특집논문

탐색적 요인분석의 오·남용 문제와 교정*

Misuse of Exploratory Factor Analysis and Its Remedies

김청택[a)]
Cheongtag Kim

요인분석은 사회과학에서 널리 사용되는 기법이기도 하지만 연구자들이 분석과정에서 필요한 결정들을 빈번하게 잘못내리는 기법이기도 하다. 본 논문에서는 요인분석을 수행할 때 어떠한 결정들이 필요하고 어떻게 결정하는 것이 바람직한지에 대하여 논의하였다. 특히 요인분석의 잘못된 적용이 대부분 요인분석을 주성분분석의 틀 속에서 이해하기 때문이라는 관점을 취한다. 따라서 이 두 방법의 통계적 차이점을 살펴보고 주성분분석의 틀에서 요인분석을 해석할 때 어떠한 문제가 있는지를 분석하였다. 이 논문은 또한 자료 적합방식, 적합도 지수, 요인의 수의 결정방법, 결측치와 범주변수들의 처리에 대한 이슈들도 논의하였다.

주제어: 탐색적 요인분석, 주성분분석, 구성체, 설명분산, 요인의 수, 요인회전 방법

Though the factor analysis is widely used in social science research, researchers make frequently unreasonable decisions when conducting these analyses. The study reviews and recommends which decisions must be made when conducting factor analyses.

* 이 논문은 2015년도 정부(미래창조과학부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업임(NRF-2015R1A5A7037508).

a) 서울대학교 심리학과 교수 김청택.
E-mail: ctkim@snu.ac.kr

This study contends that one of the main reasons of the misuse is because researchers understand exploratory factor analysis(EFA) in the frame of principal component analysis(PCA). The paper provides methodological and statistical comparisons between two methods and the consequences of interpreting EFA results in the frame of PCA. This article also address the issues of data fitting method, fit indices, choosing number of factors, and treating missing data and categorical data.

Key words: exploratory factor analysis, principal component analysis, explained variance, number of factors, factor rotation methods

Ⅰ. 서론

요인분석은 심리학, 교육학, 경영학, 사회학 등의 분야에서 설문지나 검사지를 이용하여 개인이나 집단의 특성을 연구하는 데 널리 사용되는 자료분석기법 중의 하나이다. 요인분석모형에서는 공통요인들과 측정된 변수들 간의 관계를 반영하는 요인부하량과 공통요인에 의하여 설명되지 않는 요인, 즉 고유요인으로 측정변수 간의 공분산(혹은 상관)을 예측한다. 요인들은 관찰될 수 없는 변수이기 때문에 잠재변수라 불린다.

연구자들은 이 분석법을 이용하여 문항들에 대한 반응을 유발하는 가설적 구성체(construct) 혹은 요인(factor)들을 찾아내고자 한다. 설문지를 구성하는 문항들이 몇 개의 요인에 의하여 설명될 수 있는지, 각각의 문항들이 어떠한 요인들과 연관되어 있는지에 대하여 판단하게 된다. 이 분석법은 t-검증이나 분산분석처럼 주어진 자료에 일정한 공식을 대입하여 결과를 내는 것이 아니고, 통계적 모형을 만들고 자료를 이에 적합시켜서 해를 구하게 된다. 따라서

분석에서 사용된 자료 적합방법이나 회전방법식 등에 따라 상이한 결과를 내놓기 때문에 통계적인 모형을 정확하게 이해하여 적용하지 않으면 잘못된 결론을 내리기 쉽다. 이 논문에서는 요인분석의 각 단계에서 내려야 될 결정을 살펴보고, 이때 합리적인 결정이 무엇인지에 대하여 논의할 것이다.

탐색적 요인분석을 이용하는 연구자에 의해 결정되어야 할 요소들은 다음과 같다. 어떠한 자료 적합방법(data fitting method)을 사용할 것인지, 어떤 기준에 의하여 적절한 요인 수를 결정할 것인지, 어떠한 회전방법을 사용할 것인지, 그리고 어떤 모형 적합도를 사용할 것인지 등이 있다. 연구자들이 흔히 잘못하는 선택은 주성분분석(principal component analysis)을 사용하고, 1보다 큰 고유치가 몇 개인지로 요인의 수를 결정하고, 직교회전인 VARIMAX 방법으로 회전하고, 소위 설명분산(explained variance)으로 요인분석모형의 적합도를 판단하는 것이다. 이에 대한 문제점은 이순묵(1994), Floyd & Widaman (1995), Fabrigar et al.(1999), Preacher & MacCallum(2003) 등에 의하여 꾸준히 지적되었음에도 불구하고 많은 연구자들은 여전히 위에서 제시된 방법으로 요인분석을 적용하고 있다.

이순묵(1994)에 따르면, 1990년에서 1993년 사이에 심리학 관련 학회지에 보고된 요인분석 중 90%가 넘게 주성분분석을 하였고 그 틀에서 결과를 해석하였다. 15년이 지난 2007년도의 한국심리학회지: 사회와 한국심리학회지: 산업 및 조직에 게재된 논문들에 대한 조사에 따르면 이러한 문제점은 개선되기는 하였지만, 여전히 오용의 사례는 남아있다(김청택 2008). 탐색적 요인분석을 사용한 11편의 논문 중 3편은 여전히 주성분분석을 사용하였다. 또한 요인분석이나 주성분분석에 상관없이 모형의 양호도는 설명분산에 의존하고 있었으며, 모든 논문에서 고유치에 의해 계산된 개별 요인의 설명분산을 회전 후에도 적용시키는 오류를 저지르고 있었다.

이러한 오용의 이유는 여러 가지가 있겠지만, 사용자들이 요인분석기법을 정확하게 이해하지 못한 채 통계패키지가 제공하는 기본선택을 맹목적으로 따르기 때문으로 보인다. 이 논문에서는 오사용의 주원인이 요인분석을 주성분분석의 틀 속에서 이해하고 해석하기 때문이라는 점에 초점을 맞추어 논의를 진행할 것이다. 먼저 요인분석의 논리를 개관하고 많이 저지르는 오용의 사례

를 설명한 다음 마지막으로 요인분석의 방법에 대한 지침을 제공할 것이다.

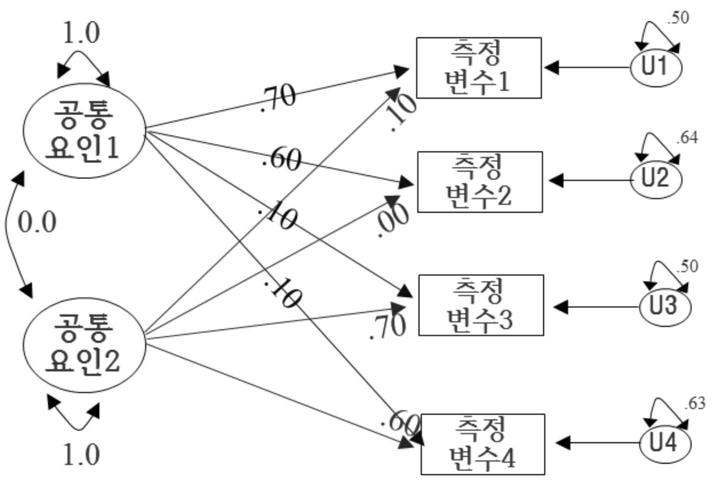
II. 요인분석의 원리

요인분석기법은 심리학의 지능이론과 함께 발달되어 왔다. 최초의 요인분석기법은 Spearman(1904)에 의하여 개발되었다. 그는 지적 능력검사 점수가 g-요인과 s-요인에 의하여 결정된다고 주장하였다. g-요인은 지력(intellectual power)으로 모든 인지적 과제를 수행할 수 있는 능력을 반영하는 요인이며, s-요인은 검사에 특수한 특성을 반영하는 요인이다. 그는 g-요인이 여러 능력간의 상관을 설명할 수 있음을 주장하였고, 이를 검증하기 위하여 Spearman의 2요인 요인분석모형을 제안하였다. 그는 지능의 하위검사들 간의 상관계수를 분석하여 모든 지능검사 점수에 영향을 미치는 공통요인이 있음을 보여주었다(Hart & Spearman 1912). 이후 이러한 Spearman의 요인분석모형은 지능이론과 더불어 다양한 형태로 발전되어 Thurstone(1935, 1938)의 다중요인모형(multiple factor model) 혹은 공통요인모형(common factor model)으로 완결되었다. 이 모형에서는 Spearman과는 달리 지능이 언어이해, 단어 유창성, 수 능력. 공간시각화능력, 연합기억, 지각속도, 추리의 일곱 가지 공통요인으로 구성되어 있다고 보았다.

