

요인분석의 모형 및 기초구조계산법

공통요인분석인지 주성분분석인지, 왜 어느 것을 택했는지를 분명히 한다. 기초구조의 추출을 위해 사용된 방법(예: 주축분해법, 최대우도법, ...)을 밝힌다.

변수의 공통분 추정방법

공통요인모형에서 기초구조를 구할 때 주축분해법 및 그에 유사한 방법을 쓸 경우에는 공통분의 추정방법을 명시해야 한다. 저자가 1990-1993까지 4개년간 국내 심리학회지의 일반, 임상, 상담, 사회, 발달, 그리고 산업및 조직의 6개 학술지에 발표된 공통요인분석 응용논문을 검토하였는데 어느 곳에도 공통분의 추정방법은 보고되어 있지 않았다.

요인의 수효를 결정한 방법 및 그 요인들이 추출될 때 전체분산 중 설명되는 분산의 퍼센트[8]

요인의 수를 결정하기까지 적절한 절차를 따랐는지 보고한다. 또한 그렇게 해서 추출된 요인들이 전체 분산 또는 전체 공통 분산의 몇 %를 설명하는지를 보고함으로써 추출되는 요인구조가 가지는 자료에 대한 설명력을 보인다. 단, 반복적인 주축분해나 최대우도법 사용시 설명분산의 퍼센트는 자동적으로 100%이므로 보고할 필요가 없다.

필요한 경우 축소상관행렬에서 구한 모든 요인의 고유치를 보고한다.

수학적으로는 고유치의 수효가 변수의 수효만큼 계산되지만 독자들이 참고해야될 정도, 즉 요인의 수효를 결정하는데 참고될 만한 수효의 고유치를 보고하도록 한다. 기초구조를 구하는데 주축분해법을 쓸 경우 고유치의 보고는 필수적이다. 연구자의 글을 읽는 사람들은 이것을 보고서 필요한 요인의 수효 결정이 합리적인지 대략적인 검토를 할 수 있다. 그러나 최대우도법을 쓸 경우 꼭 필요하진 않다. 국내의 학술지에서는 요인수효 결정과정을 검토할 수 있도록 충분히 보고하지는 않고 단지 추출되는 요인들에 대한 고유치만 보고하는 경향이 있다.

기초구조에 대한 회전 방법

국내의 학술지에서 대체로 회전의 이름(예: VARIMAX, Orthoblique)을 정확히 밝히고 있으나 경우에 따라서는 "직교회전" 또는 "사각

8) 여기서의 분산 역시 주성분분석이면 글자 그대로 변수의 분산(공통분산+고유분산)을 의미하고, 공통요인분석이면 공통분산을 의미한다.

회전"으로만 표시해서 정작 무슨 회전인지 알 수 없는 경우가 있다.

최종요인행렬 및 요인상관행렬

직각회전을 했을 경우 최종요인행렬은 하나 뿐이고 요인간 상관은 영이다. 그러나 사각회전을 했을 경우 구조계수행렬과 형태계수행렬을 제공하는 것이 바람직하다. 지면이 허락하지 않을 경우 최소한 구조계수행렬만이라도 보고해야 한다. 사각회전의 경우 요인간 상관행렬 또한 반드시 보고되어야 한다. "구조계수행렬 = (형태계수행렬) × (요인간 상관행렬)"이므로 구조계수행렬과 요인간 상관행렬만 있어도 읽는 사람은 형태계수행렬을 구할 수 있다. 사각회전시 요인 해석에서 주로 참고되는 행렬은 구조계수행렬이다(Brogden, 1969; Gorsuch, 1983). 물론 형태계수행렬도 함께 참고가 된다.

최종구조에서 요인을 해석하는 방침

보통은 계수가 0.3 - 0.4 이상인 계수에 대응하는 측정변수들의 의미를 반영하도록 요인을 해석하지만 0.3 - 0.4를 너무 경직된 기준으로 사용할 필요는 없다.

최종구조가 직각구조일 경우 각 요인이 설명하는 분산(요인분산), 및 그 비율(요인분산퍼센트)

최종구조가 직각구조이면 각 요인간 상관이 영이므로 요인이 설명하는 분산의 비율은 개별 요인의 중요성으로 해석될 수 있다. 그러나 사각회전을 했을 경우 요인간 상관 때문에 각 요인이 설명하는 분산을 해석함에 특별한 주의가 필요하다.

요인점수를 계산했을 경우 사용된 방식

SAS의 경우 회귀방식밖에 없지만 SPSS의 경우 여러 가지 방식이 있다.

이상의 보고사항중 지면을 가장 많이 차지할 가능성이 있는 부분은 원상관행렬과 최종요인행렬이다. 나머지 사항들은 간단히 언급하거나 계산결과만 몇 줄 보고할 수도 있는 내용들이다. 지면관계상 위의 사항을 모두 보고할 수 없을 경우, 어느 것을 보고하고 안하느냐 하는 것은 요인분석이 그 연구에서 얼마나 중추적인 역할을 한 방법론인가에 달려있다. 또한 연구자로서는 보고하지 못한 사항들에 대해서 독자가 요구할 경우 반드시 제공할 수 있어야 한다. 독자가 요구하기 전에 논문 심사자(출판을 위해 심사받을 경우)의 요청이 먼저 있을 것이다.

다음은 국내 학술지에서 이상의 필요적 보고사항들이 어떻게 보고되고 있는지 알아보기로 한다.

방 법

심리학회에서 발간되는 여러 학회지(일반 및 각 분과회지)중에서 非실험적인 연구방법을 사용한 논문들을 비교적 쉽게 발견할 수 있는 일반, 사회, 산업, 임상, 상담, 및 발달의 6개 학술지를 90년호부터 93년호까지의 4개년分을 검토하여 요인분석이 응용된 22개의 논문들을 조사하였다. 하나의 논문에서 요인분석의 응용을 한번 또는 그 이상으로 하고 있으므로 22개 논문에서 보고된 요인분석응용 사례는 총 61개가 된다. 검토함에 있어서 각 논문들이 다음의 사항들을 적절하게 보고하였는지 중점을 두었다: 변수수효, 원상관행렬, 표본크기, 사용된 소프트웨어, 요인분석 모형, 기초구조 계산법, 요인수효 결정법, 추출된 전체요인에 의해 설명되는 분산%(설명분산%), 고유치 보고, 기초구조 회전법, 최종요인행렬, 요인해석 방침, 각 요인이 설명하는 요인분산%, 요인점수 계산방식. 결과는 **표 1**에 제시된다.

표 1에서 검토된 22개의 논문들에서 요인분석이 본 연구 또는 예비연구에 사용된 것을 구분해서 기재하지는 않았다. 요인분석이 사용된 과제가 예비연구였을 경우, 본 연구에 사용될 경우에 비해서 보고 요망사항이 단순해질 수 있는 부분이 있겠으나 그것은 각 연구 수행자의 판단에 따라 정도의 차이가 있을 것이다.

표1. 탐색적 요인분석이 응용된 논문

( ): 저자의 추정, V: VARIMAX, J: Jennrich와 Sampson의 사각회전(direct oblimin),
HK: Harris-Kaiser의 사각회전(orthoblique case 2), O: 보고, X: 보고 누락, -: 보고 불필요,
r: 표본크기 ÷ 변수수효, PA: 주축분해법, ML: 최대우도법

저자및 연도	변수수효	상관행렬	표본크기	r	소프트웨어	요인분석모형	기초구조계산법	요인수효결정법	설명되는분산%	고유치	기초구조회전법	최종요인행렬	요인해석방침	각요인분산%	요인점수계산방식
김동기(1991) 종교정향 연구	66	X	153	2.3	SAS	X	(PA)	고유치≥1	X	X	X	X	X	X	(회귀)
김애순, 윤진(1991) 중년기 위기감 척도	100	X	299	3.0	SAS	공통요인	PA	스크리검사 평행성검사 누적분산비율 해석가능성	0	부분	HK	X	요인계수 .34~.40 이상	X	-
김원형(1993) 조직에서의 동일시,내재화, 몰입	53	X	513	9.7	SAS	공통요인	PA	고유치>1 스크리검사 누적분산비율	X	부분	V	큰계수만	요인계수 .3 이상	X	-
김준곤등(1993) 대학생활만족	63	X	830	13.2	X	X	X	X	X	부분	X	큰계수만	X	X	-
박동건(1992) JDI(번역판)의 하위척도들에서 일차원성 검토															
직무자체	18	X	355	19.7					X	부분	-	X		-	-
임금	9	X	355	39.4					X	부분	-	X		-	-
승진기회	9	X	355	39.4					X	부분	-	X	요인계수의 상대적 크기	-	-
감독	18	X	355	19.7	SPSS	공통요인	PA	1번 요인과 나머지 요인들의 고유치 비교	X	부분	-	X		-	-
동료작업자	18	X	355	19.7					X	부분	-	X		-	-
직무일반	18	X	355	19.7					X	부분	-	X		-	-
박병관등(1991) MMPI와 16PF MMPI 기초한 성격장애척도	11	X	94	8.5	X	주성분	(PA)	고유치>1, 스크리검사, 해석가능성	0	부분	V	0		0	-
MPDS와 16PF 1차요인	27	X	94	3.5	X	주성분	(PA)	고유치>1	0	부분	V	0	요인계수 .3 이상	0	-
MPDS와 16PF 2차요인	17	X	94	5.5	X	주성분	(PA)	고유치>1	0	부분	V	0		0	-
박세영(1991) 조직 공정성	34	X	246	7.2	X	주성분	PA	고유치≥1	0	부분	V	큰계수	X	0	-
직무만족	42	X	246	5.9	X	주성분	X	X	0	부분	V	큰계수	X	0	-
신현균, 원호택(1991) 비합리적 사고검사	X	X	288	X	X	X	X	X	0	X	X	X	X	0	-
오경자, 이혜련(1990) 아동행동목록(BCL)								1.직각회전과 사각회전 결과에서 일관성있는 부분, 2.요인계수가 .3이상인 문항이 6개 이상인 요인추출							
4- 5세	108	X	188	1.7	X	주성분	(PA)		X	X	V,J	X	요인계수 .3 이상	X	-
6-11세	115	X	582	5.1	X	주성분	(PA)		X	X	V,J	X		X	-
12-16세	117	X	192	1.6	X	주성분	(PA)		X	X	V,J	X		X	-
이상민(1991) 여성의류연구	25	X	240	9.6	X	(공통	ML	고유치>1,	0	-	V	0	X	0	-
치약연구	25	X	240	9.6	X	요인)	ML	해석가능성	X	-	V	0	X	0	-
이상희, 김재현(1993) 상담시간 평가															
상담자	12	X	108	9.0	SAS	(주성분)	PA	X	0	부분	V	X	X	0	-
내담자	12	X	108	9.0	SAS	(주성분)	PA	X	0	부분	V	X	X	0	-
상담직후 정서															
상담자	12	X	108	9.0	SAS	(주성분)	PA	X	0	부분	V	X	X	0	-
내담자	12	X	108	9.0	SAS	(주성분)	PA	X	0	부분	V	X	X	0	-
이연희, 박광배(1992) 절망척도의 일차원성 검토	20	X	1619	81.0	X	주성분	(PA)	고유치>1 스크리검사 해석가능성	X	부분	X	0	변수의 평균과 요인계수간 상관, 큰계수와 작은계수 비교	0	-

