

# 8장 인자분석

덕성여자대학교 정보통계학과

김 재 희



덕성여자대학교  
DUKSUNG WOMEN'S UNIVERSITY

## 8.1 서론

▶ 인자분석(factor analysis)은 공분산 구조를 몇 개의 관측 불가능한 '인자(factor)'로 설명하려는 것으로 변수들 간에 내재하고 있는 공통의 구조를 파악하고, 데이터의 특성을 몇 개의 인자로 축약하여 설명하고자 하는 것이 인자분석의 목적이라 할 수 있다.

▶ 변수들에 내재되어 있는 공통의 인자로서 데이터를 설명함으로써 분석에 필요한 변수의 차원을 줄일 수 있다.

▶ 인자분석은 특히 경영학, 사회학 관련분야에서 활발히 사용.

이러한 분야에서는 변수의 개수가 많고 관계가 연계되어 있어, 변수간의 구조(structure)와 변수 내 연관성(interrelationship of the variables)에 대한 집중적 해석이 매우 중요.

▶ 인자분석은 Spearman(1904)에 의해 시작. **단일인자모형** 고려.

Garnett(1919), Thurstone(1931) 등에 의해 여러 개의 인자를 동시에 고려하는 **다중 인자모형 (multiple common factor model)**으로 확장.

## 8.2 다중 인자모형

$p \times 1$  확률벡터  $X = (X_1, X_2, \dots, X_p)'$ 가 다음과 같이 모평균벡터  $\mu$ 와 모공분산행렬  $\Sigma$   $X$ 는  $m$ 개의 공통인자(common factor)  $F_1, F_2, \dots, F_m$ 과 특수인자(specific factor)  $\epsilon_1, \epsilon_2, \dots, \epsilon_p$ 의 선형결합으로 표현된다고 가정한다:

$$\begin{aligned} X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \dots + l_{1m}F_m + \epsilon_1 \\ X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \dots + l_{2m}F_m + \epsilon_2 \\ &\vdots \\ X_p - \mu_p &= l_{p1}F_1 + l_{p2}F_2 + \dots + l_{pm}F_m + \epsilon_p \end{aligned}$$

여기서  $l_{ij}$ 는  $i$ 번째 변수의  $j$ 번째 인자의 적재값이며  $F_1, \dots, F_m$ 와  $\epsilon_1, \dots, \epsilon_p$ 는  $p + m$  개의 관측 불가능한(unobservable) 확률변수이다. 이와 같이  $m$ 개의 공통인자를 포함한 인자모형을  $m$ -인자모형이라고 부르기로 한다.

## 8.2.1 직교인자모형

$m$ -인자 직교인자모형(orthogonal common factor model):

$$X - \mu = LF + \epsilon$$

$$p \times 1 \quad (p \times m)(m \times 1) \quad p \times 1$$

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}, \quad F = \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{pmatrix}, \quad L = \begin{pmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{pmatrix}, \quad \epsilon = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_p \end{pmatrix}$$

여기서  $\mu_i$ 는  $X_i$ 의 기대값

$\epsilon_i$ 는  $i$ 번째 특수인자(specific factor)

$F_j$ 는  $j$ 번째 공통인자(common factor)

$l_{ij}$ 는  $i$ 번째 변수의  $j$ 번째 인자의 적재값

$L = \{l_{ij}\}$ 은 인자적재행렬(matrix of factor loadings)

## 직교인자모형에서의 가정(assumption)

$$E(F) = \mathbf{0}_{m \times 1}$$

$$Cov(F) = E(FF') = I_{m \times m}$$

$$E(\epsilon) = \mathbf{0}_{p \times 1}$$

$$Cov(\epsilon) = E(\epsilon\epsilon') = \Psi_{p \times p} = \begin{pmatrix} \psi_1 & 0 & \cdots & 0 \\ 0 & \psi_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_p \end{pmatrix}$$

$F$ 와  $\epsilon$ 는 서로 독립이다(즉 인자들이 서로 직교한다). 즉

$$Cov(\epsilon, F) = E(\epsilon F') = \mathbf{0}_{p \times m}$$

이와 같은 가정을 만족하는 모형을 직교인자모형이라고 한다.

예를 들어, 변수가 5개( $p = 5$ )이고 공통인자가 2개( $m = 2$ )인 경우 인자 모형은 다음과 같이 표현될 수 있다:

$$\begin{aligned}X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \epsilon_1 \\X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \epsilon_2 \\X_3 - \mu_3 &= l_{31}F_1 + l_{32}F_2 + \epsilon_3 \\X_4 - \mu_4 &= l_{41}F_1 + l_{42}F_2 + \epsilon_4 \\X_5 - \mu_5 &= l_{51}F_1 + l_{52}F_2 + \epsilon_5.\end{aligned}$$

## 8.2.2 인자모형에서 공분산행렬의 분해

$$\begin{aligned}\blacktriangleright (X-\mu)(X-\mu)' &= (LF+\epsilon)(LF+\epsilon)' \\ &= (LF+\epsilon)((LF)'+\epsilon') \\ &= LF(LF)' + \epsilon(LF)' + LF\epsilon' + \epsilon\epsilon'\end{aligned}$$

$$\begin{aligned}\Sigma = Cov(X) &= E(X-\mu)(X-\mu)' \\ &= LE(FF')L' + E(\epsilon F')L' + LE(F\epsilon') + E(\epsilon\epsilon') \\ &= LL' + \Psi\end{aligned}$$

: 공분산행렬을 인자적재행렬과 특수인자의 공분산행렬로 인자분해된다.