다중요인모형에서 관찰변수(검사점수) X는 공통요인에 의하여 설명되는 부분과 그 변수에만 영향을 미치는 고유요인(unique factor)의 부분으로 분리된다. 즉

$$X_i = \lambda_{i1} z_1 + \lambda_{i2} z_2 + \ldots + \lambda_{im} z_m + u_i = \sum_{j=1}^{m} \lambda_{ij} z_j + u_i \tag{1}$$

여기에서 X_i는 i 번째 측정변수를, z_j는 j번째 공통요인점수를, u_i는 i번째 측정변수와 관련된 고유요인을 나타낸다. λ_{ij}는 요인점수가 관찰점수를 변화시키는 기울기를 말하고 요인부하량이라 명명된다. 요인의 수는 m으로 표기된다.



〈그림 1〉 요인분석모형의 경로도

요인분석모형은 <그림 1>과 같이 경로도의 형식으로도 표현될 수 있다. 경로도에서 원은 공통요인(잠재변수), 사각형은 측정변수, 화살표는 기울기, 양쪽 방향 화살표는 공분산(상관계수)을 나타낸다. 예컨대 측정변수3는 다음과 같이 표현된다.

$$측정변수3 = .10(공통요인1) + .70(공통요인2) + U3 \quad (2)$$

측정변수3은 두 요인에 의해서 설명되는데 공통요인1보다는 공통요인2에서 설명되는 부분이 더 크다. 공통요인들에 의해서 설명되지 않는 부분은 $U3$에 의해서 나타난다. 식(1)과 식(2)는 요인점수들이 독립변수이고 측정변수가 종속변수인 중다회귀분석모형과 동일하나, 다른 점은 요인점수 z_i가 관찰되지 않는다는 것이다. 따라서 회귀분석과 동일한 방법으로 식(1)의 요인부하량을 추정하는 것이 가능하지 않다.

요인분석에서는 측정변수들 간의 공분산행렬을 계산하고 식(1)로부터 유도되는 공분산구조를 유도한 다음, 이 둘의 차이를 가장 작게 만드는 모수치들(요인부하량, 고유분산)을 추정한다. 이때 공통요인점수와 고유요인점수의 상관이

0이며, 고유요인들 간의 상관도 0이라고 가정한다. X_i들 간의 공분산 행렬을 $S(p \times p)$라 표기하고 S를 예측하는 모형을 $\hat{\Sigma}$라 하면, $\hat{\Sigma}$은 다음과 같이 유도된다.

$$\hat{\Sigma} = \Lambda \Phi \Lambda' + D_{\psi} \tag{3}$$

여기서 Λ는 $p \times m$ 요인부하량 행렬이며, Φ는 요인점수들간의 상관행렬이고, D_{ψ} 는 고유요인들간의 공분산 행렬로 대각행렬이다. 위의 예에서 Λ, Φ, D_{ψ}는 다음과 같다.

$$\Lambda = \begin{pmatrix} .70 & .10 \\ .60 & .00 \\ .10 & .70 \\ .10 & .60 \end{pmatrix} \quad D_{\psi} = \begin{pmatrix} .50 & 0 & 0 & 0 \\ 0 & .64 & 0 & 0 \\ 0 & 0 & .50 & 0 \\ 0 & 0 & 0 & .63 \end{pmatrix} \quad \Phi = \begin{pmatrix} 1.00 & 0.00 \\ 0.00 & 1.00 \end{pmatrix} \quad \hat{\Sigma} = \begin{pmatrix} 1.00 & .42 & .14 & .13 \\ .42 & 1.00 & .06 & .06 \\ .14 & .06 & 1.00 & .43 \\ .13 & .06 & .43 & 1.00 \end{pmatrix}$$

실제로 관찰된 변수들간의 상관계수 S는 위의 $\hat{\Sigma}$와 동일하지 않다. S와 $\hat{\Sigma}$ 사이의 차이 (discrepancy)를 최소화하는 방식으로 $\hat{\Sigma}$ 즉 Λ와 Φ 그리고 D_{ψ}를 추정하게 된다. 만약 요인들 간의 상관계수가 0이면 Φ가 항등행렬이 되어 다음과 같이 된다.

$$\hat{\Sigma} = \Lambda \Lambda' + D_{\psi} \tag{3-1}$$

요인분석은 공분산구조에 대한 모형으로 식(3)에 의해서 표현될 수 있지만, 식(1)에 의하여서는 완벽하게 표현될 수는 없다. 따라서 요인분석기법을 이해하기 위해서는 위의 행렬을 이해하는 것이 필수적이다. 요인분석모형에 근거하여 식(1)을 다시 해석해 보면 다음과 같다. 요인의 평균은 0이고 분산은 1이다. X의 절편은 0이고 분산은 1이다. 즉 독립변수와 종속변수가 모두 표준점수이다. 따라서 식(1)은 표준화된 요인들로 표준화된 관찰변수를 예언하는 표준화된 중다회귀분석이다.

X_i의 분산은 공통요인에 의해 발생하는 분산 즉 공통분산(common variance)

과 고유요인에 의해 발생하는 분산 즉 고유분산(unique variance)의 합으로 구성되며 전체 X_i의 분산 중에서 공통분산이 차지하는 비율을 공통분(communality)이라 한다.

Ⅲ. 선택1: 주성분분석 대 요인분석

SPSS의 요인분석 프로그램에서는 주성분분석을 하여야 할지 요인분석을 하여야 할지를 선택하게 되어 있다. 사실 두 방법은 전혀 다른 통계적 기법인데도 불구하고 동일한 분석방법으로 분류되어 있다는 점에서 문제가 있다. 더구나 주성분분석이 기본선택으로 설정되어 있다. 설문지나 검사지를 사용하는 대부분의 연구에서 주성분분석을 사용하는 것은 잘못된 선택이다. 주성분분석이 무엇인지, 어떤 점에서 요인분석과 다른지에 대하여 보다 구체적으로 살펴보자.

1. 주성분분석

주성분분석은 Pearson(1901)에 의해 제안되고 Hotelling(1933)에 의해 정리된 기법이다. 요인분석이 심리측정학자들에 의하여 개발되고 발전된 것과 대조적으로 주성분분석은 통계학자들에 의해 개발되고 발전되었다. 요인분석은 관찰점수가 공통요인과 고유요인에 의해 예측되는 통계적 모형인데 비하여 주성분분석은 통계적 모형을 가지고 있지 않으며 단순히 자료의 차원 축소 방법이다. 주성분분석은 관찰변수가 다수일 때 관찰변수의 차원보다 적은 차원으로 자료를 기술하려는 기법이다. 예컨대 자동차의 품질을 나타내는 가속시간, 연비 등의 100개의 변수들이 있다고 할 때 이를 요약하여 소수의 지표(성분)들로 기술하는 기법이다.

p개의 관찰변수가 두 개의 성분에 의하여 표현될 수 있다고 가정하면 i번째 제품의 p개의 관찰값은 다음과 같은 식으로 기술된다.

$$
\begin{aligned}
x_{i1} &= \lambda_{11} z_{1i} + \lambda_{12} z_{2i} \\
x_{i2} &= \lambda_{21} z_{1i} + \lambda_{22} z_{2i} \\
&\cdots \\
x_{ip} &= \lambda_{p1} z_{1i} + \lambda_{p2} z_{2i}
\end{aligned}
$$

이를 m개의 성분을 가진 행렬식으로 나타내면

$$x_{p\times 1} = \Lambda_{p\times m} z_{m\times 1}$$

로 표현된다. 여기서 z는 각 개인의 성분점수를, Λ은 성분부하량을, p는 관찰변수의 수를, m은 성분의 수를 나타낸다. 만약 성분의 수가 변수의 수만큼 있으면 $(p=m)$ 성분점수로 관찰변수를 완벽하게 복원할 수 있다. 그러나 성분의 수가 관찰변수의 수보다 작으면 $(m<p)$, 당연히 성분점수로 관찰변수를 복원할 때 정보의 손실이 있게 된다. 주성분분석에서는 x들의 정보를 가능한 한 최대로 유지하면서 차원을 축소한다. 이때 관찰변수들 간의 공분산행렬을 스펙트럼분해(spectral decomposition)을 이용하여 분해하면, 성분 부하량(Λ)과 각 개인별로 p개의 성분점수(z)를 구할 수 있다.