(표 1 계속)

저자및 연도	변수 수효	상관 행렬	표본 크기	r	소프트 웨어	요인분 석모형	기초구조 계산법	요인수효 결정법	설명 되는 분산 %	고유치	기초구조 회전법	최종요 인행렬	요인해석 방침	각 요인 분산 %	요인점 수계산 방식
이영호, 송종용 (1991) BDI 자료															
대학생	21	X	159	7.6	X	(공통	(ML)	χ^2 검증	0	-	V	0	X	0	-
연수원교육생	21	X	256	12.2	X	요인)	(ML)	χ^2 검증	0	-	X	0	X	0	-
임영란, 안창일 (1992) MMPI내용척도															
정상인	15	X	890	59.3	X	주성분	X	X	X	부분	V	0	요인계수 .3 이상	0	-
대학생	15	X	273	18.2	X	주성분	X	X	X	부분	V	0		0	-
환자	15	X	450	30.0	X	주성분	X	X	X	부분	V	0		0	-
장재윤(1993) 기회/위협의 시나리오에 대한 타당도 검토	6	X	90	15	X	X	X	X	X	X	X	0	X	X	-
전겸구, 김교헌 (1991) 생활스트레스 척도															
대인관계	40	X	301	7.5	X	주성분	(PA)	고유치≥1	0	X	V	X	X	X	-
당면과제	40	X	301	7.5	X	주성분	(PA)	고유치≥1	0	X	V	X	X	X	-
전겸구, 이민규 (1992) CES-D 한국판	20	X	220	11.0	X	주성분	(PA)	고유치>1 스크리검사	0	부분	V	0	X	0	-
전현민, 노명래 (1991)															
자아개념검사	100	X	301	3.0	X	주성분	(PA)	고유치≥1	0	X	X	큰	요인계수 .4 이상	0	(결정)
정신건강검사	96	X	301	3.1	X	주성분	(PA)	고유치≥1	0	X	X	계수만		0	
차재호, 정지원 (1993)															
집합주의 20대	32	X	300	9.4	X	주성분	(PA)	고유치≥2	0	X	V	큰	요인계수 .3 이상	0	(결정)
50대	32	X	300	9.4	X	주성분	(PA)	고유치≥2	0	X	V	계수만		0	
최광선(1993)								모두							몇개
성격인지	20	X	62	3.1	X	주성분	(PA)	고유치>1	X	X	V	0	X	0	문항
자질	30	X	62	3.1	X	(주성분)	(PA)	해석가능성	X	X	X	0	X	0	점수
최미례, 김중술 (1990) KWIS평가치분석															
정신병 집단	11	X	78	7.1	X	주성분	(PA)	선행연구의	0	부분	V	0	요인계수의 상대적	0	-
신경증 집단	11	X	80	7.3	X	주성분	(PA)	결과를 따름	0	부분	V	0	크기	0	-
최윤경등(1993) MMPI 연구															
표준척도															
환자 남	15	X	244	16.3	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
여	15	X	256	17.1	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
정상인 남	15	X	149	9.9	X	주성분	(PA)	고유치>1	X	부분	(V)	큰계수		0	-
여	15	X	159	10.6	X	주성분	(PA)	고유치>1	X	부분	(V)	만	요인계수	0	-
대학생 남	15	X	161	10.7	X	주성분	(PA)	고유치>1	X	부분	(V)		의 상대적	0	-
여	15	X	46	3.1	X	주성분	(PA)	고유치>1	X	부분	(V)		크기 비교	0	-
내용척도															
환자 남	15	X	244	16.3	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
여	15	X	256	17.1	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
정상인 남	15	X	149	9.9	X	주성분	(PA)	고유치>1	X	부분	(V)	큰계수	요인계수	0	-
여	15	X	159	10.6	X	주성분	(PA)	고유치>1	X	부분	(V)	만	의 상대적	0	-
대학생 남	15	X	161	10.7	X	주성분	(PA)	고유치>1	X	부분	(V)		크기 비교	0	-
여	15	X	46	3.1	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
성격장애척도															
환자 남	15	X	244	16.3	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
여	15	X	256	17.1	X	주성분	(PA)	고유치>1	X	부분	(V)		요인계수	0	-
정상인 남	15	X	149	9.9	X	주성분	(PA)	고유치>1	X	부분	(V)	큰계수	의 상대적	0	-
여	15	X	159	10.6	X	주성분	(PA)	고유치>1	X	부분	(V)	만	크기 비교	0	-
대학생 남	15	X	161	10.7	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-
여	15	X	46	3.1	X	주성분	(PA)	고유치>1	X	부분	(V)			0	-

* 공통요인분석에서 기초구조를 구하기 위해 주축분해를 할 경우 공통분의 추정치가 필요하다. 표 1에서 제시된 공통요인분석중 어느 것도 이 추정치에 대해서는 언급하지 않았다.

결 과

표 1을 보면 요인분석이 응용되는 현실을 쉽게 파악할 수가 있다. 우선 표 1을 심리학의 각 분과별로 정리하면 표 2와 같다.

표 2. 각 분과誌별 요인분석을 응용한 연구 (90년-93년)

	임상	산업	사회	일반	상담	발달	합계
논문수	10	5	3	2	1	1	22
요인분석 사례	37	7	5	7	4	1	61

그러면 앞서의 '요인분석후 보고사항'의 부분에서 언급된 순서대로 관행과 문제점을 논의하기로 한다.

원상관행렬, 표본크기

요인분석은 측정변수간의 상관계수행렬을 분석해서 소수의 잠재변수를 추출하는 방법이다. 그런데 원상관행렬을 보고한 경우를 보면, 61개의 사례중 전무하였다. 원상관행렬이 보고되지 않으면 그 논문의 심사자는 요인분석의 부분에 대해서 심사를 할 수 없게 된다. 즉 심사자가 논문에 실린 결과를 확인하기 위해 재분석을 할 수가 없게 되는 것이다. 물론 심사자가 편집진에게 부탁해서, 편집진이 연구자로부터 원자료나 상관행렬자료를 입수해서 심사자에게 전달하여 재분석이 되고 그 결과로 논문이 심사에서 통과되어 학술지에 실렸을 가능성을 배제할 수는 없다. 그러나 요인분석이 주요 방법론으로서 사용되었을 경우, 변수의 수효가 극단적으로 많지 않는 한 상관행렬이 논문에 보고되므로서 심사자나 독자가 재분석을 할 수 있어야 한다. 재분석을 할 때 논문의 저자가 실시한 방식은 물론 또 다른 방식으로도 해보아서 최적의 解가 보고되었는지를 심사자는 평가해야 한다. 표 1에 제시된 61개의 사례에서 전혀 상관행렬이 보고되지 않았음은 논문의 제출자들이 심사자들로부터 받을 수 있는 학문적 평가와 조언의 기회가 상실되어 왔음을 의미한다. 이것은 한편 학술지가 논문의 독자에 대한 서비스를 충분히 하고 있지 못한 것으로 이해될 수도 있다.

표본크기는 61개의 사례에서 모두 보고되었다. n개의 변수가 있을 때 필요한 표본크기가 어느 정도이어야 하는지에 대한 통일된 의견은 없다. Gorsuch(1983, p.332 참조)는 1개의 변수에 대해 관찰수효가 적어도 5배 이상이어야 한다고 하고 Nunnally(1978, p.276 참조)는 10배 이상이어야 한다고 주장한다. 표본크기가 변수수효의 5배가 안되어도 별 문제가 없다고 주장하는 사람은 없다. 요인구조가 확실하게 들어있는 경우는 표본크기가 작아도 별 무리없이 요인구조가 드러나겠지만 글자 그대로 탐색적인 경우에는 표본이 크지 않을 경우 요인구조의 발견이 어려워진다. 표 1에 제시된 61개의 사례중 60개의 사례에서 변수수효를 제시하고 있으므로 표본크기를 변수수효로 나눈 비율을 구할 수 있었다. 표 1에 있는 r을 정리하면 표 3과 같다.