▶ 변수와 공통인자간의 공분산을 구하면

$$Cov(X, F) = E[(X-\mu)F'] = LE(FF') + E(\epsilon F') = L$$

▶ 원래 변수의 분산은 인자적재값들로 표현된다.

$$\begin{aligned}\sigma_{ii} = Var(X_i) &= l_{i1}^2 + \dots + l_{im}^2 + \psi_i \\ &= h_{ii}^2 + \psi_i \\ &= \text{공통분산} + \text{특수분산} ,\end{aligned}\quad i = 1, 2, \dots, p$$

- ▶ 공통성(communality) 또는 공통분산(common variance):

$$h_i^2 = l_{i1}^2 + \dots + l_{im}^2$$

- ▶  $\psi_i$  : 특수성 또는 특수분산(specific variance)

- ▶ 변수들간의 공분산은  $i = 1, 2, \dots, p, \quad k = 1, 2, \dots, p, \quad i \neq j$ 에 대해

$$Cov(X_i, X_k) = l_{i1}l_{k1} + \dots + l_{im}l_{km},$$

- ▶ 변수와 공통인자간의 공분산은

$$Cov(X_i, F_j) = l_{ij}$$



### 8.2.3 인자적재행렬의 비유일성

인자의 개수가  $m > 1$  인 경우 인자적재행렬은  $m \times m$  직교행렬  $T$  를 이용하여 표현.  
직교행렬  $T$  는  $TT' = T'T = I$  를 만족하므로

$$X - \mu = LF + \epsilon = LTT'F + \epsilon = L^*F^* + \epsilon$$

여기서  $L^* = LT$

$$F^* = T'F$$

$$E(F^*) = T'E(F) = 0$$

$$\text{Cov}(F^*) = T'\text{Cov}(F)T = T'T = I_{m \times m}$$

분산을 공분산행렬  $L^*$ 를 이용해 표현하면

$$\Sigma = LL' + \Psi = LTT'L' + \Psi = L^*L^{*'} + \Psi$$

직교행렬  $T$  형태에 따라 인자적재행렬은 변화가능, 인자적재행렬은 항상 유일하지는 않다.

## 8.2.4 인자모형의 척도불변성

확률벡터  $X$ 에, 정칙행렬  $A$ 를 이용하여 다음과 같이 선형변환을 했을 때

$$Z = AX + b$$

다중인자모형은

$$Z - \mu_Z = AX - A\mu = ALF + A\epsilon = L_Z F + \epsilon_Z$$

여기서  $\mu_Z = A\mu + b$

$$L_Z = AL$$

$Z$ 의 공분산행렬을 구하면

$$\text{Cov}(Z) = \Sigma_Z = A\Sigma A' = ALL'A + A\Psi A' = L_Z L_Z' + \Psi_Z$$

여기서  $\Psi_Z = A\Psi$ .

: 인자적재행렬에도 마찬가지로 변환이 적용된다.

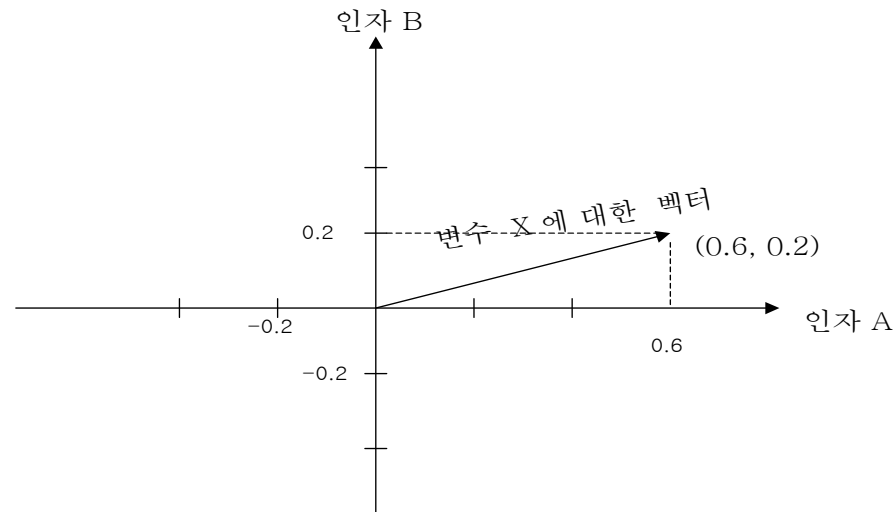
## 8.2.5 인자모형의 기하학적 표현

▶ 인자를 축(axis)으로 변수를 선(line)/벡터(vector)로 표현하고 좌표평면에 그림.

《예제 8.1》 두 개의 인자적재값이 각각 0.6, 0.2인 변수의 기하학적인 표현에 대해 알아보자.