이를 기하학적으로 해석하면 다음과 같다. p개의 변수를 가진 N명의 자료가 있다고 가정하자. 모든 자료는 p차원상에 한 점으로 표상할 수 있다. p개 축은 각각 변수를 나타낸다. 이 축을 직교회전하여 새로운 축을 구성하는데, 자료의 분산을 가장 많이 반영하는 축을 첫 번째 축으로, 그 축에 직교하면서 두 번째로 자료의 분산을 많이 반영하는 축을 두 번째 축으로 삼는다. 이러한 방식으로 p개의 새로운 축이 생성된다. 즉 새로운 p개의 변수가 생성되게 되는데, 이 변수의 분산은 고유치와 일치하게 된다. 차원을 축소하는 과정은 고유치가 낮은 축들을 삭제함으로써 $(p>m)$ p차원을 m차원으로 줄일 수 있다.

요인분석에서 요인점수는 잠재변수인데 반하여 주성분분석에서 성분점수z는 관찰변수의 선형결합으로 표현될 수 있다.

$$z = (\Lambda'\Lambda)^{-1}\Lambda' x$$

위의 예에서 첫 번째 성분점수를 품질지수로 삼으면, 이는 100개의 품질을 나타내는 변수들의 1차 함수로 계산될 수 있다. 요인분석에서는 요인점수가 진점수이고 이 진점수의 분포를 추정한다면, 주성분분석에서는 성분점수는 측정변수들로부터 직접적으로 계산될 수 있지만, 분포를 가지고 있지는 않다.

2. 잠재변수(요인)와 성분점수

요인분석이 심리측정이론가들에 의해서 개발되면서 잠재변수의 개념이 등장하였다. 이 변수는 실제로 관찰되지는 않지만 측정변수에 영향을 주는 변수이다. 예컨대 지능은 직접적으로 관찰되지는 않지만 지능검사 점수들에 영향을 준다. 이때 지능은 잠재변수이고 지능검사 점수는 측정변수이다. 심리학에서 잠재변수의 개념은 행동주의와 함께 등장한 조작주의(operationalism)와 밀접하게 관련된다. 심리학에서 태도와 능력과 같은 직접적으로 관찰되지 않는 개념을 가설적 구성체라 명명하는데 이에 대응되는 변수가 잠재변수이다. 이 가설적 구성체는 관찰 가능한 행동들에 의하여 발현되는데(manifest), 관찰 가능하게 발현된 행동으로 개념을 정의하는 것이 조작적 정의이다. 이에 대한 변수가 발현변수(manifest variable), 측정변수(measured variable), 혹은 관찰변수(observed variable)이다.

3. 요인분석과 주성분분석의 비교

학술지에서 심리검사나 척도를 분석하면서 주성분분석을 적용시키는 것을 쉽게 발견할 수 있다. 주성분분석은 요인분석과 동일한 분석방법이 아니며, 심리검사나 척도에서 주성분분석은 적절한 기법이 아니다. 요인분석이 공분산 혹은 상관계수 행렬에 대한 통계적 모형인데 반하여 주성분분석은 자료의 차원을 축소하는 기법으로 성분점수를 산출하는 것이 주목적이다. 요인분석과 주성분분석이 전혀 다른 통계적 기법이기 때문에 비교하는 것이 용이하지 않으나 다음과 같은 차이점이 있다.

첫째, 요인분석의 경우는 요인점수가 잠재변수이고 관찰변수들이 요인점수

에 의하여 결정되게 된다. 반면 주성분분석에서는 성분점수가 관찰변수에 의하여 결정된다. 이는 심리측정의 관점에서 중요한 차이이다. 예컨대 우리가 관심을 가지는 구성체가 공격성이라고 하고 이를 측정하는 변수로 공격성 척도 점수 A, B, C가 있다고 하자. 요인분석에서는 공격성이 증가하면 A, B, C의 점수가 증가하게 된다는 것을 예측하고, 주성분분석에서는 A, B, C 점수를 적절하게 선형결합하면 공격성이 결정된다고 본다. 심리측정의 맥락에서 보면 A, B, C의 세 점수가 모든 공격성의 행동영역(domain)을 모두 포함할 수 없으므로 세 가지 점수의 선형결합으로 공격성 점수를 정의하는 것은 무리가 있다(Cronbach & Meehl 1955; Margenau 1955). 반면 공격성이 A, B, C 점수를 결정한다는 요인분석의 이론은 A, B, C가 모든 영역을 포함하고 있지 않더라도 크게 무리가 없다. 따라서 심리검사와 척도에서 구성개념에 대한 추론을 위해서는 요인분석이 더 적절하다고 할 수 있다.

둘째, 요인분석에서는 관찰변수들 간의 상관계수(혹은 공분산)들을 가장 잘 예측할 수 있는 모형을 찾는다. 요인분석모형에서는 관찰변수의 분산은 공통요인에 의하여 설명되지 않는 부분은 고유요인에 의하여 설명되기 때문에 각 측정변수의 분산은 모두 설명된다. 따라서 자료의 적합과정에서 변수들 간의 공분산을 잘 예측하는 모형을 찾게 된다. 반면 주성분분석에서는 관찰변수들 간의 상관계수들뿐만 아니라 분산도 적합할 수 있는 해를 찾는다. 따라서 주성분분석은 요인분석보다 상관계수들을 잘 예측하지 못 하게 된다.

셋째, 요인분석은 통계적 모형을 가정하고 이 모형에 자료를 적합시키는 방법인데 반하여, 주성분분석은 통계적 모형이 없으며 단순히 자료 주도적으로 차원을 축소시키는 방법이다. 따라서 요인분석에서는 분포이론을 적용시킬 수 있고 분포이론에 근거한 모형적합도 지수를 산출하는 것이 가능하다. 이에 반하여 주성분분석에서는 분포이론의 도입이 가능하지 않으며, 따라서 적합도 지수를 산출하는 것이 가능하지 않다.

이 두 모형의 가장 큰 차이는 요인이 측정변수들의 함수로 정의될 수 있는지의 여부이다. 요인분석의 측정변수는 반영적 지표함수(reflective indicator), 주성분분석의 측정변수는 형성적 지표변수(formative indicator) 혹은 인과적 지표변수(causal indicator)로 불리기도 한다. 이러한 맥락에서 요인분석의 목적

이 척도개발이라면, 주성분분석의 목표는 지수(지표)의 개발로 볼 수 있다. 요인분석에서는 요인점수가 계산될 수 없기 때문에 지표의 개발에 사용될 수 없다(학자들에 따라서 요인점수를 추정하는 경우도 있다).

Ⅳ. 선택 2: 요인분석의 자료 적합방법

요인분석은 통계적 모형이기 때문에, 모형이 예측하는 자료의 패턴과 실제 자료의 패턴의 차이를 최소화 하는 방식으로 모형의 모수치를 구하게 된다. 이때 차이를 정의하는 다양한 방식이 존재하며, 어떤 자료 적합방식을 사용하느냐에 따라서 모수치에 대한 추정치가 달라진다. 자료 적합방법은 관찰된 공분산 행렬(S)과 모형에 의하여 함축된(implied) 행렬($\hat{\Sigma}$) 사이의 불일치를 최소화하는 요인분석모형($\hat{\Sigma}= \Lambda\Phi\Lambda' + D_{\psi}$)을 찾아내는 것이다. 요인분석에서 많이 사용되는 방법으로 최대우도법(Maximum Wishart Likelihood), 일반화된 최소제곱법(Generalized Least Square), 최소제곱법(Ordinary Least Square)이 있다. 각각의 불일치 함수는 다음과 같다.