표 3. '표본크기/변수수효'의 비율

비 율	사 례 수
5.0 미만	12
5.0 - 10.0	22
10.0 이상	26
합 계	60
평균: 12.96, 최소: 1.6, 최대: 81.00, 표준편차: 13.39	

표 3을 보면 평균적으로는 표본크기 對 변수수효의 비율이 12.96이지만 5.0미만의 비율을 보이는 사례가 12개나 된다. 적어도 이 12개의 사례에서는 그 解의 안정성을 확보하는데 어려움이 있을 수 있다. 즉 유사한 방법으로 그런 크기의 표본을 뽑아서 요인분석을 할 경우 아주 다른 解를 구할 가능성이 대단히 크다. 또한 그 비율이 5.0-10.0 사이의 사례도 20개나 되는데 이 경우 역시 자료內에 요인구조가 뚜렷하지 않은 경우라면, 새로운 자료를 구해서 분석할 때 상이한 해석이 도출될 가능성이 항상 존재함을 인식해야 할 것이다.

분석에 사용되는 컴퓨터 소프트웨어

분석에 사용된 소프트웨어의 보고는 심사자가 재분석을 통해 분석결과를 확인하는데 반드

시 필요하다. 표 1에서 보면 사용된 소프트웨어를 보고하지 않은 경우가 거의 전부이고 예외적으로 몇 개의 연구에서 보고하고 있다. SAS를 사용한 사례가 7개 사례이고 SPSS를 사용한 사례가 1개로 보고되고 있다. 저자의 관찰에 의하면 국내에서 요인분석시에 대체로 SPSS를 사용하는 것으로 알고 있다. 단지 저자들이 연구에서 밝히지 않았을 뿐으로 생각된다. 보다 정교한 요인분석을 위해서는 SAS가 SPSS보다 더 유용하지만 처음 사용하기에는 SPSS가 편리하다. SAS의 경우 분석과정에서 선택적인 의사결정을 연구자가 제공해야하는 어려움이 있고 SPSS는 프로그램 내에 기능들이 일반적인 것으로 고정되어 있어서 초보자들의 의사결정을 대신하는 경우가 많다. 그러나 SPSS를 사용할 경우 주성분분석이나 공통요인분석에 상관없이 주성분분석의 고유치가 제공되므로(MacCallum, 1983) 요인의 수효결정에 있어 연구자를 오도할 우려가 있으므로 공통요인분석을 할 때는 특별히 주의해야 한다.

요인분석의 모형 및 기초구조계산법

표 1을 보면 요인분석의 모형으로서는 주성분모형이 44개 사례에서, 공통요인모형이 단지 12개 사례에서만 사용되었다고 보고되고 있다. 나머지 5개 사례에서는 어떤 모형인지 보고되지 않았다. 우선 모형이 보고되지 않으면 뽑힌 요인이 주성분인지 공통요인인지 독자로서는 판단할 수가 없다. 또한 모형의 선택에 있어서 왜 주성분모형을 선택했으며 또는 공통요인모형을 선택했는가에 대한 분명한 설명없이 그냥 주성분분석 또는 공통요인분석을 했음을 보고하고 있다. 사회과학에서 내용이 축약된 어떤 구조를 뽑아내기 위해서는 자료내 측정의 오차 및 구성개념에 포함될 수 없는 변수의 특수성을 배제해야 하므로 공통요인분석이 바람직함을 밝힌 바 있다. 그렇다면 연구자가 굳이 주성분모형을 쓸 때는 그 목적이 어디에 있는지를 분명히 밝혀야 할 것이다. 그 목적이란 자료의 축소, 고유요인이 별로 없는 물리적 자료의 내용을 요약, 또는 기타 합당한 목적이어야 할 것이다. 주성분분석이 통계학에서 발달하여 오랜 역사를 지니고 있기 때문에 사회과학도들이 통계학과에 가서 요인분석의 수업을 들을 경우 주성분분석을 배우게 되고 그 익숙함에서 오는 영향이 사회과학에 공통요인분석의 보편화 속도를 느리게 하는 것은 사실이지만 이제는 사회과학에서 공통요인분석을 가르치고 있으므로 방향전환이 크게 진전될 것으로 보인다.

기초구조의 계산법으로서는 주축분해가 단연코 선호된다. 1950년대와 60년대에 요인분석의 발달에 중심적 위치를 차지한 Henry Felix Kaiser(1927-1992)는 40년간을 요인분석에 외골수로 집중적 연구를 하였으나 최대우도법의 사용에 대해서는 끝까지 호감을 보이지 않았다. 그것은 최대우도법에서 사용되는 반복 알고리즘이 공통분을 과다추정하여 Heywood 사례(공통분이 1이상이 됨)를 만들어 내는 것에 대한 실망때문이다. 물론 최대우도법에는 측정변수에 대한 분포의 가정이 적절할 때 요인의 수효를 검증해 볼 수 있는 장점이 있다. 따라서 최대우도법의 사용에는 나름대로의 목적이 있다.

표 1에서 보면 49개 사례에서 주축분해법의 사용을, 4개 사례에서 최대우도법의 사용을 보고하고 있다. 나머지 8개 사례에서는 어떤 방식으로 기초구조를 계산했는지의 보고가 없다. 이 경우 역시 논문에서의 결과를 심사자나 독자가 검토할 수 없는 경우가 된다. 같은 자료에 상이한 계산법을 사용해서 일관성있게 나오는 정보가 신뢰할 만한 정보이다. 따라서 기초구조계산시에 두 방법을 사용해 보는 것은 의미있는 노력이다. 그러나 최대우도법을 쓸 때는 측정변수간 분포의 가정이 크게 위반되지 않을 것을 전제로 한다. 심리학에서 검사를 개발할 경우 낱개의 문항이 하나의 측정변수가 되며 이 때는 다변량정규분포의 가정이 성립되기 쉽지 않다. 낱개 문항에 대한 응답은 치우친(skewed) 분포를 보이는 경우가 많기 때문이다. 이런 경우엔 최대우도법의 결과는 그야말로 '탐색적' 의미에서라면 참고가 되겠지만, 그 이상의 의존을 해서는 안될 것이다. 또한 최대우도법은 현재로서는 공통요인분석을 위한 것만이 표준적 통계패키지(예: SAS, SPSS등)에 들어있고 주성분분석을 위한 것은 아직 들어있지 않다.

변수의 공통분 추정방법

일단 공통요인모형을 선택하고 기초구조계

산법으로서 주축분해법을 사용하게 되면 축소상관행렬을 만들기 위해서 각 변수의 분산(1.0)을 공통분으로 대체할 필요가 있다. 이 때 가장 많이 쓰이는 것은 그 변수를 종속변수로 하고 나머지 다른 변수들을 예측변수로 하여 회귀분석을 할 때 얻어지는 R^2, 즉 다중상관제곱치(Squared Multiple Correlation, 줄여서 SMC로 부름)이다. Dwyer(1939)는 어느 변수의 공통분은 그 변수의 SMC와 같거나 그 보다 큰 값이라고 했으나 Jackson 과 Tweed (1980)는 공통분이 SMC보다 항상 크다고 증명하고 있다. 그러나 뽑고자 하는 공통요인의 수효에 비해 측정변수의 수효가 클수록 각 변수의 SMC는 그 변수의 공통분에 근접해 가므로 SMC가 보편화된 추정치가 되었다. 그러나 이것만이 추정방식은 물론 아니다.

예를 들면 원상관행렬이 逆을 갖지 않을 경우 SMC가 계산되지 않으므로 각 변수가 다른 나머지 변수들과 가지는 상관계수의 절대치중 가장 큰 값을 공통분추정치로 사용하는 수가 많고 경우에 따라서는 연구자가 직접 적절한 값을 제공할 수도 있다. 그런데 공통분의 추정방식에 따라 축소상관행렬의 대각선 값이 달라지게 되므로 그 방식을 밝히는 것이 필요하다. 그러나 공통요인모형이라 해도 기초구조계산법으로서 최대우도법을 택하면 공통분의 추정치를 연구자가 제공할 필요가 없다. 단지 "몇 개의 요인이 있을 것이다"하고 가정을 제공하면 그 가정을 검증하는 과정의 부산물로서 공통분의 값이 계산되어 나온다.

표 1에서 보면 공통요인모형에서 주축분해법을 사용한 8가지 사례에서 어떤 방식으로 공통분의 추정치를 구했는지 밝히지 않고 있다. SPSS에서는 기본적으로(default) SMC를 공통분의 추정치로 사용하고 있고 SAS에서는 연구자가 제공하지 않을 경우 자동적으로 1.0을 공통분의 추정치로 사용한다[9].

요인수효 결정방법

요인의 수효결정방법은 요인분석의 모형과 기초구조계산법에 따라서 달라진다. **표 1**을 보면 공통요인모형에서 주축분해법을 선택한 세 가지 논문(김애순과 윤진, 1991; 김원형, 1993; 박동건, 1992)에서 보고된 8개 사례가 있는데 이런 경우 요인의 수효결정을 위해 스크리검사, 평행성분석, 누적분산비율, 및 해석가능성을 추천할 수 있다. 김원형(1993)의 연구에서 스크리검사와 누적분산비율이 사용되고 있다. 김원형의 연구에서 사용된 "고유치>1"의 기준은 카이저방식으로서 주성분모형에서 주축분해법이 사용될 때 많이 적용되어 왔던 기준이지만 공통요인분석에서는 사용되지 않는다. 의미있는 공통요인 가운데는 고유치가 1.0보다 작은 것이 있을 수 있기 때문이다. 김애순과 윤진(1991)의 연구는 4가지 방식을 균형있게 사용함을 보이고 있다. 박동건(1992)의 연구는 6개의 공통요인분석 사례를 제시하고 있으나 多요인구조를 뽑는 경우가 아니고 1차원성을 검토하기 위해서 고유치를 사용하고 있는 경우이므로 多요인구조를 위해 요인수효결정을 하는 경우와 다른 경우이다. 이것은 이 논문의 주제가 아니며 Hattie(1985)의 연구가 좋은 지침을 제공하고 있다.