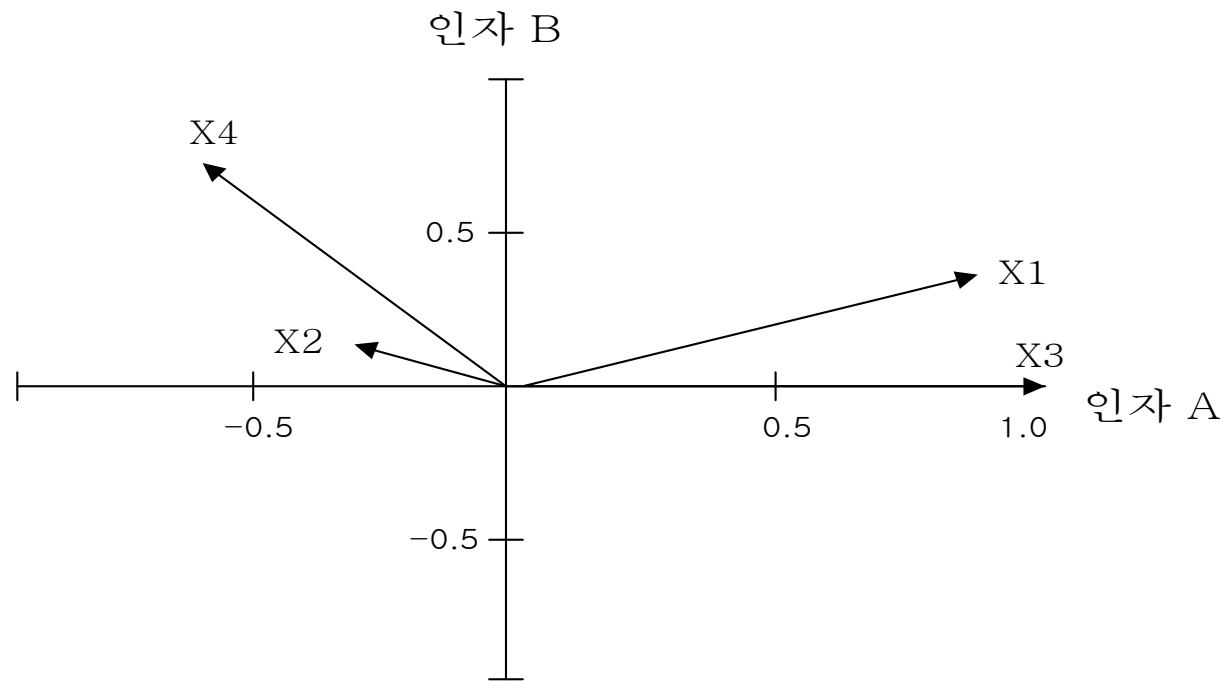
인자 A에 대한 비중은 0.6이고 인자 B에 대한 비중은 0.2이다.

변수 벡터의 길이  $h^2 = 0.6^2 + 0.2^2 = 0.4$  : 공통인자에 의한 분산(공통성)



[그림 8.1] 두 개의 직교인자와 한 개 변수에 대한 그림

《예제 8.2》 4개 변수가 2 개의 공통인자 A, B로 표현되는 인자모형에서 인자적재값을 이용하여 그래프로 그린 결과 [그림 8.2]와 같을 때 그림이 나타내는 바를 알아보자.  
(벡터 길이 = 공통성)



[그림 8.2] 2 개의 직교인자와 4 개 변수에 대한 그림

## 8.3 인자적재와 공통성에 대한 추정방법

(1) 주성분법 (2) 주축인자법 (3) 최대우도법

### 8.3.1 주성분법

공분산행렬을 스펙트럼분해하면

$$\begin{aligned}\Sigma &= \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \cdots + \lambda_p e_p e_p' \\ &= [\sqrt{\lambda_1} e_1 | \sqrt{\lambda_2} e_2 | \cdots | \sqrt{\lambda_p} e_p] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \sqrt{\lambda_2} e_2' \\ \vdots \\ \sqrt{\lambda_p} e_p' \end{bmatrix}\end{aligned}$$

으로 표현할 수 있다. 여기서  $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$  는  $\Sigma$ 의 고유값  $e_i$ 는  $\Sigma$ 의 고유벡터로  $\Sigma e_i = \lambda_i e_i$ 을 만족한다.

만약  $m = p$ 이면 즉  $\psi_i = 0, i = 1, \dots, p$  라 하면

$$\sum_{p \times p} = LL' + 0 = LL'$$

이 되어 인자적재값과 주성분계수의 값이 일치하게 된다.

근사적으로  $\Sigma$ 를  $m$ 개의 인자로 표현할 수 있어

$$\begin{aligned} \Sigma &\approx \lambda_1 e_1 e_1' + \lambda_2 e_2 e_2' + \dots + \lambda_p e_m e_m' \\ &= [\sqrt{\lambda_1} e_1 \mid \sqrt{\lambda_2} e_2 \mid \dots \mid \sqrt{\lambda_m} e_m] \begin{bmatrix} \sqrt{\lambda_1} e_1' \\ \vdots \\ \sqrt{\lambda_m} e_m' \end{bmatrix} = LL'_{(p \times m) \ (m \times p)} \end{aligned}$$

공분산행렬이  $\Sigma = LL' + \Psi$  이므로

특수성:  $\psi_i = \sigma_{ii} - \sum_{j=1}^m l_{ij}^2, \quad i = 1, \dots, p$

- ▶ 표본공분산행렬  $S = \hat{\Sigma}$ 을 이용한 주성분법에 의한 인자적재행렬 추정방법  
 $S$ 가 고유값  $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \dots \geq \hat{\lambda}_p \geq 0$ , 고유벡터  $e_1, e_2, \dots, e_p$ 를 가진다고 할 때

$$S = \tilde{L}\tilde{L}' + \tilde{\Psi}$$

$$\tilde{L} = [\sqrt{\hat{\lambda}_1} e_1 | \sqrt{\hat{\lambda}_2} e_2 | \dots | \sqrt{\hat{\lambda}_m} e_m] \quad \tilde{\Psi}_{p \times p} = \begin{pmatrix} \tilde{\psi}_1 & 0 & \dots & 0 \\ 0 & \tilde{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{\psi}_p \end{pmatrix}$$

$$\text{공통성: } \tilde{h}_i^2 = \tilde{l}_{i1}^2 + \tilde{l}_{i2}^2 + \dots + \tilde{l}_{im}^2 \quad \text{특수성: } \tilde{\psi}_i = s_{ii} - \tilde{h}_i^2$$