최대우도법

$$f_{ML}= \log|\hat{\Sigma}| - \log|S| + tr\left(\hat{\Sigma}^{-1}S\right) - p = \frac{1}{2}tr\left[\hat{\Sigma}(S-\hat{\Sigma})\right]^2$$

일반화된 최소제곱법

$$f_{GLS} = \frac{1}{2}tr\left[S^{-1}(S-\hat{\Sigma})\right]^2$$

최소제곱법

$$f_{OLS} = \frac{1}{2}tr\left[(S-\hat{\Sigma})^2\right]$$

이밖에도 주축분해법이 있다. 이 방법에서는 불일치 함수에 의하여 모형을

적합시키는 것이 아니라 소위 축소된 상관계수행렬($S-D_{\psi}$)의 차원을 축소하여 요인을 구한다. 만약에 고유분산 D_{ψ}이 알려져 있다면 스펙트럼분석에 의하여 요인부하량 Λ이 다음과 같이 구해진다.

$$S-D_{\psi} = UDU' = \Lambda\Lambda'$$

$$\Lambda = UD^{1/2}$$

문제는 고유분산을 알지 못한다는 것이다. 따라서 먼저 여러 가지 방법으로 고유분산을 초기값을 정하고 요인부하량 Λ을 구한다. Λ이 구해지면 식(3-1)에 의하여 D_{ψ}가 구해진다. 반복적으로 Λ와 D_{ψ}를 계산하여 안정된 결과를 얻게 되는 것이 주축분해법이다.

자료가 정규분포를 따르면, 더 엄격하게 말해서 공분산 행렬이 Wishart 분포(Wishart distribution)를 따르면, 가장 바람직한 방법은 최대우도법일 것이다. 최대우도법은 분포에 기반했기 때문에 추정치의 정확성을 판단하거나 모형의 양호도를 판단할 때 검증통계량을 사용할 수 있다는 장점이 있다. 물론 분포에 대한 가정이 위배되면 이 통계량은 편향될 수 있다. 그러나 정규분포를 굳이 따르지 않더라도 첨도가 표준 정규분포의 첨도인 3을 크게 벗어나지 않는다면, 최대우도법을 사용하더라도 편향이 크지 않는 특징이 있어서 이 방법의 적용범위를 더 넓히고 있다(Amemiya & Anderson 1990; Anderson & Amemiya 1988; Browne & Shapiro 1988; Shapiro 1987; Yuan and Bentler 1999). 다른 추정방법은 분포를 가정하지 않기 때문에 검증통계치를 계산하는 것이 가능하지 않다. 다만, GLS의 경우 표본의 크기가 커지면 최대우도추정량에 근접하기 때문에(Browne 1973), 검증통계치를 사용할 수 있다. 따라서 최대우도법이 해를 제공하지 못할 경우에는 GLS를 사용하여 해를 구할 수 있다.

1. 서열척도의 요인분석

설문지나 검사지의 문항을 이용하여 요인분석을 하는 경우, Likert 척도와 같은 서열척도나 이분변수를 포함하는 경우가 흔히 있지만, 많은 통계패키지

가 이를 적절하게 처리하지 않고 있다. 요인분석은 정규분포, 더 나아가서 위사트 분포에 근거하고 있다. 이분변수나 서열척도인 경우, 각 변수는 정규분포를 따르지 않고 공분산은 위사트 분포를 따르지 않으므로 이 분포에 기반을 둔 최대우도법으로 적합시킬 수 없다. 특히 이분/서열 척도의 경우에는 Pearson 상관 행렬도 아니기 때문에 이에 대한 적절한 처리가 필요하다.

이분/서열 척도인 경우에는 다분상관계수(polychoric correlation)를 계산할 수 있다. 이 상관계수는 원래 연속적인 차원을 몇 개의 범주로 나누어서 사람들이 반응한다는 가정하에서 계산된다. 예컨대 "당신은 호주제 폐지를 찬성하십니까?"라는 질문에 대하여 "아주 찬성한다-찬성한다-찬성도 반대도 하지 않는다-반대한다-아주 반대한다"의 반응을 수집하고 그 점수를 1점부터 5점까지 할당하였다고 하자. 이 점수는 연속변수(등간척도)가 아니다. 그러나 찬성-반대 심리적 크기는 연속변수로 볼 수 있다. 이 변수를 z라 하면 z와 Likert 척도 사이에는 다음과 같은 법칙에 의하여 반응이 산출될 수 있다.

$$
\begin{array}{ll}
z < a1 & \rightarrow \text{점수1} \\
a1 < z < a2 & \rightarrow \text{점수2} \\
a2 < z < a3 & \rightarrow \text{점수3} \\
a3 < z < a4 & \rightarrow \text{점수4} \\
a4 < z & \rightarrow \text{점수5}
\end{array}
$$

즉 네 개의 역(threshold; $a1$, $a2$, $a3$, $a4$)에 의하여 점수가 산출된다. 다분상관계수는 서열척도를 이용하여 연속 잠재변수 z들 간의 상관을 계산하는 방식이다.

다분상관행렬을 사용해 요인분석을 하는 경우에는 Wishart 분포를 사용한 최대우도법을 사용하여 추정할 수 없다. 이 경우에 사용가능한 방법이 Browne이 제시한 ADF(Asymptotically Distribution Free) 방법이다(Browne 1982; 1984). 이 방법은 가중최소제곱법(Weighted Least Square; WLS)의 일종으로 불일치 함수는 다음과 같다.

$$f_{adf} = (s-\sigma)' W^{-1}(s-\sigma)$$

여기서 s와 σ는 각각 관찰된 공분산행렬과 함의된 공분산행렬을 벡터의 형태로 재배열한 것이며, W는 $\mathbf{s}$의 공분산행렬이다. 가중치 행렬 W는 변수의 수가 많아지면 기하급수적으로 늘어나기 때문에 가중치 행렬에서 대각행렬만을 사용하는 DWLS(Diagonal Weighted Least Squares) 방식을 사용하기도 한다.

기존의 프로그램에서는 이러한 분포의 문제를 소홀히 처리하고 있다. 그 이유의 하나가 요인분석을 주성분분석의 일종으로 보았기 때문으로 추측된다. 주성분분석은 단순한 차원 축소이기 때문에 분포에 대한 고려가 필요하지 않다. 그러나 요인분석은 통계적 모형으로 분포가 모형을 적합시키는 데 중요한 요소가 되고 이에 대한 고려 없이는 정확한 분석이 가능하지 않다.

SPSS나 SAS와 같은 범용 프로그램은 이분/서열척도를 처리하는 방법을 제공하고 있지 않다. 그러나 공분산 구조모형의 프로그램인 LISREL, MPLUS에서는 WLS방법으로 추정하는 방법을 제공하고 있고, CEFA(Comprehensive Exploratory Factor Analysis; Browne et al. 2004)등을 사용하면 다분상관행렬을 분석할 수 있다.

요약하면 다음과 같다. 일반적으로 정규분포에서 크게 벗어나지 않는다면 최대우도법이 가장 바람직하다. 최대우도법이 해를 구하지 못한다면 GLS 방법이 그다음으로 고려될 수 있다. 척도가 이분척도나 서열척도인 경우에는 이를 처리할 수 있는 방법을 적용하여야 한다.

Ⅴ. 선택 3: 모형의 양호도 판단

최대우도법과 일반화된 최소제곱법으로 요인분석을 한 경우에는, 우도비검증(likelihood ratio test)을 이용하여 연구모형이 맞다는 영가설에 대한 가설검증을 할 수 있다. 이 우도비의 로그값에 -2를 곱한 값은 카이제곱분포를 따르게 된다. 이때 자유도는 (전체 공분산의 개수 - 파라미터의 수 + 제약의 수)가 되어 다음과 같이 된다.

$$df = \frac{1}{2}\left[(p-m)^2 - (p+m)\right]$$

이 영가설을 기각하면 연구모형이 틀린 것으로 된다. 그러나 가설검증의 문제는 표본의 크기가 커지면 영가설이 기각될 확률이 높아진다는 것이다. 그리고 대부분의 경우에 영가설은 사실이 아니므로 가설검증을 사용하는 데에는 한계가 있다. 이 방법을 대체하기 위하여 RMSEA, ECVI, NNFI, NFI, CFI 등의 적합도 지수가 개발되고 그에 대한 기준도 제시되었다. RMSEA의 경우, .05보다 작으면 좋은 모형, .08보다 작으면 합리적인 모형, .10보다 크면 좋지 않은 모형으로 분류되는데(Browne & Cudeck 1983), Hu & Bentler(1999)는 .06을 기준으로 삼기도 한다. 전통적으로 NNFI, NFI, CFI 등의 경우는 .90보다 크면 양호한 모형 그렇지 않으면 좋지 않은 모형으로 평가하지만(Marsh, Hau & Wen 1999), Hu & Bentler(1999)는 .95를 기준으로 좋은 모형과 그렇지 않은 모형을 구분하기도 한다. 따라서 다양한 적합도 지수를 모두 고려하여 모형의 양호도를 판단하여야 한다.