다음은 주성분모형에서 주축분해법을 사용한 경우를 보기로 한다. **표 1**에서 이 범주에 해당하는 사례는 41개이다. 이들 사례의 대부분이 원래는 공통요인분석을 했어야 했는데 주성분 분석을 하였으므로 이 경우의 사례가 이처럼 많은 것이다. 주성분모형에서 요인수효의 결정을 위해서는 스크리검사, 평행성분석, Velicer(1976)의 MAP방법, 및 카이저방식이 사용될 수 있다. 이런 계량적 방식이외에 해석가능성은 항상 고려되어야 한다. 41개의 사례에서 어떤 방식이 사용되는지는 **표 4**에 요약되어 있다. 이미 언급된 바와 같이 카이저방식은 더 이상 사용되기 어려운 정도로 신랄한 비판을 받아왔으므로 가장 약한 비중을 두어야 할 것이다. 그런데 **표 4**를 보면 총 41개 사례중 카이저방식만 사용한 사례가 25개로서 가장 많은 사례수를 차지한다. 연구 사례에서 '고유치 >1', 또는 '고유치 ≥ 1'으로 언급된 것이 카이저방식(엄밀히 말하면 고유치 > 1.0)으로 분류되었다. 스크리검사가 단독으로 쓰이지 않고 다른 방식과 함께 사용되고 있는 사례수는 3개이다. 평행성분석은 전혀 사용되지 않았으나 국내에 Hays(1987)의 비상업용 프로그램인 PARALLEL 프로그램이 보급되어 있으므로 앞으로 많이 사

9) 따라서 SAS에서 공통요인분석시에 공통분의 추정방식에 대한 기재를 명시하지 않으면 주성분분석의 解를 얻게 된다.

용될 수 있을 것이다. MAP 방법 역시 전혀 사용되지 않았으나 Reddon(1985)의 비상업용 프로그램이 보급되면 쉽게 사용될 수 있을 것이다.

해석가능성만이 사용된 경우는 오경자와 이혜련(1990)의 연구에서 보고된 3개 사례가 있다. 이 3개 사례에서 요인의 수효결정에 다음의 절차를 따랐다: 첫째, 직각회전과 사각회전 결과에서 일관성있게 유지되는 요인들을 추출; 둘째, 그 요인의 세로줄(column)에서 요인계수가 0.3이상인 문항이 여섯개 이상 있을 것. 이 두가지 모두 해석가능성을 고려한 지침이라고 할 수 있다. 그런데 첫째 절차의 경우 일관성있게 유지되는 요인들을 보기 위해서는 우선 요인들간에 상관이 있을 것인지 또는 무시할 수 있는 정도일지를 결정해서 그에 따라 사각 또는 직각회전의 실시를 결정했어야 할 것이다. 사각회전이 올바른 선택일 경우 Orthoblique와 Direct Oblimin 방식으로 돌려서 일관성 있게 추출되는 부분을 최종구조로 할 수 있을 것이다. 보통은 m개 요인이 나올 것이라고 예상될 때 m±1개의 요인수효를 주고서 결과를 구한 후 해석가능성을 비교하게 된다. 둘째 절차의 경우 Thurstone(1947)이래로 내려오는 '요인계수는 0.3-0.4 이상이고, 1개의 공통요인을 뽑으려면 3개 이상의 변수를 기초로 한다'는 전통을 좀 더 보수적으로(6개 이상 변수) 적용한 것으로 볼 수 있다.

표 4에서 보면 고유치가 2이상인 요인을 추출하는 방식은 차재호와 정지원(1993)의 연구에서 2개 사례가 보고되고 있다. 이 경우 그 논리에 대한 설명이 없어서 왜 '고유치>2'를 기준으로 했는지 알 수가 없다. 이미 Cliff (1988)가 지적했듯이 카이저방식 즉 '고유치>1'의 기준조차 흘로 설 땅이 없어지고 있는 형편이다. 즉 고유치가 1보다 작은 주성분도 필요하다면 추출될 가치가 있는 것이다. 왜냐하면 이런 주성분도 신뢰도가 양수, 즉 주성분의 분산(variance)에서 眞점수의 비율이 '영'을 상회할 수 있기 때문이다. 선행연구들의 결과에서 3개 요인이 보고되고 있음을 이유로 최미례와 김중술(1990)의 논문(2개 사례)에서도 3개 요인을 추출하기로 했다는 보고를 하고 있다. 이 경우는 이 글의 주제가 되고 있는 탐색적 요인분석보다는 확인적 요인분석에 해당하는 경우로 보인다. 즉 선행연구결과를 볼 때 3개 요인이 관례이면 현재의 연구에서도 그러한 구조가 나올 것인지 가설을 주어 검증하는 확인적 분석이 검사의 타당도에 대한 증거로서 더 유용할 것으로 보인다.

표 4에서 보면 요인수효결정의 여러 가지 방식을 복합적으로 사용한 연구는 아주 일부분임을 알 수 있다. 앞으로는 이러한 복합적 사용이 보다 권장되어 합리적인 의사결정에 보탬이 되어야 할 것이다.

요인수효결정의 경우중 마지막으로 언급되

표 4. 주성분모형에서 주축분해법의 경우 요인수효 결정방식

방 식	사례수	비 고
스크리검사만 사용	0	
평행성분석만 사용	0	
MAP방법만 사용	0	
카이저방식만 사용	25	가장 많은 사례수
해석가능성만 사용	3	오경자와 이혜련(1990)
스크리검사, 카이저방식	2	전겸구와 이민규(1992), 박병관등(1992)*
카이저방식, 해석가능성	2	최광선(1993)
스크리검사, 카이저방식, 해석가능성	1	이연희와 박광배(1992)
고유치 ≥ 2	2	차재호와 정지원(1993)
선행연구들의 결과따름	2	최미례와 김중술(1990)
미보고	4	
합 계	41	

* 박병관등의 연구에서 성격장애 척도의 경우만 스크리검사와 카이저방식을 사용하였음.

어야 할 것은 공통요인에서 최대우도법으로 기초구조를 계산하는 경우이다. 표 1에서 보면 이런 경우는 이상빈(1991) 및 이영호와 송종용(1991)의 연구뿐이다. 이경우 χ^2검증, 터커-루이스 계수, 및 해석 가능성을 사용하여 요인의 수효를 결정한다. 그런데 이상빈(1991)에서는 카이저 방식만을 사용했고 이영호와 송종용(1991)에서는 χ^2검증만을 사용했다. 카이저 방식은 공통요인 방식에서 사용되는 방식이 아니고, χ^2검증은 표본크기에 민감한 문제가 있음을 언급한 바 있다. 터커-루이스 계수는 통계패키지로써 SAS를 쓸 경우 출력에서 제시되고, SPSS를 쓸 경우 연구자가 직접계산해서 보고할 수 있다. 앞으로 이 계수의 사용이 적극 권고되어야 할 것이다.

설명분산%

추출되는 요인들이 전체분산(주성분모형에서) 또는 전체공통분산(공통요인모형에서)의 몇%를 설명하는가의 정보가 바로 설명분산%이다. 기초계산법으로서 일회의 주축분해를 했을 때는 반드시 보고되어야 하는 정보이다. 그러나 반복적인 주축분해 또는 최대우도법을 사용할 경우 추출되는 요인들이 전체분산(공통분산)의 100%를 설명하도록 짜맞춰지므로 설명분산%는 자동적으로 100%이다. 즉 보고할 필요가 없다. 그런데 분석을 위한 패키지로써 SPSS를 사용할 경우 모형이나 기초구조계산법이 어떤것이냐에 관계없이 주성분모형의 주축분해시에 준한 방법으로 '설명분산%'와 '요인분산%'를 계산해서 제공하는 오류를 가지고 있다. 표 1에서 공통요인 모형으로 분석된 연구들에서 분산%의 값들을 보면 SPSS가 사용되었고 이러한 오류 속에서 그 값들이 제공된 것임을 알 수 있다[10].

주축분해를 할 경우 분석상관행렬을 1회만 분해(factoring)하여 주축을 뽑아내면 단일주축분해이고 반복적으로 분해하여 뽑아내면 반복주축분해이다. 주성분모형일 경우 이 주축은 주성분으로 불리고, 공통요인모형에서는 공통요인으로 불린다. 주성분분석은 항상 단일분해이다. 그러나 공통요인분석에서는 단일분해를 할 수도 있고 반복분해를 할 수도 있다. 이것은 공통요인모형에서 공통분값의 비결정성 때문이다. 주성분모형의 경우 분석상관행렬에서 대각선의 값은 1.0으로 고정되어 있으나 공통요인모형에서는 추정된 공통분의 값을 원상관행렬의 대각선에 대체해 넣어 만든 축소상관행렬을 분해하여 기초해를 뽑아낸다. 그런데 기초해의 행렬 즉 기초요인행렬에서 각 변수의 가로줄에 있는 값들을 제곱하여 더하면 그 변수의 계산된 공통분값이 된다. 처음에 대입된 공통분의 값이 완벽한 추정치가 아니므로 이 추정치는 기초해에서 구한 계산된 공통분과 일치하지 않게 마련이다. 따라서 계산된 공통분을 다시 원상관행렬의 대각선에 공통분추정치로 대입한 후 그 축소상관행렬을 분해한다. 즉 2차분해를 하여 새로운 기초요인행렬을 얻는다. 이 행렬에서 구한 공통분의 값이 2차분해에서 사용된 공통분추정치와 일치하면 반복과정은 정지되지만 그렇지 않으면 계속할 수가 있다. 물론 적당한 반복수효를 설정해서 일정 횟수이상의 반복을 못하게 할 수도 있다. 이러한 반복분해에는 장단점이 있으나 解의 성질로 보아 단일분해가 더 분명한 의미를 가지므로 단일분해가 더 선호되고 있다.