- ▶  $i$ 번째 인자에 의해 설명되는 분산 비율:  $\frac{\hat{\lambda}_i}{s_{11} + s_{22} + \dots + s_{pp}}$

- ▶ 표본상관행렬  $R$ 을 사용했을 경우  $i$ 번째 인자에 의해 설명되는 분산 비율:  $\frac{\hat{\lambda}_i}{p}$

### 8.3.2 주축인자법

주성분법에 의해 공통성 또는 특수성을 추정한 후 이에 대한 정보를 이용하여 다시 인자적재행렬을 구한다.

$p \times p$ 인 공분산행렬, 상관행렬로부터 주성분법의 결과를 이용하여 조정된(adjusted) 공분산행렬, 조정된 상관행렬에 대해 각각 다음의 근사적인 관계식

$$S - \hat{\Psi} \approx LL'$$

$$R - \hat{\Psi} \approx LL'$$

여기서  $L$ 은  $p \times m$ 인 인자적재행렬.

$$S - \hat{\Psi} \text{의 대각선상의 성분은 공통성 : } \tilde{h}_i^2 = s_{ii} - \tilde{\psi}_i, \quad i = 1, 2, \dots, p$$

$$R - \hat{\Psi} \text{의 대각선상의 성분은 공통성: } \tilde{h}_i^2 = 1 - \tilde{\psi}_i, \quad i = 1, 2, \dots, p$$



$R - \hat{\Psi}$ 에 대해서, 공통성에 대한 초기추정량은

$$\tilde{h}_i^2 = R_i^2 = 1 - \frac{1}{r^{ii}}$$

여기서  $r^{ii}$ 는  $R^{-1}$ 의  $i$ 번째 대각선상의 성분이며  $R_i^2$ 은  $X_i$ 와 나머지  $(p-1)$ 개의 변수들과의 다중상관계수제곱(squared multiple correlation: SMC)이 된다.

마찬가지로  $S - \hat{\Psi}$ 에 대해서, 공통성에 대한 초기추정량은

$$\tilde{h}_i^2 = s_{ii} - \frac{1}{s^{ii}} = s_{ii} R_i^2$$

로 쓸 수 있다. 여기서  $s^{ii}$ 는  $S^{-1}$ 의  $i$ 번째 대각선상의 성분이다.

이와 같은 과정을, 공통성의 추정치가 수렴할 때까지, 반복한다.

데이터에 따라서는 반복적 계산에도 불구하고 공통성이 수렴하지 않을 경우도 있으며 계산상의 문제로 인해 공통성이  $s^{ii}$  값보다 커질 경우도 있으므로 결과를 주의해서 볼 필요가 있다.

### 8.3.3 최대우도법

- ▶ 공통인자  $F$ 와 특수인자  $\epsilon$ 이 정규분포를 따른다고 가정
- ▶  $n$ 개의 확률표본  $X_1, \dots, X_n$ 에 대한 우도함수

$$\begin{aligned}
 & L(\mu, \Sigma | X_1, \dots, X_n) \\
 &= (2\pi)^{-\frac{np}{2}} |\Sigma|^{-\frac{n}{2}} \times \exp \left\{ -\frac{1}{2} \text{tr} \left[ \Sigma^{-1} \left( \sum_{j=1}^n (X_j - \bar{X})(X_j - \bar{X})' + n(\bar{X} - \mu)(\bar{X} - \mu)' \right) \right] \right\} \\
 &= (2\pi)^{-\frac{(n-1)p}{2}} |\Sigma|^{-\frac{(n-1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left[ \Sigma^{-1} \sum_{j=1}^n (X_j - \bar{X})(X_j - \bar{X})' \right] \right\} \\
 &\quad \times (2\pi)^{-\frac{p}{2}} |\Sigma|^{-\frac{1}{2}} \exp \left\{ -\frac{n}{2} (\bar{X} - \mu)' \Sigma^{-1} (\bar{X} - \mu) \right\}
 \end{aligned}$$

$\Sigma = LL' + \Psi$  이므로  $L$ 에 대한 유일성 조건

$$L' \Psi^{-1} L = \Delta \quad (\text{대각행렬})$$

하에서 수치해석알고리즘 활용한 반복계산통해  $L$ 과  $\Psi$ 의 최대우도추정량  $\hat{L}, \hat{\Psi}$ 를 구한다.

### 8.3.4 Heywood 상황

- ▶ 인자에 대한 공통분산은 다중상관계수의 제곱이 되므로 그 범위는 0에서 1 사이가 된다.
- ▶ 최대우도법 이용시 반복적 수치적 계산과정때문에 공통성이 1 이상인 현상이 나타난다.

\* Heywood 상황 : 공통성이 1인 경우

\* ultra-Heywood 상황 : 공통성이 1을 초과하는 경우

=> 특수분산의 추정값이 0 또는 음수(-)가 되어 인자분석의 결과가 유효하지 않게 된다.