대부분의 상업적인 소프트웨어에서는 적합도 지수는 제시하지 않고 설명분산(explained variance)을 제공하기 때문에, 많은 논문들에서 설명분산으로 모형의 양호도를 판단하기도 한다. 이는 주성분분석에서 사용될 수 있는 개념으로 요인분석모형의 적합도를 판단하는 기준으로 사용하는 데 한계가 있다. 보다 구체적으로 살펴보면 다음과 같다.

설명분산은 다음과 같은 식을 통해서 구할 수 있다. 앞에서 설명했던 주성분분석에서, 전체 분산 중 i번째 축에 의하여 설명될 수 있는 분산은 다음과 같다.

$$\text{설명분산} = \frac{l_i}{\sum_{i=1}^{p}}$$

여기서 l_i는 i번째 고유치이다. 즉 전체 분산은 모든 고유치들의 합이 되고 각 축이 설명하는 분산은 그 축의 고유치가 된다. 그리고 각 축이 직교하므로 첫 번째부터 m번째 축까지 설명하는 분산은 다음과 같다.

$$설명분산 = \frac{\sum_{i=1}^{m}}{\sum_{i=1}^{p}}$$

요인분석에서는 주성분분석과는 달리 고유요인이 존재하는데 이 고유요인을 고려하지 않고 표본상관행렬을 분해하는 것은 의미가 없다. 만약 고유요인의 분산이 0이라면 이 두 방법은 일치하게 되는데 이러한 가정은 현실적이지 않다.

이 방법을 사용할 때 가장 큰 문제는, 회전을 시키면 각 축이 설명하는 정도와 고유치에 의하여 계산되는 각 축의 설명분산이 일치하지 않는다는 점이다.[1] 첫 번째 고유치에 의해서 계산되는 설명분산은 회전된 첫 번째 요인이 설명하는 분산과 무관하다. 이는 주성분분석을 해석할 때도 마찬가지이다. 학술지에 보고된 많은 논문들이 여전히 이러한 잘못된 해석을 반복하고 있다.

이러한 문제점을 해결하기 위하여, SPSS와 SAS는 최근에 요인부하량의 제곱합을 관찰변수의 수로 나눈 값을 결과에 포함시켰다. 회전된 요인부하량 행렬을 사용하면 회전 후의 소위 설명분산을 계산할 수 있는 장점이 있다. 더 정확히 이 수치를 해석하면 다음과 같다. 요인들 간의 상관이 없다면, 요인부하량의 제곱합은 각 관찰변수의 공통분(communality)의 합과 동일하다. 즉 식 (3-1)로부터 다음이 성립한다.

$$\hat{\Sigma} - D_{\psi} = \Lambda\Lambda'$$

위 식에서 좌변의 대각요소들은 각 변수의 공통분 즉 각 변수가 공통요인들에 의하여 설명되는 분산을 나타내고 있다. 그리고 우변의 대각요소는 각 요인

---

1) 원래 주성분분석에서는 회전이 허용되지 않았다. 축의 순서가 설명량을 반영하고 있기 때문에 축을 회전하는 것은 논리적으로 가능하지 않다. 주성분분석을 다중다변량회귀분석을 이용하여 유도한 다음 축의 회전에 대하여 타당한 것으로 받아들여져서 회전이 허용되었다.

부하량의 제곱합으로 구성되므로 요인분석모형에서는 요인부하량의 제곱합이 공통분이 된다. 관찰변수가 요인분석에서 설정한 요인들에 의하여 모두 설명된다면 공통분의 합은 관찰변수의 개수와 일치하게 된다. 따라서 요인부하량의 제곱합을 관찰변수의 수로 나누면, 전체 분산 중에서 공통요인에 의해 설명되는 분산의 합이 된다. 그리고 각 요인별로 요인부하량의 제곱합을 구하면, 동일한 맥락 속에서 각 요인의 공통분이 전체 분산에서 얼마나 차지하고 있는지가 된다.

다시 한 번 강조하지만 이 수치는 모형의 적합도와는 직접적으로 관련되어 있지 않다. 이 분산이 작다고 할 때, 그 이유는 요인의 수를 원래보다 작게 정했기 때문일 수도 있고(즉 요인분석모형이 잘못되었기 때문일 수도 있고), 요인의 수는 정확하게 정했으나 고유분산의 크기가 0이 아니기 때문일 수도 있고, 혹은 이 두 이유 모두 때문일 수도 있다.

또한 요인부하량의 제곱합에 의한 해석은 직교회전을 한 경우에만 가능하다. 만약 사교회전을 한다면, 요인부하량의 제곱합이 더 이상 공통분의 합과 일치하지 않기 때문에 이러한 해석을 할 수 없다. SAS 프로그램의 경우는 소위 구조행렬(structure matrix)을 사용하여 각 요인에 대한 설명분산을 새롭게 계산한다. 먼저 아래와 같이 새로운 Λ^*를 계산한다.

$$\hat{\Sigma} - D_\psi = \Lambda\Phi\Lambda' = \Lambda\Phi^{1/2}\Phi^{1/2}\Lambda' = \Lambda^*\Lambda^{*'} \quad whrere \quad \Lambda^* = \Lambda\Phi^{1/2}$$

이를 구조행렬 (structure matrix)이라 한다. 이 행렬에서 구성요소는 각각 요인점수와 관찰변수사이의 상관을 나타낸다. 직교회전과 마찬가지로 Λ^*로 구성된 모형에는 Φ가 없으므로 전체 구성요소의 제곱합은 전체 공통분과 동일하게 된다. 요인들간의 상관을 가정하고 있지 않은 모형과 동일한 방식으로 각 요인별로 부하량을 제곱하여 합하면 각 요인의 설명분산을 계산할 수 있다.

이 설명분산이 높으면 공통분이 대부분의 분산을 설명하고 있기 때문에 좋은 모형에 가깝다고 할 수 있으나, 이 분산이 작다고 하여 반드시 나쁜 모형이라고 할 수 없다. 공통분이 크지 않지만 자료를 잘 설명할 수 있는 모형도 있기 때문이다.

요인분석에서는 모형에 대한 양호도를 크게 세 가지 기준에 의하여 판단할 수 있다.

첫째는 위에서 제시된 적합도 지수이다. 자료가 모형에 근접하지 못하면 좋은 않은 모형으로 결론을 내리면 된다. 제안된 적합도 지수의 기준들이 전문가들에 의해 제시되었지만, 통계적 분포이론에서 나온 것이 아니고 경험에서 얻은 일반 원리 (rule of thumb)이거나 제한된 상황에서 실시된 모의실험의 결과에서 나온 것이다. 이는 중요한 기준이기는 하지만 절대적인 기준이 될 수는 없다.

둘째는 여러 모형에서의 상대적인 우월성이다. 탐색적 요인분석에서는 요인의 수에 따라서 하나 이상의 모형들이 존재하게 되는데, 어떤 모형이 더 합리적인 모형인지는 AIC나 BIC 등의 지표를 사용하여 판단할 수 있다. 이 통계치는 적은 값을 가질수록 더 합리적인 모형으로 평가할 수 있다.

셋째는 해석가능성이다. 적합된 요인부하량의 형태가 상식이나 연구자의 지식으로 설명될 수 있는지가 중요한 기준이 된다. 두 모형이 유사한 적합도를 보이면 해석가능성에 의하여 최종적으로 판단하는 것이 바람직하다.

Ⅵ. 선택 4: 요인부하량 행렬의 회전 방법

요인부하량 행렬(Λ)은 그림 1에서 각 요인점수가 각 측정변수에 미치는 영향력, 즉 요인부하량으로 구성된 행렬이다. 요인분석의 적합과정에서 풀이된 해에서 고유분산까지는 정해지지만 요인부하량까지는 결정되지 않는다. 요인들 간의 상관이 0이라 가정하면 함축 행렬 ($\hat{\Sigma}$)은 $\hat{\Sigma} = \Lambda\Lambda' + D_\psi$로 표현된다. 여기서 직교표준(orthonormal) 행렬 T가 있다고 가정하면 $TT' = I$가 성립되므로, $\Lambda\Lambda' = \Lambda TT'\Lambda' = \Lambda_r\Lambda_r{}'$가 성립되어 두 개의 서로 다른 요인부하량 행렬 Λ와 Λ_r은 동일한 함축 공분산 행렬을 산출하게 된다. 이러한 방식으로 계속해서 새로운 Λ_r을 만들어내,면 동일한 함축 공분산 행렬($\hat{\Sigma}$)을 유발하는 요인부하량 행렬은 무수히 많이 존재하게 된다. 요인들 간의 상관이 있다고 가정하

면, 더 이상 T가 직교표준이 되지 않게 되므로 $\Lambda\Lambda' = \Lambda T(T^{-1}T^{-1\prime})T'\Lambda' = \Lambda_r \Phi \Lambda_r{}'$이 되어 요인점수들 간의 상관계수행렬 ($\Phi = T^{-1}T^{-1\prime}$)이 새롭게 나타나게 된다.