연구자가 공통요인모형에서 주축분해를 하고자 할 때 통계패키지로서 SAS를 쓰게 되면 단일분해와 반복분해를 선택해서 실시할 수 있다. 그러나 SPSS를 쓸 경우 단일분해는 안되고 자동적으로 반복분해가 실시되어 解가 제공된다. 따라서 SPSS에서 공통요인모형에 주축분해법을 사용할 경우 설명분산의 %는 보고할 필요가 없다. 자동적으로 100%이다. 그러나 SAS에서 주성분분석, 공통요인분석, 또는 SPSS에서 주성분분석을 할 때는 설명분산%가 보고된다.

표 1에서 보면 SAS를 사용한 4개의 논문중 김원형(1993)과 김동기(1991)는 설명분산%의 보고가 없고 김애순과 윤진(1991) 및 이상희와 김계현(1993)의 경우에만 보고가 있다. SPSS를

10) 예로써 이상빈(1991)의 여성의류 연구사례에서 최대우도법을 사용한 경우에 설명분산%로서 47.7%를 보고하고 있다. 이경우의 설명분산%는 원래가 100%이므로 보고될 필요가 없는 것인데 이 값을 구할 때 분모를 전체 공통분산이 아닌 전체분산(여기선 25)으로 나누었으므로 100%가 아닌 47.7%가 나온 것이다. 또한 요인분산%가 1번 요인의 경우 25.1%로 보고되고 있는데 이것도 1번 요인의 분산을 전체 분산(여기선 25)으로 나누어주는 SPSS의 오류 때문에 그렇게 된 것이다.

사용했다고 보고한 논문은 박동건(1992)뿐인데 공통요인모형에 주축분해법을 사용했으므로 설명분산%는 보고되어 있지 않다. 물론 보고할 필요가 없는 경우이다. 나머지 17개의 연구에서는 어떤 소프트웨어인지 보고가 없는데 이상빈(1991) 및 이영호와 송종용(1991)의 연구에서는 최대우도법을 사용하고 있으므로 이들 만큼은 설명분산%의 보고가 필요 없다. 그런데 이상빈의 논문에서 여성의류 연구사례와 이영호와 송종용의 논문에 있는 두 사례 모두에서 설명분산%가 100%가 아닌 값으로 보고되어 있다. 이미 언급한 바와 같이 SPSS를 사용해서 얻은 오류로 생각된다.

고유치 보고

기초구조의 계산법으로서 최대우도법을 쓸 때는 고유치의 보고가 필수적인 것은 아니다. 그러나 주축분해법을 사용할 경우 요인의 수효 결정에서 고유치는 자료內의 정보중에서 각각의 주축이 뽑아내는 정보의 양으로 해석된다. 그런 맥락에서 스크리검사, 평행성분석, 누적분산비율기준등이 개발된 것이다. 따라서 요인의 수효결정에 참고될 수 있는 충분한 수효의 고유치를 보고해야 한다. **표 1**에서 보면 최대우도법을 사용한 두 논문(이상빈, 1991; 이영호와 송종용, 1991)에서는 고유치의 보고가 꼭 필요한 것은 아니다. 그러나 주축분해법을 사용했다고 보고한 16개의 논문에서는 독자들이 재검토하기에 충분한 수효의 고유치가 보고되어야 하겠으나 10개의 연구(박세영, 1991; 김원형, 1993; 김애순과 윤진, 1991; 이연희와 박광배, 1992; 박동건, 1992; 최미례와 김중술, 1990; 최윤경등, 1993; 박병관등, 1991; 전겸구와 이민규, 1992; 이상희와 김계현, 1993)에서만 고유치가 '부분적'으로 보고되어 있다. 여기서 '부분적'이라 함을 대체로 추출된 요인에 대해서만 고유치를 보고하고 있음을 의미한다. 그 논문에서 요인분석이 주된 연구방법이었고, 요인의 수효결정에 대한 논리를 논문의 심사자나 독자가 재검토할 수 있으려면 추출된 요인을 포함해서 보다 충분한 수효의 요인에 대한 고유치가 보고되어야 한다.

기초구조회전법

표 1을 기초로 하여 **표 5**를 작성하였다. 총 61개의 사례중 기초구조의 회전이 불필요하였던 연구(박동건, 1992)를 제외한 55개 사례 중에서 약 85.5%인 47개 사례에서 회전방법이 보고되었다. 보고된 47개 사례중 약 91.5%인 43개 사례에서 VARIMAX를 유일한 회전방법으로 사용하였다. 사각회전가운데 Harris와 Kaiser (1964)의 orthoblique는 SAS를 사용한 김애순과 윤진(1991)에서 보고되었다.

그러면 **표 1**에 제시된 43개의 VARIMAX회전 사례, 김애순과 윤진(1991)의 orthoblique회전, 그리고 오경자와 이혜련(1990)의 VARIMAX 및 Direct Oblimin회전에 대해서 검토해 본다. 회전하기전에 자료내 요인구조를 알아보기 위해 기초행렬을 참고하여 PREPLOT을 해 볼 수 있고 회전후 변수 모임이 요인축상에 잘 수용되는지를 보기 위해서 PLOT을 해보는 것에 대해서 앞서 언급한 바 있다. 우선 43개의 VARIMAX회전 사례에서, 과연 자료내 요인들이 직각구조를 보이는지에 대해서 논문들에서

표 5. 기초구조의 회전에 사용된 방법

		사례수	비 고
직각회전	VARIMAX만 사용	43	
사각회전	Oblimin만 사용	0	
	Orthoblique만 사용	1	김애순과 윤진(1991)
VARIMAX와 Oblimin		3	오경자와 이혜련(1990)
미보고		8	
보고 불필요		6	박동건(1992)
합 계		61	

검토를 보이고 있지는 않다. 회전방식의 결정은 연구자가 자료의 성질을 충분히 이해한 후에 결정하며 회전후 PLOT을 통해 변수모임들이 요인축들에 충분히 수용되는지를 조사하므로서 연구자도 자신감을 가지고 그 구조를 제시할 수 있는 것인데 그에 대한 논의가 전혀 제시되지 않고 있다. 또한 김애순과 윤진의 사각회전이나 오경자와 이혜련의 사각회전에서도 파라메터값이 얼마로 주어졌는지, 회전후 PLOT을 볼 때 요인들이 변수모임들을 얼마나 잘 반영하는지에 대한 언급이 없다. 또한 사각회전시에는 요인간 상관행렬이 주어져야 하는데 오경자와 이혜련의 연구에서는 그에 대한 정보가 제공되어 있지 않다. 오경자와 이혜련의 연구에서 VARIMAX와 Direct Oblimin을 실시하였으나 요인추출은 VARIMAX의 결과에서 실시하였다. 그런데 각 요인을 나타내는 척도간의 상관이 무시할 수 없는 정도의 큰 값으로서 별도의 표에 제시되고 있음은, 실제로 요인구조가 직각구조가 아닌 사각구조임을 스스로 입증하는 것이다. 즉 요인추출을 VARIMAX의 결과에서 하고 있는 것이 오류임을 나타내고 있다.

다시 VARIMAX에 대한 비평으로 돌아가기로 한다. 저자는 VARIMAX를 사용한 14개의 논문중 최종요인행렬을 온전하게 제공하고 있는 8개의 연구에 대해서 사각회전을 적용해 봄으로서 과연 그들이 정말로 직각구조인지 검토해 보고자 한다. 기초解와 회전후의 解와는 수학적으로 동등한 입장에 있는, 일종의 동치모델(equivalent model)이다. 즉 기초요인행렬과 VARIMAX회전에 의해서 나온 최종요인행렬은 서로 순서가 바뀌어서 제시될 수도 있는 관계에 있다. 그러므로 VARIMAX회전후 산출된 최종요인행렬에 사각회전을 적용함은 기초요인행렬에 사각회전을 적용함과 같은 효과를 가진다. 저자가 사각회전(이미 구해진 factor pattern을 주고 사각회전하는 경우 SAS의 PROMAX가 이용가능함)을 적용한 연구 및 최종요인 행렬들은 **표 6**과 같다.

우선 PROMAX의 기능을 보기 위해서 세개의 요인으로 구성된 직각구조와 사각구조(세 요인간 상관을 0.3, 0.4, 0.5로 하였음)에 대응하는 12개 측정변수간의 상관행렬(모집단행렬)을 분석해 보았다. 분석은 두 차례에 걸쳐 이루어졌다. 첫번째 분석으로서 직각구조에 대응하는 상관행렬을 공통요인모형에서 단일주축분해하여 얻은 기초요인행렬에 PROMAX회전을 실시했으나 요인간 상관은 모두 "0"을 보이는 직각구조를 보여주었다. 즉, 기초구조에 PROMAX 회전을 적용시켜도 원자료가 직각구조에 대응하고 있으므로 사각구조가 나오지는 않았다. 이것은 직각구조에 사각회전을 적용한다고 반드시 요인간 상관이 만들어지는 것은 아니라는 Tabachnik와 Feidell(1983)의 주장에서 예측되는 결과이다.