#### 이러한 상황이 발생하는 원인

- (1)공통성에 관한 부적절한 초기 추측값
- (2)공통인자수가 너무 많거나
- (3)공통인자수가 너무 적거나
- (4)안정적인 추정값을 계산하기에 부족한 양의 데이터가 있거나
- (5)공통인자모형 적합이 부적절한 데이터 등

## 8.4 인자개수 결정

### (1) 총 변이에의 공헌도(percentage of total variance)

공통인자들에 의해 설명되는 분산비율이 전체 변이의 70 ~ 90%가 되도록 정한다.

### (2) 평균 고유값(average eigenvalues)

고유값들의 평균

$$\bar{\lambda} = \sum_{i=1}^p \frac{\lambda_i}{p}$$

을 구한 후 고유값이 평균값 이상이 되는 주성분을 보유한다.

상관행렬을 사용한 경우 평균 고유값은 1 이 된다.

### (3) 스크리 그래프

2차원 좌표축에 (고유값순서, 고유값크기)로  $(i, \lambda_i)$  점을 찍고 점간을 선분으로 연결한다.

가파른 정도를 보고 가파른 부분에 해당하는 고유값까지로 인자의 개수를 정한다.

#### (4) 인자 개수에 대한 검정

관심 있는 귀무가설은 인자의 개수가  $m$ 개를 만족하는 모형이다.

$$H_0 : \Sigma = LL' + \Psi$$

( $L$  은  $p \times m$  행렬 즉 인자가  $m$ 개인 인자모형을 만족한다.)

$$H_1 : \Sigma \neq LL' + \Psi$$

데이터에 대해 다변량 정규분포 가정을 할 수 있는 경우, 우도비 이용

$$LR = \left( \frac{|\hat{\Sigma}|}{|\hat{\Sigma}_0|} \right)^{\frac{n}{2}}$$

여기서  $H_0$  하에서 최대우도추정량  $\hat{\Sigma}_0 = \hat{L}\hat{L}' + \hat{\Psi}$  과

전체 가설 공간  $H_0 \cup H_1$  하에서의 최대우도추정량  $\hat{\Sigma}$ 를 구한 후 우도비를 이용.

Bartlett(1951)은 다음의 검정통계량

$$T_m = \left( n - \frac{2p + 4m + 11}{6} \right) \ln \left( \frac{|\hat{L}\hat{L}' + \hat{\Psi}|}{|S|} \right)$$

$H_0$  하에서 근사적으로 자유도  $v = \frac{1}{2} [(p-m)^2 - p - m]$  인 카이제곱분포를 따른다

유의수준  $\alpha$ 에서 검정법은

$$T_m \geq \chi_v^2(\alpha) \text{이면 } H_0 \text{를 기각한다.}$$

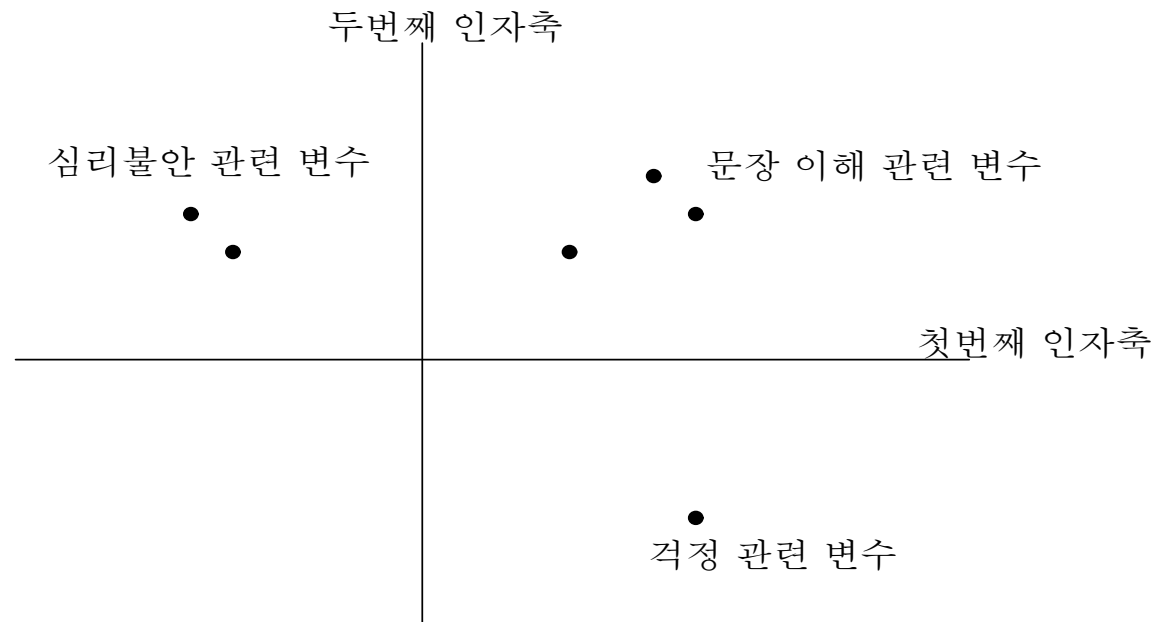
$H_0$ 가 기각되지 않으면  $\hat{L}\hat{L}' + \hat{\Psi}$ 은 적절한 인자모형이 되므로 모형에서 가정한 인자의 개수로 인자모형을 설정하고 해석하면 된다.