관찰자료를 동일한 정도를 설명할 수 있는 무수히 많은 요인부하량 행렬(Λ_r) 들이 존재하는데, 연구자들은 선호하는 패턴을 정의한 후 이에 가장 부합되는 Λ_r을 찾아내는 것이 회전(rotation)이다. 가장 대표적으로 사용되는 회전기준은 열절약(row parsimony) 방식과 행절약(column parsimony)방식이다. 전자는 하나의 열에서 높은 부하량이 적고 낮은 부하량이 많은 것을 말하며, 후자는 하나의 행에서 높은 부하량이 적고 낮은 부하량이 많은 것을 말한다.

위에서 살펴본 바와 같이 요인들 간의 상관을 0이라고 가정하는 회전방법이 직교회전법(orthogonal rotation)이고, 요인들 간의 상관을 있을 것으로 가정하는 방법이 사교회전법(oblique rotation)이다. 요인분석에서 가장 많이 사용되는 회전방법은 직교회전법인 VARIMAX 방식이다. 이 방법 역시 상용프로그램에서 기본값으로 되어 있다. 만약 연구자가 열절약에 비중을 더 많이 둔다면 다른 방식을 사용하는 것이 바람직할 것이다. 특히 회전과 관련하여 연구자가 주의할 점은 직교회전법을 사용해야 하는 경우가 사회과학 연구에서는 빈번하게 발생하지 않는다는 것이다. 한 연구에서 여러 개의 요인들을 다룰 때에는 요인들 간의 상관을 가정하는 것이 이론적으로 더 타당한 경우가 많다. 이러한 경우에는 사교회전법을 적용시켜야 한다. 요인분석법을 사용하는 연구들에서 회전과 관련하여 흔히 관찰되는 잘못된 분석의 하나는 다음과 같다. 먼저 직교회전의 방식으로 요인분석을 한 다음, 각 요인에 해당하는 문항들의 합산점수로서 하위요인점수를 계산한다. 그리고 그 점수들 간의 상관계수를 구하여 요인들 간의 상관인 것처럼 해석한다. 직교회전을 하였다는 것은 요인들 간의 상관이 0이라고 가정한다는 것을 의미한다. 따라서 이러한 방식으로 요인들 간의 상관계수를 내는 것은 의미가 없다. 보다 바람직한 방식은 사교회전의 결과로 산출되는 요인점수들 간의 상관행렬을 보고하는 것이다.

요인회전법 중에서 유용하지만 많이 사용되지 않는 방법으로 표적회전법(target rotation; Browne 1972a, 1972b)이 있다. 회전은 자료를 동일하게 설명하는 무한히 많이 존재하는 요인부하량 중에서 하나를 선택하는 방법이므로

연구자가 바람직한 요인부하량 행렬의 패턴을 가장 잘 알고 있다. 연구자가 직접 낮은 부하량을 가질 것으로 예상되는 부하량 즉 λ_{ij}를 명세하는 것이다. 회전을 하는 과정에서 명세된 부하량을 작게 하는 방식으로 최종 Λ_r를 찾아내게 된다. 표적회전법은 직교회전이나 사교회전 모두에서 가능한 방법이다. 많은 연구자들이 표적회전이 주관적인 판단이 개입된다는 생각으로 꺼려하기도 하는 동시에 표적회전을 제공하는 소프트웨어가 많지 않기 때문에 사용하지 않는다. 그러나 회전이 일어나는 과정을 고려하면, 이 회전법은 다른 회전법에 비하여 결코 더 주관적이라 할 수 없다. 다른 회전법도 연구자가 원하는 요인의 구조를, 예컨대 행절약이나 열절약과 같이, 미리 명세해 두기 때문이다.

요약하면 다음과 같다. 대부분의 경우에 직교회전보다는 사교회전이 바람직하다. 다양한 회전법 중에서 자신의 연구에 맞는 회전법을 찾아내어 적용시켜야 한다. 예컨대 설문지가 일부 문항들이 두 개 이상의 요인들과 관련되게 설계되었다면, 소프트웨어에서 제공하는 회전법으로는 이러한 구조를 발견할 수 없다. 소프트웨어에서 적절한 회전법을 찾을 수 없으면 표적회전법을 선택하는 것이 바람직하다.

Ⅶ. 선택 5: 요인의 수의 결정

요인의 수에 따라서 자료에 대한 해석이 달라지기 때문에 탐색적 요인분석에서 요인의 수를 결정하는 일은 중요하다. 가장 많이 사용되는 방법은 표본상관계수 행렬의 고유치가 1보다 큰 값이 몇 개 있는지로 정하는 것이다. 몇몇 소프트웨어에서는 기본값으로 정해져 있다. 이 방법은 Guttman(1954)에 의해 처음 제안되었다. 그는 모집단의 상관계수 행렬의 최소 차원(rank)이 1보다 큰 고유치의 개수와 일치됨을 보여주었다. 이러한 발견을 Kaiser(1961)가 표본상관행렬에 적용시켰다(Cattell 1966). 이 법칙은 Kaiser-Guttman 고유치 법칙으로 알려져 있다. 그러나 이 법칙에 따라서 요인의 수를 결정하면, 작은 표본에서는 요인의 수를 과대평가하는 경향이 있으며, 조건에 따라 요인의 수를 과소 혹은 과대평가하는 경향이 있다(Linn 1968; Cattell & Vogelman 1977; Hakstian, Rogers & Cattell 1982; Zwick & Velicer 1982, 1986; Preacher et al. 2015). 따라서 이 방법은 요인의 수를 정하는 근사치로 사용될 수 있으나 최종 요인 수를 결정하는 기준으로 사용되기는 적절하지 않다.

고유치 기준에 대한 보완으로 평행검사 기법이 제안되기도 하였다(Horn 1965; Zwick & Velicer 1986). 이 기법에서는 표본상관계수 행렬의 크기와 동일한 무선 상관행렬을 생성한 다음 고유치를 계산하고 평균 고유치보다 큰 고유치의 개수로 요인의 수를 추정한다. 이 기법은 표집의 오차를 고려하는 방법이기 때문에 Kaiser-Guttman방법보다는 더 논리적인 방법이다. 이 방법을 제공하는 소프트웨어는 많지는 않지만, SAS, Mplus 등에서 분석할 수 있다(Hayton, Allen & Scarpello 2004; Ledesma & Valero-Mora 2007).

두 번째로 스크리 검사(scree test)에 의한 방법이 Cattell(1966)에 의하여 제안되었다. 이 방법은 다음과 같다. 각 고유치는 각 차원이 자료를 어느 정도 설명하는지를 반영하고 있다. 새로운 차원을 추가하는 것이 설명분산을 유의미하게 증가시키지 않으면 요인을 포함하지 않는 방법이다. 따라서 고유치를 내림차순으로 정렬한 다음 새로운 차원을 더하여도 고유치가 크게 증가되지 않는 시점에서 요인의 수를 정하게 된다. 바꾸어 말하면 고유치 곡선에서 변곡점을 찾아내고 그 변곡점을 요인의 수로 결정하는 방법이다. 이 방법은 요인의 수를 판단하는 하나의 지표로 널리 사용된다.