다음은 두번째 분석으로서 사각구조에 대응하는 상관행렬을 공통요인모형에서 단일주축분해하고 일단은 VARIMAX회전을 하여 최종요인행렬 및 PLOT을 얻었다. 이 PLOT의 모양을 보면 축과 축사이에 변수들의 모임이 산포되어 있어서 상관행렬내의 요인구조가 직각구조가 아님을 의미하였다. 따라서 최종요인행렬에 사각회전으로서 PROMAX회전을 실시하고

표 6. VARIMAX회전 후의 최종요인행렬이 제시된 연구들

연 구	PROMAX가 적용된 최종요인행렬
박병관 등(1991)	p.60, 표 3
이상빈(1991)	p.125, 표 5
이상희, 김계현(1993)	p.36, 표 1의 상담자에 의한 상담평가 부분
이영호, 송종용(1991)	p.105, 표 7
임영란, 안창일(1992)	p.49, 표 3의 '정상인 집단' 부분
장재윤(1993)	p.119, 표 3의 '시나리오1' 부분
전겸구, 이민규(1992)	p.70, 표 2의 '한국' 부분
최미례, 김중술(1990)	p.199, 표 4의 '정신병' 부분

PLOT을 얻었다. 그 결과로 요인간 상관행렬은 원래의 요인간 상관을 그대로 재현하였고 PLOT에서는 변수의 모임들이 각 요인축상에 가지런히 위치함을 보여주었다. 즉 원래가 사각구조일 경우, VARIMAX로 회전해서 얻은 최종요인행렬에 사각회전을 실시하면 올바른 요인구조를 얻을 수 있음을 나타낸다.

표 6에 있는 최종요인행렬들 가운데 박병관 등(1991)의 최종요인행렬에 PROMAX회전을 시킨 프로그램 예를 **표 7**에 소개한다.

또한 직각구조인 8개의 최종요인행렬을 PROMAX방식으로 사각회전시킨 결과로 얻은 요인상관행렬을 **표 8**에 제시한다.

표 8을 보면 장재윤의 자료에서만이 요인간 상관이 영에 가까워서 VARIMAX회전이 타당했음을 보여주고 있고 나머지의 자료에서는 양수 또는 음수의 크거나 작은 요인간 상관을 보여주고 있다. **표 6**의 자료들이 표본자료에 기초한 것이고 저자가 실시한 PROMAX회전 역시 다른 회전방법과 마찬가지로 '우발성의 이론화'를 범할 수 있는 가능성을 가지고 있으므로 **표 8**의 요인간 상관이 실제의 상관보다 어느 정도 다를 수 있음을 감안해야 한다. 그럼에도 불구하고 이 결과는 연구자들에게, 자료의 성질을 고려하지 않은 편리성 위주의 VARIMAX회전이 올바른 선택은 아니었음을 보여주기에 충분하다. 참고로 박병관등의 최종요인행렬을 2차원공간에 점찍어 본 **그림 1**에, 저자가 PROMAX로 돌린 결과를 점찍은 것은 **그림 2**에 제시되어 있다.

표 7. 박병관등(1991)의 최종요인행렬에 적용한 PROMAX회전 프로그램

```
/* VARIMAX SOLUTION IS READ AS DATA */
OPTIONS LS=75;
DATA OLD(TYPE=FACTOR);
INPUT _TYPE_ $ _NAME_ $ v1-v11 ;
CARDS;
PATTERN FACTOR1  .16  .09  .88  .77 .38 .84 .60 .88 .77 .30  .25
PATTERN FACTOR2 -.93 -.79 -.03 -.05 .36 .03 .43 .19 .51 .79  .93
PATTERN FACTOR3 -.23 -.53  .14 -.03 .80 .22 .56 .22 .25 .45 -.02
;
RUN;
PROC FACTOR preplot R=PROMAX PRE=NONE plot ;
RUN;
```

표 8. "표 6"의 최종요인행렬을 사각회전한 결과의 요인상관행렬

박병관등	이상빈	이상희, 김계현	이영호, 송종용
1.0	1.0	1.0	1.0
.21 1.0	.03 1.0	.34 1.0	.49 1.0
.40 .56 1.0	.43 .27 1.0		.49 .45 1.0
	.15 .28 .34 1.0		

임영란, 안창일	장재윤	전겸구, 이민규	최미례, 김중술
1.0	1.0	1.0	1.0
.47 1.0	.06 1.0	-.33 1.0	.50 1.0
		.48 -.21 1.0	.51 .49 1.0
		.32 -.30 .23 1.0	

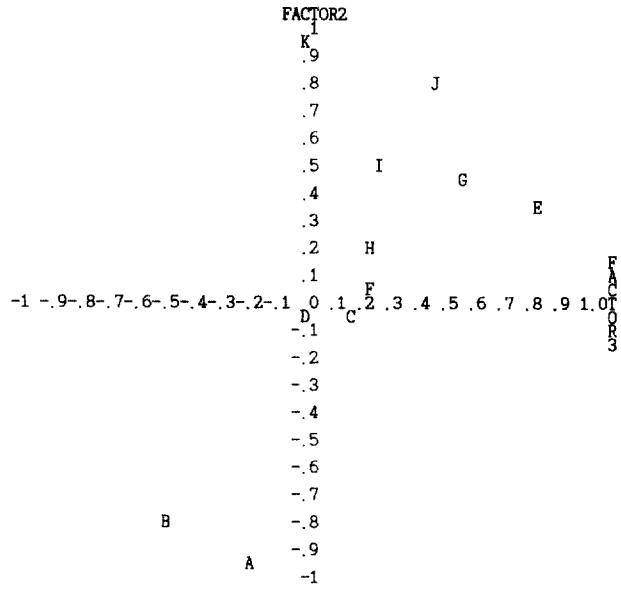


그림 1. 박병관 등(1991)의 최종요인행렬에서 요인 2와 요인 3에 대한 PLOT

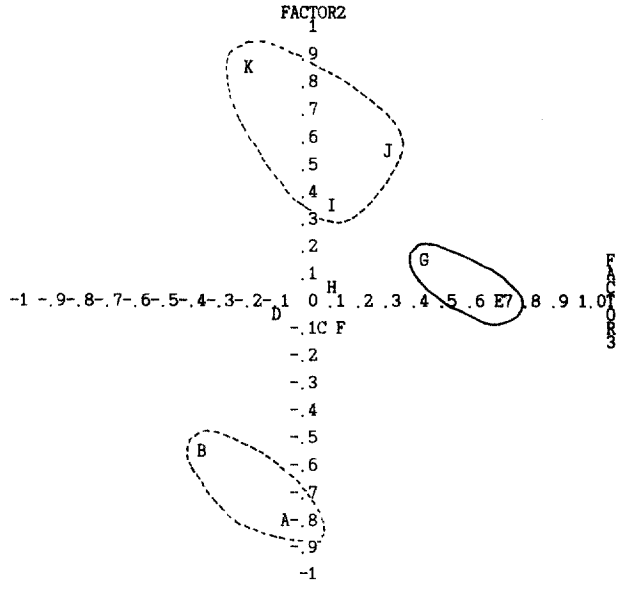


그림 2. PROMAX 회전후의 PLOT

그림 1은 박병관 등이 VARIMAX 회전 후 제시한 최종요인 행렬에서 요인2와 요인3의 부분만을 요인공간(factor space)으로 표시한 것이다. 즉, 요인2와 요인3의 2차원 평면에 11개의 변수(박병관 등의 글에 있는 변수들을 처음부터 끝까지 A, B, · · · K로 지정 함)를 점찍어 본 것이다. 원점을 중심으로 C, D, F 및 H가 있는데 이들은 요인2나 요인3이 아닌 다른 요인- 여기서는 나머지 요인이 요인1 밖에 없음 - 을 나타내는 변수들이다. 1 상한에 원점으로부터 좀 떨어진 곳에 또 한무리의 점들 - E, G, I, J - 이 있는데 이들은 요인2와 요인3을 나타내긴 하되 이 두 요인이 상관되어 있으므로 현재의 VARIMAX 회전만으로는 어떤 요인이 어떤 변수의 모임에 근접하는지를 나타낼 수 없는 경우이다. 만일에 요인2의 축을 오른쪽으로 약간, 요인3의 축을 왼쪽으로 충분히 회전한다면 각 요인축이 어떤 변수들을 요약하는가를 좀더 분명히 나타낼 수 있는데 이처럼 요인 축간의 직각구조를 깨트리게 되면 사각구조가 되고 요인간 상관이 산출된다. 그림 2는 PROMAX회전(Hendrickson & White, 1964)을 한 후의 PLOT이다. 원점을 중앙으로 C, D, F, H가 몰려 있는 것은 그림 1에서와 같다. 여기서도 요인공간은 직각좌표상에 표시되고 있다. 그러나 이것은 요인간 상관이 0.56(표 8 참조)인 경우의 요인공간을 편의상 직각좌표에 나타냈을 뿐이다. 여기서 요인2의 축에 가까이 A, B, I, J, K가 있다. 또 요인3의 축에 가까이 E와 G가 있다. 박병관등은 VARIMAX 회전 후 요인2를 내-외향성, 요인3을 의존성으로 해석하였는데, 사각회전을 하고 보니 요인2와 요인3의 상관이 0.56이나 되고, 요인3은 겨우 2개의 변수밖에 요약하지 못하고 있다. 이경우는, 원래 요인2와 요인3이 하나의 요인인데 분리된 요인으로 취급되어 3개로 요인구조를 추출한 것이 아닌가 하는 의문을 가져볼 수 있다. 즉, 요인수효결정에 있어서 '해석 가능성'의 지침을 적용해 볼 수 있는 경우라고 본다.