$H_0$ 가 기각되면  $\hat{L}\hat{L}' + \hat{\Psi}$ 은 부적절한 인자모형이 되므로 인자의 개수를 늘려 다시 검정한다.

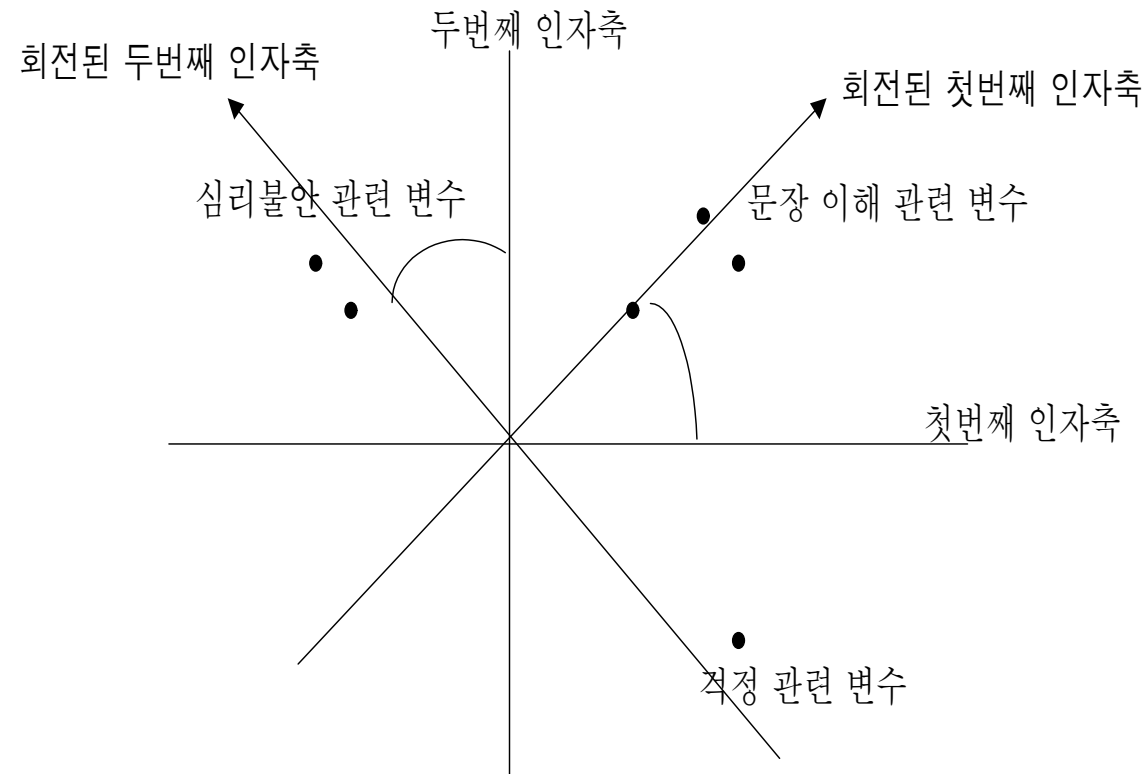
## 8.5 인자적재의 회전

적절한 변환을 통해 인자에 대한 해석이 용이한 단순한 구조를 만들 필요가 있다.

회전변환의 기하학적 목적은 **인자적재값이 인자축에 놓이게 하는 것**이다. 인자의 축에 놓이는 변수 즉 그 인자에서만 비중인 높은 변수들을 파악해 인자를 해석할 수 있게 된다.