세 번째로 적합도 지수를 이용하여 요인의 수를 정하는 방법이 있다. 이에 대한 구체적인 절차는 잘 알려져 있지 않다. 원칙은 적합도 지수상에서 좋은 모형으로 평가되면 그 모형을 선택하는 것이다. 그런데 이러한 지표를 사용하는데 고려해야 할 점이 있다. 요인의 수가 많아짐에 따라 복잡한 모형이 되어서 카이제곱의 값은 항상 작아지므로 카이제곱에 근거한 많은 적합도 지수를 그대로 사용할 수는 없다. RMSEA, AIC, BIC, ECVI 등의 적합도 지수는 복잡한 모형에 대하여 벌점(penalty)을 주어 모형의 복잡성을 고려한다. RMSEA와 ECVI 등을 연구한 Browne & Cudeck(1993)에서 보면 여러 적합도들이 표본의 크기와 모형의 복잡성에 따라 다른 방식으로 작동한다. 여기서 제안할 수 있는 한 가지 방법은 1요인모형에서 출발하여 요인을 하나씩 증가하여 RMSEA가 .08보다 작아지는 지점에서 멈추는 것이다.

여러 가지 다양한 모형 선택의 방법이 제시되었지만 절대적인 기준이 있을 수 없다. 요인의 수에 대한 가장 바람직한 선택 방법은 위에서 제시된 모든 통계치를 고려하고 검사에서 원래 측정하고자 하는 구성체의 특성을 고려하여 최종적으로 연구자가 결정하는 것이다. 이러한 판단이 다소 주관적이라고 생각되면, 확인적 요인분석을 통하여 자신의 가설을 확인하면 된다.

Ⅷ. 선택 6: 결측치 처리방법의 선택

실제 자료를 처리하는 과정에서 결측치가 발생하는 경우가 많으나 요인분석의 경우, 이 결측치의 처리방법에 대한 논의가 상대적으로 적은 것으로 보인다. SPSS등의 통계분석 프로그램에서는 요인분석에서 크게 두 가지 결측치 처리방식을 제공하는데, 하나는 결측치만 분석에서 제거하는 소위 짝제거(pair-wise deletion)이고, 다른 하나는 하나의 변수에서라도 결측치가 생기면 그 사례를 모두 제거하는 사례제거(list-wise deletion)이다. 전자의 경우 분석자료인 공분산행렬이 양정치행렬(positive definite)이 되지 못하는 경우가 발생하므로 모형을 적합시키는 과정에서 계산이 불가능할 수 있다. 대부분의 경우 사례제거를 하여 양정치행렬을 만든 다음 분석한다. 그러나 구체적으로 요인분석과 관련하여 연구된 바는 없지만, 일반적으로 사례제거의 방식은 MCAR(Missing Completely at Random, Little & Rubin 2002)이 아닌 경우에는 추정치를 왜곡시킬 가능성이 높다는 것이 알려져 있다(Schafer & Graham 2002; Graham, Cumsille & Elek-Fisk 2003). 또한 사례제거 방식은 자료수집의 상황에 따라서 사례 수를 현저하게 줄일 수 있기 때문에 바람직하지 않을 수 있다.

이러한 문제를 완화시킬 수 있는 방법으로 FIML(Full Information Maximum Likelihood) 방법이나 다중대치법(Multiple Imputation) 방법들을 사용할 수 있다. FILM은 개인별로 우도를 구한 다음 이를 모두 합하여 전체 우도를 계산하는 방식이다. 개인별로 결측치가 있는 부분을 제외한 모형을 구성하여 우도를 구성하기 때문에 개인마다 다양한 결측치의 형태를 가지더라도 처리할 수

있는 이점이 있다.

대치법은 결측치를 다른 변수의 자료를 이용하여 채워 넣는다. 이때 결측치의 변수의 특성과 개인의 특성을 동시에 고려하여 가장 적절한 값으로 대치하게 된다. EM 알고리즘을 사용한 대치방법이 빈번히 적용된다(Rubin 1976' Pigott 2001). 이 방식은 결측된 값이 다른 변수의 값에 의하여 예측되기 때문에 변수들 간의 관계에 대한 통계모형의 경우에 그 관계의 크기를 과대추정할 가능성이 있다. 이러한 문제를 해결하기 위한 방식이 다중대치법이다. 이 방법은 표집오차를 포함하는 방식으로 한 결측치에 대하여 두 개 이상의 대치값을 만들어 내고, 서로 다른 대치값에서 나온 결과들을 통합하여 최종적인 추정치를 내는 방식이다.

요약하면 결측치가 발생하면 그 사례를 지우고 분석하기보다는 다중대치법이나 FIML 등의 방식을 사용하여 처리하는 것이 덜 편향된 결과를 산출한다.

Ⅸ. 선택 7: 소프트웨어의 선택

대부분의 상용 통계 패키지(예컨대 SPSS, SAS)가 요인분석 프로그램을 포함하고 있다. SPSS의 경우 요인분석의 기본방법으로 주성분분석을 제공하고 있으며, 적합도 지수를 포함하고 있지 않고, 다양한 회전법을 제공하고 있으나 표적회전법을 제공하지 않고 있다. 따라서 이러한 소프트웨어를 사용하여 요인분석을 하는 것은 바람직하지 않다. SAS의 경우는 평행분석, 표적회전, 추정치에 대한 표본오차 등을 제공하고 있어서 추천할 수 있는 소프트웨어이나 모형의 평가에 대한 통계치를 충분하게 제공하지 못 하며 여전히 성분분석결과의 형식을 따르고 있는 단점이 있다.

그밖에 요인분석을 위하여 특별히 개발된 소프트웨어가 있다. 그 중 최근에 가장 심리측정이론가들에 의해 선호되는 프로그램이 CEFA(Comprehensive Exploratory Factor Analysis; Browne et al. 2004)이다. 이 프로그램은 모형적합도 지수를 제공하며, 표적회전방식을 포함하고, 서열척도를 위한 다분상관

계수를 다룰 수 있다. 특히 이 소프트웨어는 요인부하량과 요인점수들 간의 상관계수에 대한 표준오차를 계산하는 절차가 포함되어 있어서, 이들에 대한 검증통계치와 신뢰구간을 제공한다. 또한 이 프로그램은 공개프로그램이다.

그 다음으로 MPLUS 프로그램은 서열척도와 이분척도의 분석 등 다양한 방식의 요인분석을 제공한다. 특히 다중대치법에 의한 결측치의 처리를 할 수 있는 현재까지 유일한 프로그램이다. 이 프로그램은 원래 구조화된 방정식모형을 처리하기 위하여 개발되었으나, 탐색적 요인분석도 수행할 수 있다.

Ⅹ. 맺는말

이 논문에서는 요인분석을 수행하는 과정에서 저지르기 쉬운 오류를 나열하고 이때 어떠한 선택이 합리적인지에 대하여 논의하였다. 흔히 저지르는 오류는 요인분석 대신에 주성분분석을 이용하는 것이다. 이 두 분석법은 통계적으로 동일한 기법이 아니기 때문에 목적에 맞게 사용되어야 한다. 또 다른 문제는 요인분석을 하면서 주성분분석처럼 해석하는 경우이다. 주성분분석과는 달리, 요인분석에서 표본공분산 행렬의 고유치의 크기로 소위 설명분산을 계산하는 것이 크게 의미가 없지만, 요인분석모형을 평가하는 데 흔히 이 설명분산를 사용한다. 더 나아가서는 회전 후에는 요인과 각 고유치(설명분산)는 상관이 없는데도 불구하고 이 둘을 짝짓기도 한다. 지난 20여 년 동안 이 분야에서 많은 개선이 이루어졌지만, 여전히 이러한 잘못된 관행은 계속되는 것으로 보인다.

이밖에도 요인분석에서 반드시 고려되어야 할 요소들로, 적합도 지수의 선택, 요인의 수의 선택 그리고 결측치의 처리방법이 있다. 이에 대한 적절한 선택들이 논의되었다. 연구의 목적과 자료의 특성에 따라 다양한 방식으로 요인분석이 수행될 수 있기 때문에 가장 좋은 방식의 분석법이 있을 수는 없다. 그러나 사회과학 분야에서 행해지는 전형적인 연구에서 추천되는 분석방법을 굳이 추천하자면 다음과 같다. 자료의 적합방식은 최대우도방법이나 일반화된

최소제곱법을 사용한다. 범주형 척도를 포함하는 경우에는 다분상관계수를 사용하여 ADF에 의한 적합방법을 선택하는 것이 바람직하다. 요인의 수에 대한 판단은 고유치나 평행분석, 스크리 검사, 적합도 지수 등을 이용하여 후보요인의 수를 결정한 다음, 회전된 요인부하량을 연구자가 검토하여 해석 가능성이 가장 높은 해를 선택한다. 회전방법은 사교회전을 기본으로 이용하며 표적회전법을 적극적으로 이용한다. 마지막으로 설명분산과 같은 주성분분석의 잔재는 요인분석에서는 사용하지 않는 것이 바람직하다.