최종요인행렬

최종요인행렬은 요인구조의 회전방식에 따라 다르게 제공된다. 직각회전의 경우 최종요인행렬은 한가지뿐이다. 요인의 해석은 이 행렬을 참고하여 이루어진다. 그러나 사각회전의 경우 최종요인행렬로서 형태계수행렬과 구조계수행렬이 별도로 계산된다. 또한 이 두 최종요인행렬과 더불어 요인간 상관행렬이 계산된다. 사각회전에서는 구조계수행렬과 요인간 상관행렬을 요인해석시에 주된 자료로, 형태계수행렬을 보조자료로 참고하게 된다. 그러므로 회전방식에 관계없이 최종요인행렬 및 요인간 상관행렬(사각회전시)을 온전히 제공하는 것이 심사자나 독자에게 요인해석의 결과를 검토할 수 있게 해 준다. 표 1에서 보면 총 61개 사례중 17사례에서만 온전하게 보고하고 있다.

요인 해석 방침

요인해석에 사용되는 최종요인행렬은 직각회전시에는 하나이고 사각회전시에는 두개다. 즉 형태계수행렬과 구조계수행렬이다. 형태계수는 일종의 표준회귀계수이고, 구조계수는 일종의 상관계수이다. 사각회전시 요인해석을 위해서는 구조계수가 우선적으로 사용된다. 직각회전시의 최종요인행렬은 형태계수행렬인 동시에 구조계수행렬이다. 이것은 요인간 상관이 영인 경우라서 표준회귀계수와 상관계수가 같아지기 때문이다. 따라서 요인의 해석을 위해 어떤 회전이냐에 관계없이 '구조계수'가 사용된다해도 일반성을 상실하지 않는다. 구조계수는 일종의 상관계수이므로 전통적으로 0.3 - 0.4 이상을 해석의 대상으로 해 온 것이다. 그러나 요인해석에서 이 값에 너무 기계적으로 얽매이지 말고, 요인의 내용에 대한 확실한 이해를 바탕으로 보다 적절한 해석을 해야 할 것이다. 표 1에서 보면 22개의 논문 중 11개에서 즉, 50%정도에서 요인해석 방침을 제시하고 있다. 대체로 요인계수의 값이 0.3이나 0.4를 기준으로 하는 전통에서 크게 벗어나지 않으면서 요인계수의 상대적 크기를 고려하는 융통성을 보이고 있다. 그런데 하나의 공통요인을 해석하기 위해서는 적어도 3개 이상의 변수가 큰 요인계수를 보이고 있어야 하는 것이 오래된 전통이다. 그러나 척도개발 내지 검증을 할 경우라면 하나의 척도가 하나의 요인을 정의할 수 있다. 사회과학에서 하나의 척도라면 최소한

문항의 수효가 5개 이상은 되어야 신뢰도와 내용타당도 확보에 도움이 될 것이다. 따라서 요인분석에서의 측정변수가 개별 문항이라면 요인계수가 .3 - .4 이상인 문항이 5개 이상은 되어야 한다는 요구가 무리한 것은 아니다.

요인분산%

각 요인의 중요성은 요인의 분산이 전체분산 또는 전체공통분산 가운데 몇 %나 차지하는 가로 표시된다. 주성분모형의 경우 전체분산은 측정변수의 수효와 같은 값이다. 공통요인모형의 경우 단일주축분해에서는 축소상관행렬의 고유치의 合이 전체공통분산이다. 그러나 반복주축분해나 최대우도법(여기서도 반복알고리즘이 있음)에서는 반복과정이 끝난 상태에서 얻어진 공통분총합(각 측정변수의 공통분을 모두 합한 것)이 전체공통분산이다. 그런데 이러한 요인분산%는 직각회전일 때에 구하기가 쉽다. 사각회전일 경우 요인간에 상관이 있으므로, 각 요인이 독자적으로 설명하는 분산(독자분산)과 다른 요인들과 공동으로 설명하는 분산(공유분산)으로 나누어 보고해야 할 것이다. 어느 요인의 독자분산은 모든 다른 요인들의 설명력을 감안한, 즉 통제한(partialled out) 후 그 요인에 의해 독특하게 설명되는 분산이다. 표 1에서 보면, 공통요인모형에서는 요인분산%의 보고가 전혀 되어있지 않다. 주성분모형의 경우 거의 대부분의 연구에서 요인분산%를 보고하고 있다. 앞서 '설명분산%'의 부분에서 언급했듯이 소프트웨어로서 SPSS를 쓸 경우 공통요인모형인데도 불구하고 요인분산%의 계산에서 전체공통분산이 아닌 전체분산을 분모로 하여 계산된 결과가 제시되는 오류가 있음을 지적해 둔다. 표 1에서 이상빈(1991)과, 이병호와 송종용(1991)에서는 이렇게 하여 잘못 계산된 요인분산%가 보고되고 있다.

요인점수계산방식

요인점수계산방식에는 가장 많이 쓰이는 회귀방식을 비롯하여 Bartlett방식, Anderson-Rubin방식, 및 이상적 변수방식이 있어 각기 장단점을 지니고 있다. 표 1에서 요인점수를 계산한 경우는 그리 흔하지 않다. 주성분모형일 경우에는 요인점수 계산방식으로서 어떤 것을 쓰느냐에 무관하게 같은 값(주성분점수라고 함)이 결정되어 나오므로 실제로 계산방식을 보고할 필요가 없다. 전현민과 노명래(1991) 및 차재호와 정지원(1993)이 분명히 밝히지는 않았지만 이런 경우로 추정되므로 해당되는 칸에 (결정)이라고 기재하였다. 김동기(1991)에서 요인점수를 계산했다고 보고하고 있는데 그가 사용한 SAS에서는 회귀방식밖에 없다. 그러나 요인분석모형이 공통요인모형일 때나 회귀방식으로 구했다는 보고를 할 필요가 있을 것이다. 최광선(1993)에서는 주성분분석을 하면서, 요인행렬에 있는 각 요인의 세로줄에서 요인계수가 큰 문항들에 대해 응답자가 보이는 점수를 합하여 그 요인의 점수로 계산했음을 보고하고 있다. 그러나 이 점수는 실은 요인점수가 아니라 단지 그 요인을 재는 척도의 점수라고 하는 것이 정확하다. 이러한 척도점수에는 眞점수와 오차점수가 함께 들어 있다. 주성분분석의 경우에는 주성분점수가 수학적으로도 산뜻하게 결정되므로 굳이 최광선(1993)의 방법을 권할 이유는 없을 것으로 보인다.

결 론

요인분석의 개발은 이미 60년대말에 대부분의 기술적인 문제가 해결되고 그 이후로 사용에 있어서의 기본 절차들이 정착되어 왔다. 심리학계의 학술지에 발표된 요인분석의 응용논문들을 볼 때 요인분석의 두 가지 모형 가운데 주성분모형의 사용이 아직도 공통요인모형의 사용을 압도하고 있음은 심리측정 분야가 응용연구자들에게 충분히 서비스되고 있지 못함으로 이해된다. 주성분분석은 이미부터 통계학에 출발점을 두고 보급 및 활용되어 왔으나, 공통요인분석의 경우 심리학에서의 개념구성(예 : Spearman의 일반지능 정의)을 하기 위한 수단으로써 개발되어오면서 행동자료의 특성을 이해하지 않고는 충분한 사용과 해석을 하기 어려운 부분들이 많이 첨가되었다. 따라서 공통요인분석을 학습하는 것은 주성분분석을 학습하는 것 보다 훨씬 더 많은 '행동자료에 대한

사회과학적 이해'를 필요로 한다. 분석시에 주성분모형과 공통요인모형에 대한 올바른 선택이 이루어진 후의 절차는 대체로 공통요인모형일 때가 더 복잡성을 띄게 된다. 우선 공통요인모형일 경우 국내에서 가장 널리 쓰이는 통계패키지인 SPSS를 연구자들이 사용할 때마다 본인이 손으로 계산해서 수정해야할 사항들이 많다. 요인의 수효결정은 어떤 모형이 선택되느냐에 관계없이 신중을 기해야 하는 단계이다. 여러 가지 방법을 모두 사용해 가면서 가장 적절한 의사결정을 내려야 하는데 아직은 그러한 바람직한 관행이 자리잡혀 있지 않다. 특히 주성분모형일 경우 카이저방식에 지나치게 의존하고 있는 현상을 시급히 벗어나야 한다. 이미 80년대 말까지에는 카이저방식에 대한 신랄한 비평이 쏟아졌고 거기 대해서 Kaiser 본인은 물론 다른 어느 누구에 의해서도 카이저방식은 옹호되지 않았다. 이제 주성분모형에서 스크리검사, 평행성분석, MAP방식, 및 해석가능성이 요인수효 결정에 주된 방식으로 이용되어야 할 것이다.

일단 기초구조가 산출되면 해석을 하기 위해서 회전을 시켜야 한다. 요인회전에 있어서 직각회전이냐 사각회전이냐에 대한 선택은 연구자가 편리를 위해 자의로 선택하는 것이 아니다. 자료의 성질 및 요인 분석이 사용되는 목적을 이해하면 어떤 회전이 적절한지 결정될 수 있다. 직각구조가 들어 있는 자료에 사각회전을 적용할 경우 요인간 상관이 반드시 산출되는 것은 아니다. 표본자료에서는 어떤 통계절차에서건 '우발성의 이론화'가 존재할 수 있다. 단지 연구자가 내용에 대한 이해를 가지고 그러한 현상을 할인해서 해석할 수 있어야 한다. 한편 사각구조가 들어 있는 자료에 직각회전을 할 경우 요인구조가 분명하게 들어나지 않으므로 많은 변수들이 여러개의 요인에서 큰 계수를 보이면서 요인의 의미해석에 어려움을 제공할 수 있다.