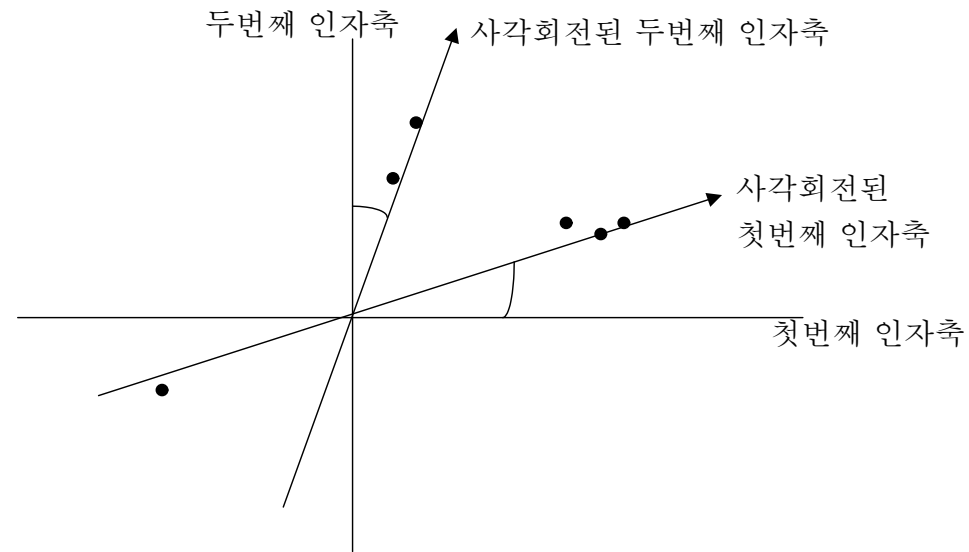


[그림 8.3] 개인 심리검사에 대한 6개 변수와 2개 인자축

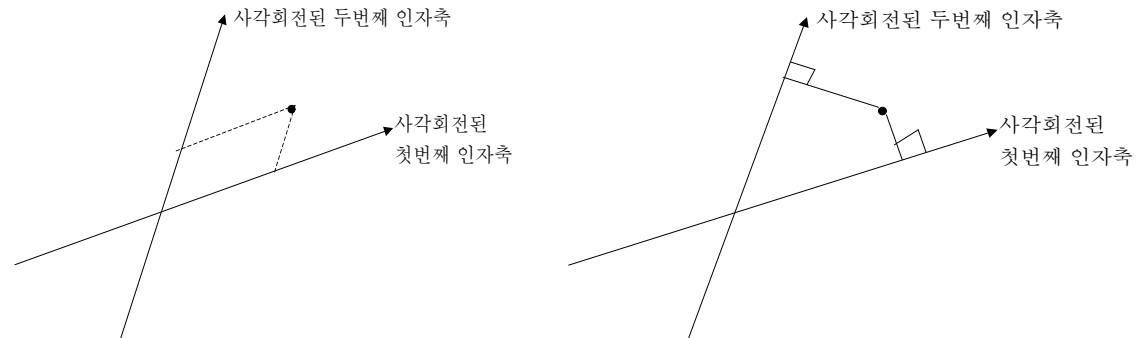


[그림 8.4] 개인 심리검사에 대한 4 개 변수와 직교 회전된 인자축





[그림 8.5] 5개 변수와 2개 사각회전된 인자축



[그림 8.6] 인자패턴(factor pattern)과 인자구조(factor structure)

## 8.5.1 직교회전

### (1) 베리맥스 회전(Varimax Rotation)

- ▶ Kaiser(1958)
- ▶ 인자적재값들의 제곱을 취해 이들의 분산을 최대화

$$\text{Maximize } \sum_{j=1}^m (\hat{l}_{ij}^{*2} - \overline{\hat{l}_i^{*2}})^2 = \left[ \sum_{i=1}^p \hat{l}_{ij}^{*4} - \left( \sum_{i=1}^p \hat{l}_{ij}^{*2} \right)^2 / p \right], \quad j = 1, \dots, m$$

$$\text{여기서 } \overline{\hat{l}_j^{*2}} = \sum_{i=1}^p \hat{l}_{ij}^{*2} / p, \quad \hat{l}_{ij}^* = \frac{\hat{l}_{ij}}{h_i}$$

- ▶ 베리맥스방법은

$$V = \frac{1}{m} \sum_{j=1}^m \left[ \sum_{i=1}^p \hat{l}_{ij}^{*4} - \left( \sum_{i=1}^p \hat{l}_{ij}^{*2} \right)^2 / p \right]$$

를 최대화하는 직교행렬  $T$ 를 찾아 변환하는 방법이다.

## (2) 쿼티맥스 회전(Quartimax Rotation)

▶ 인자적재행렬의 각 행에 놓인 적재값 즉 한 변수가 각 공통인자에서 차지하는 비중의 제곱에 대해 분산을 최대화하도록 변환하는 방법

$$\text{Maximize } \sum_{j=1}^m (\hat{l}_{ij}^2 - \hat{l}_i^2)^2, \quad i = 1, 2, \dots, p$$

여기서  $\hat{l}_i^2 = \sum_{j=1}^m \hat{l}_{ij}^2 / m$  이다.

## 8.5.2 비직교회전

- ▶ 해석의 편리를 위해 비직교회전(nonorthogonal: oblique rotation)
- ▶ 사각회전을 통해 비록 회전된 축이 서로 직교하지는 않으나 축 위에 인자적재값들이 놓이게 될 수 있어 해석상 이점을 얻을 수 있다.
- ▶ 사각회전에서 변환행렬  $Q$ : 비직교행렬(nonorthogonal), 비정칙(nonsingular) 행렬

$$F^* = Q' F$$

$$L^* = LQ$$

이고

$$Cov(F^*) = Q' I Q = Q' Q \neq I$$

으로 즉 비직교행렬에 의해 변환된 인자  $F^*$ 는 독립이 아니다.

비직교회전인 사각회전을 이용하는 방법으로 Oblimin, Oblimax, Covarmin, Quatimin, Biquartimin, Orthoblique, Promax 방법 등이 있다.

## 8.6 인자점수

공통인자를 파악한 후 이들에 대한 점수를 얻고자 할 때

### 8.6.1 가중최소제곱법

인자모형  $X - \mu = LF + \epsilon$ 에 대한 적합으로 인자적재행렬  $L$ 과 평균벡터  $\mu$  그리고  $Var(\epsilon_i) = \psi_i$  를 구한 후 이들의 추정값을 이용하여 인자점수를 추정할 수 있다.

Bartlett(1937): 가중최소제곱법을 제안

$$\begin{aligned}\sum_{i=1}^p \frac{\epsilon_i^2}{\psi_i} &= \epsilon' \Psi^{-1} \epsilon \\ &= (X - \mu - Lf)' \Psi^{-1} (X - \mu - Lf) \\ &= f' L' \Psi^{-1} Lf + f' L' \Psi^{-1} (X - \mu) + (X - \mu)' \Psi^{-1} fL \\ &\quad - (X - \mu)' \Psi^{-1} (X - \mu)\end{aligned}$$

을 최소화하도록

$$\hat{f} = (\hat{L}'\hat{\Psi}^{-1}\hat{L})^{-1}\hat{L}'\hat{\Psi}^{-1}(X - \hat{\mu})$$

을 구한다. 그러면 각 개체에 대해  $j = 1, \dots, n$

$$\begin{aligned}\hat{f}_j &= (\hat{L}'\hat{\Psi}^{-1}\hat{L})^{-1}\hat{L}'\hat{\Psi}^{-1}(X_j - \hat{\mu}) \\ &= \hat{\Delta}^{-1}\hat{L}'\hat{\Psi}^{-1}(X_j - \bar{X})\end{aligned}$$

를 이용하여 인자점수를 계산할 수 있다.