참고문헌

김청택. 2008. “심리학 논문에 나타난 요인분석의 문제점.” 미발표 자료.

이순묵. 1994. “요인분석의 관행과 문제점.” ≪한국심리학회지: 산업 및 조직≫ 7: 1-27.

Amemiya, Y. and T.W. Anderson. 1990. “Asymptotic Chi-Square Tests for a Large Class of Factor Analysis Models.” *Annals of Statistics* 18: 1453-1463.

Anderson, T.W. and Y. Amemiya. 1988. “The Asymptotic Normal Distribution of Estimators in Factor Analysis under General Conditions.” *Annals of Statistics* 16: 759-771.

Bartholomew, D.J 2005. “History of Factor Analysis: A Statistical Perspective.” Encyclopedia of Statistics in Behavioural Science Vol.2.: 851-858, John Wiley and Sons Limited.

Browne, M.W. 1972a. “Orthogonal Rotation to a Partially Specified Target.” *British Journal of Mathematical and Statistical Psychology* 25: 115-120.

Browne, M.W. 1972b. “Oblique Rotation to a Partially Specified Target.” *British Journal of Mathematical and Statistical Psychology* 25: 207-212.

Browne, M.W. 1973. “Generalized Least Squares Estimators in the Analysis of Covariance Structures." *ETS Research Bulletin Series* 1973: 1–36.

Browne, M.W. 1982. "Covariance Structure." In D.M. Hawkins(ed.). *Topics in Applied Multivariate Analysis*: 71-141. Cambridge: Cambridge University Press.

Browne, M.W. 1984. "Asymtotically Distribution Free Methods for the Analysis of Covariance Structures." *British Journal of Mathematical and Statistical Psychology* 37: 62-83

Browne, M.W., and R. Cudeck. 1993. "Alternative Ways of Assessing Model Fit." In K.A. Bollen & J.S. Long(eds.). *Testing structural equation models*(pp. 136-162). Beverly Hills, CA: SAGE

Browne, M.W., R. Cudeck, K. Tateneni, and G. Mels. 2004. *CEFA: Comprehensive Exploratory Factor Analysis*, Version 2.00 [Computer software and manual]. Retrieved from http://quantrm2.psy.ohio-state.edu/browne/.

Browne, M.W., R.C. MacCallum, C. Kim, B.L. Anderson, and R. Glaser. 2002. "When Fit Indices and Residuals Are Incompatible." *Psychological Methods* 7: 403-421.

Browne, M.W. and A. Shapiro. 1988. "Robustness of Normal Theory Methods in the Analysis of Linear Latent Variate Models." *British Journal of Mathematical and Statistical Psychology* 41: 193-208.

Cattell, R.B. 1966. "The Scree Test for the Number of Factors." *Multivariate Behavioral Research* 1: 245-276.

Cattell, R.B. and S. Vogelmann. 1977. "A Comprehensive Trial of the Scree and KG Criteria for Determining the Number of Factors." *Journal of Educational Measurement* 14: 289-325.

Cronbach, L.J. and P. Meehl. 1955. "Construct Validity in Psychological Tests." Psychological Bulletin 52:281-302.

Fabrigar, L.R., D.T. Wegener, R.C. MacCallum, and E.J. Strahan. 1999. "Evaluating the Use of Exploratory Factor Analysis in Psychological Research." Psychological Methods 4: 272-299.

Floyd, F.J. and K.F. Widaman. 1995. "Factor Analysis in the Development and Refinement of Clinical Assessment Instruments." *Psychological Assessment* 7: 286-299.

Graham, J.W., P.E. Cumsille, and E. Elek-Fisk. 2003. "Methods for Handling Missing Data." In J.A. Schinka & W.F. Velicer(eds.). *Research Methods in Psychology*(pp. 87-114). Volume 2 of Handbook of Psychology (I.B. Weiner, Editor in Chief). New York: John Wiley & Sons.

Guttman, L. 1954. "Some Necessary Conditions for Common Factor Analysis." *Psychometrika* 19: 149-161.

Hakstian, A.R., W.T. Rogers, and R.B. Cattell. 1982. "The Behavior of Number of Factors Rules with Simulated Data." *Multivariate Behavioral Research* 17: 193-219.

Hart, B. and C. Spearman. 1912. "General Ability, Its Existence and Nature." *British Journal of Psychology* 5: 51-84.

Hayton, J.C., D.G. Allen, and V. Scarpello. 2004. "Factor Retention Decisions in Exploratory Factor analysis: A Tutorial on Parallel Analysis." *Organizational Research Methods* 7: 191-205.

Hotelling, H. 1933. "Analysis of a Complex of Statistical Variables into Principal Components." *Journal of Educational Psychology* 24: 321-377.

Horn, J.L. 1965. "A Rationale and Test for the Number of Factors in Factor Analysis." *Psychometrika* 30: 179-185.

Hu, L.T. and P.M. Bentler. 1999. "Cutoff Criteria for Fit Indexes in Covariance Structure Analysis: Conventional Criteria versus New Alternatives." *Structural Equation Modeling* 6: 1-55.

Kaiser, H.F. 1960. "The Application of Electronic Computers to Factor Analysis." *Educational and Psychological Measurement* 20: 141-151.

Kaiser, H.F. 1961. "A Note on Guttman's Lower Bound for the Number of Common Factors." *British Journal of Statistical Psychology* 14: 1-2.

Ledesma, R.D. and P. Valero-Mora. 2007. "Determining the Number of Factors to Retain in EFA: An Easy-to-use Computer Program for Carrying out Parallel Analysis." *Practical Assessment, Research & Evaluation* 12: 1-11.

Linn, R.L. 1968. "A Monte Carlo Approach to the Number of Factors Problem." *Psychometrika* 33: 37-71.

Little, R.J.A. and D.B. Rubin. 2002. *Statistical Analysis with Missing Data* (2nd ed.). New York: Wiley.

Margenau, H. 1955. "The Competence and Limitations of Scientific Method." *Journal of Operations Research Society of America* 3: 135-146.

Pearson, K. 1901. "On Lines and Planes of Closest Fit to Systems of Points in Space." *The London, Edinburgh and Dublin Philosophical Magazine and Journal of Science, Sixth Series* 2: 559-772.

Pigott, T.D. 2001. "A Review of Methods for Missing Data." *Educational Research and Evaluation* 7: 353-383.

Preacher, K.J. and R.C. MacCallum. 2003. "Repairing Tom Swift's Electric Factor Analysis Machine." *Understanding Statistics* 2: 13-43

Preacher, K.J., G. Zhang, C. Kim, and G. Mels. 2013. "Choosing the Optimal Number of Factors in Exploratory Factor Analysis: A Model Selection Perspective." *Multivariate Behavioral Research* 48: 28-56.

Rubin, D.B. 1976. "Inference and Missing Data." *Biometrika* 63: 581-592.

Schafer, J.L. and J.W. Graham. 2002. "Missing Data: Our View of the State of the Art." *Psychological Methods* 7: 147-177.

Shapiro, A. 1987. "Robustness Properties of the MDF Analysis of Moment Structures." *South African Statistical Journal* 21: 39-62.

Spearman, C. 1904. "General Intelligence: Objectively Determined and Measured." *American Journal of Psychology* 15: 201-293.

Thurstone, L.L. 1935. *The Vectors of Mind: Multiple Factor Analysis for the Isolation of Primary Traits*. Chicago: University of Chicago Press.

Thurstone, L.L. 1938. "Primary Mental Abilities." *Psychometric Monograph* 1. Chicago: University of Chicago Press.

Yuan, K-H and P.M. Bentler. 1999. “On Asymptotic Distributions of Normal Theory MLE in Covariance Structure Analysis under Some Nonnormal Distributions.” *Statistics and Probability Letters* 42: 107-113.

Zwick, W.R. and W.F. Velicer. 1982. “Factors Influencing Four Rules for Determining the Number of Components to Retain.” *Multivariate Behavioral Research* 17: 253-269.

Zwick, W.R. and W.F. Velicer. 1986. “Comparison of Five Rules for Determining the Number of Components to Retain.” *Psychological Bulletin* 99: 432-442.

<접수 2016/1/31, 수정 2016/2/17, 게재확정 2016/2/22>