표 5에서 보면 국내 학술지에서 사각회전을 보고하는 경우는 극히 드물고 거의가 직각회전을 보고하고 있다. 요인들을 추출하여 그 요인들, 즉 이론적인 변수들 간의 네트워크를 만드는 것이 이론구조를 엮어 가는 방식이라면 그런 맥락에서 추출되는 요인들 간에 상관이 영이 될 수가 없다. 또한 관계가 없는 요인들을 사용하여 이론구조를 엮는다는 것은 불가능하다. 사각회전이 직각회전에 비해 연구자의 지식을 많이 필요로 하긴 하지만 **표 8**에서 제시된 요인간 상관들을 볼 때 사각회전이 이론 개발을 위해 제공하는 의미는 지대하다.

끝으로 이 글을 마무리하면서 가장 강조해야할 것은 학술지에 제출된 논문에 대해 심사자나 독자들이 재분석을 해볼 수 있도록 저자들은 충분한 정보를 제공해야 한다. 요인분석의 경우 그러한 정보중에서도 가장 중요한 것은 바로 상관 행렬이다. 이 행렬을 보고한 논문은 **표 1**에서와 같이 전무하다. 이 행렬이 제시되지 않고는 논문 제출자의 분석결과에 대한 확인 및 조언이 따를 수가 없다

참 고 문 헌

김동기(1991). 종교 정향과 사회현상의 부적 결과에 대한 귀인의 연구: 기독교인을 중심으로. **한국심리학회지: 사회, 6,** 22-31.

김애순, 윤진(1991). 중년기 위기감 척도 제작에 관한 연구. **한국심리학회지: 발달, 4,** 73-87.

김원형(1993). 조직 동일시, 조직 몰입, 조직 내 재화 변인들간의 탐색적 요인분석 및 확인적 요인분석에 대한 연구. **한국심리학회지: 산업 및 조직, 6,** 12-35.

김준곤, 성한기, 이춘희, 박경자(1991). 대학생활 만족의 구성요인과 예언요인. **한국심리학회지: 산업 및 조직, 4,** 154-167.

박동건(1992). 직무만족 측정의 동질성. **한국심리학회지: 일반, 11,** 129-146.

박병관, 김진영, 노주선, 안창일, 신동균, 박영숙(1991). MMPI 성격장애 척도와 16PF 검사구조분석: 성격장애 척도 타당화 예비연구. **한국심리학회지: 임상, 10,** 55-75.

박세영(1991). 대우의 형평 지각을 결정짓는 요인들. **한국심리학회지: 산업 및 조직, 4,** 22-44.

신현균, 원호택(1991). 비합리적 사고와 귀인양식에 있어서 우울집단과 불안집단 비교. **한국심리학회지: 임상, 10,** 269-285.

오경자, 이혜련(1990). 한국 아동 청소년의 문

제 행동 증후군 연구: 4-5세, 6-11세, 12-16세 남아 CBCL 자료의 분석). **한국심리학회지: 임상, 9,** 33-35.

이상빈(1991). 소비자 관여의 개념화와 측정. **한국심리학회지: 산업 및 조직, 4,** 110-145.

이상희, 김계현(1993). 상담회기 평가 질문지 (Session Evaluation Questionaire)의 타당화 연구. **한국심리학회지: 상담과 심리치료, 5,** 30-47.

이연희, 박광배(1992). 심리척도에 의한 성격집단 분류에서 예측되는 분류 오류의 확률: 항목 반응 이론의 응용. **한국심리학회지: 일반, 11,** 24-40.

이영호, 송종용(1991). BDI, SDS, MMPI-D 척도의 신뢰도 및 타당도에 대한 연구. **한국심리학회지: 임상, 10,** 98-113.

임영란, 안창일(1992). MMPI 내용척도의 타당화 연구. **한국심리학회지: 임상, 11,** 42-64.

장재윤(1993). 환경 변화 지각에 있어서의 개인차 및 친숙도 효과. **한국심리학회지: 산업 및 조직, 6,** 108-128.

전겸구, 김교헌(1991). 대학생용 생활 스트레스 척도의 개발: 제어 이론적 접근. **한국심리학회지: 임상, 10,** 137-158.

전겸구, 이민규(1992). 한국판 CES-D 개발연구. **한국심리학회지: 임상, 11,** 65-76.

전현민, 노명래(1991). 자아 개념과 정신건강과의 관계에 대한 탐색적 조망. **한국심리학회지: 임상, 10,** 114-125.

차재호, 정지원(1993). 현대 한국 사회에서의 집합주의. **한국심리학회지: 사회, 7,** 150-163.

최광선(1993). 집단행동 제어에 미치는 행동 조정 및 정서 불안 감소 방략의 효과. **한국심리학회지: 사회, 7,** 164-174.

최미례, 김중술(1990). 정신병 환자의 KWIS분산도 분석 - 신경증 환자와의 비교. **한국심리학회지: 임상, 9,** 192-205.

최윤경, 임영란, 박병관, 안창일, 최정윤, 김재환(1993). MMPI 내용척도와 성격장애 척도의 구조분석. **한국심리학회지: 임상, 12,** 1-15.

Allen, S, J., & Hubbard, R. (1986). Regression equations for the latent roots of random data correlation matrices with unities on the diagonal. *Multivariate Behavioral Research, 21,* 393-398.

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, 3,* 77-85.

Bartlett, M. S. (1951). A further note on tests of significance on factor analysis. *British Journal of Psychology, 4,* 1-2.

Brogden, H. E. (1969). Pattern, structure,and the interpretation of factors. *Psychological Bulletin, 72,* 375-378.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Researh, 1,* 245-276.

Cattell, R. B., & Dickman, K. (1962). A dynamic model of physical influences demonstrating the necessity of obligue simple structure. *Psychological Bulletin, 59,* 389-400.

Cattell, R. B., & Gorsuch, R. L. (1963). "The uniqueness and significance of simple structure demonstrated by contrasting organic 'natural structure' and 'random structure' data." *Psychometrika, 28,* 55-67.

Cliff, N. (1988). The egenvalues-greater-than-one rule and the reliability of components. *Psychological Bulletin, 103,* 276-279.

Dwyer, P. S. (1939). The contribution of an orthogonal multipe factor solution to multiple correlation. *Psychometrika, 4,* 163-171.

Gorsuch, R. L. (1983). Factor analysis (2nd Ed,). Hillsdale, NJ: Lawrence Erlbaum.

Guilford, J. P. (1981). Higher-order structure-of-intellect abilities. *Multivariate Behavioral Research, 16,* 411-435.

Guttman, L. (1954). Some necessary and sufficient conditions for common factor analysis. *Psychometricka, 19,* 149-161.

Hariss, C. W., & Kaiser, H. F. (1964). Oblique factor analytic solutions by orthogonal transformations. *Psychometrika,*

29, 347-362.

Hattie, J. (1985). Methodology review: assessing unidimensionality of tests and items. *Applied Psychological Measurement, 9*, 139-164.

Hays, R. D. (1987). PARALLEL: A Program for performing parallel analysis. *Applied Psychological Measurement, 11*, 58.

Hendrickson, A. E., & White, P. O. (1964). PROMAX: A quick method for rotation to oblique simple structure. *The British Journal of Statistical Psychology, 17*, 65-70.

Humphreys, L. G., Ilgen, D., McGrath, D., & Montanelli, R. (1969). Capitalization on chance in rotation of factors. *Educational and Psychological Measurement, 29*, 259-271.

Jackson, D. J., & Tweed, D. L. (1980). Note on the squared multiple correlation as a lower bound to communality. *Psychometrika, 45*, 281-284.

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika, 31*, 313-323.

Joreskog, K. G., & Lawley, D. N. (1968). New methods in maximum likelihood factor analysis. *British Journal of Statistical Psychology, 21*, 85-96.

Kaiser, H. F. (1960). The application of electrionic computers to factor analysis. *Educational and Psychological Measurement, 20*, 141-151.

Lawley, D. N., & Maxwell, A. E. (1971). *Factor analysis as a statistical method.* New York: Macmillan Publishing Co., Inc.

Loo, R. (1979). The orthogonal rotation of factors in clinical research: A critical note. *Journal of Clinical Psychology, 35*, 762-765.

MacCallum, R. (1983). A comparison of factor analysis program in SPSS, BMDP, AND SAS. *Psychometrika, 48*, 223-231.

Montanelli, R. G. Jr., & Humphreys, L. G. (1976). Latent roots of ramdom data correlation matrices with squared mutiple correlation on the diagonal: A Monte Carlo Study. *Psychometrica, 41*, 341-347.

Nunnally, J. C. (1978). *Psychometric Theory* (2nd ed.). New York: Graw-Hill.

Reddon, J. R. (1985). MAPF and MAPS: Subroutiness for the number of principal components. *Applied Psychological Measurement, 9*, 96.

Tabachnik, B. G., & Fidell, L. S. (1989). *Using mulivariate statistics.* San Francisco: Harper & Row.

Thurstone, L. L. (1947). Mulitiple factor analysis. Chicago: University of Chicago Press.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*, 1-10.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlation. *Psychometrika, 41*, 321-327.

Widaman, K. F. (1993). Common factor analysis versus principal component analysis: Differential bias in representing model parameters. *Multivariate Behavioral Research, 28*, 263-311.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin, 99*, 432-442.

1차 접수 : 1994년 11월 12일
수정본 접수: 1995년 2월 17일

韓國心理學會誌: 産業 및 組織
Korean Journal of Industrial and Organizational Psychology
1994. Vol. 7, No. 1.

Practice of factor analysis in Korean academic journals

Soon-Mook Lee

Sung-Kyun-Kwan University

It is necessary and required for an author to provide information for readers to evaluate the quality of his/her paper submitted to a journal for publication. In this article, factor analysis papers published in Korean academic journals of Psychology are reviewed to examine if proper information has been provided. It is uncovered that authors using factor analysis do not provide enough information for the reviewers and the readers of their research results. It is concluded that review practice should be tightened more than now in order for the articles to communicate and accumulate academic findings.