여기서  $\hat{\mu} = \bar{X}$ 를 이용하고 최대우도법의 경우에는 유일성조건

$$\hat{L}'\hat{\Psi}^{-1}\hat{L} = \hat{\Delta}$$

이 만족하도록 구한다.

Bartlett 방법은 인자점수와 인자가 높은 상관관계를 갖게 되며 추정된 인자점수는 불편추정량 (unbiased estimator)이다.

## 8.6.2 회귀적방법

인자들을 변수들의 다중회귀식으로부터 인자점수를 추정하는 방법

$$F = BX + \epsilon$$

최소제곱추정량은

$$\hat{B} = (X'X)^{-1}X'F$$

$$\hat{B} = S_{XX}^{-1}S_{XF}$$

로 구할 수 있다. 여기서  $S_{XX} = \widehat{Cov}(X)$ ,  $S_{XF} = \widehat{Cov}(X, F)$ .

인자들과 변수들이 다변량 정규분포를 따를 경우 조건부 기대값을 구하면

$$E(F|X) = \mu_F + \Sigma_{FX}\Sigma_{XX}^{-1}(X - \mu_X)$$

을 얻는다. 인자모형으로부터  $\mu_F = 0$ 이고  $\Sigma_{FX} = Cov(F, X) = L'$ 이므로

$$E(F|X) = L' \Sigma_{XX}^{-1} (X - \mu_X)$$

으로 쓸 수 있다.

$$\hat{f}_j = \hat{L}' S^{-1} (X_j - \bar{X}), \quad j = 1, 2, \dots, n$$

을 구할 수 있으며 전치된 형태(transposed)를 구하면

$$\hat{f}_j' = (X_j - \bar{X})' S^{-1} \hat{L}, \quad j = 1, 2, \dots, n$$

이 된다.



## 8.7 인자분석 전략

(1) 주성분을 이용한 인자분석을 한다. :

해석을 위해 베리맥스(varimax) 직교변환 등을 모색한다.

(2) 최대우도를 이용한 인자분석을 한다:

다변량 정규분포를 가정할 수 있다면 최대우도를 이용한 인자분석을 수행한 후 인자의 해석을 위해 베리맥스 직교변환을 모색한다.

(3) 주성분을 이용한 인자분석 결과와 최대우도를 이용한 결과를 비교한다

(4) 적절한 공통인자  $m$ 개에 대해 (1)(2)(3) 과정을 반복한다.

(5) 데이터의 크기가 큰 경우 데이터를 두 부분으로 나누어 각 부분에서 따로 인자 분석을 시행하고 결과를 비교해 인자 분석 결과의 안정성을 점검한다.

## (6) 인자의 해석:

(회전된) 인자적재행렬  $\hat{L}^*$ 의  $j$ 번째 열  $\hat{l}_{1j}^*, \hat{l}_{2j}^*, \dots, \hat{l}_{pj}^*$ 을 살펴보아 어떤 변수가 어떤 인자에서 높은 비중을 차지하는지를 본다. 각 인자에서 큰 적재값을 갖는 변수, 즉 높은 비중을 차지하는 변수들을 중심으로 인자에 대한 이름을 붙이고 해석한다.

《예제 8.3》 신체 측정 데이터에 대해 상관행렬을 이용하여 인자 분석을 하여 인자적재값, 공통성과 특수성을 구하여 [표 8.1]에 정리하였다.

공통인자인 인자1은 길이부분에 더 많이 내재하는 인자로 ‘길이 인자’로 볼 수 있으며 인자2는 무게와 둘레에 기여하는 부분이 큰 ‘무게 인자’로 볼 수 있다.

공통인자에 의해 설명되는 분산인 공통성은 각 변수에 대해 2개 인자의 인자적재값제곱의 합으로 계산되고, 특수인자에 의해 설명되는 분산은  $\psi_i = Var(X_i) - h_i^2 = 1 - h_i^2$ 으로 계산된다. 여기서 상관행렬을 이용했으므로  $Var(X_i) = 1$ 이다.

[표 8.1] 신체 측정 자료에 대한 인자분석 결과

변수	주성분법 인자1                  인자2		공통성 $h_i^2$	특수성 $\psi_i$
$X_1$ (신장)	1.00	0.00	1.00	0.00
$X_2$ (앞은키)	0.90	-0.12	0.82	0.18
$X_3$ (가슴둘레)	0.31	0.68	0.56	0.44
$X_4$ (흉골길이)	0.55	0.12	0.32	0.68
$X_5$ (머리둘레)	0.11	0.18	0.04	0.96
$X_6$ (엉덩이둘레)	0.34	0.62	0.50	0.50
$X_7$ (몸무게)	0.50	0.86	0.99	0.01

《예제 8.4》 100명 고등학생들의 성적에 대한 데이터의 상관행렬을 이용하여 인자분석을 하여 인자적재값, 공통성과 특수성을 구하여 [표 8.2]에 정리하였다.

[표 8.2] 과목성적에 대한 인자분석 결과

변수	주성분법		공통성	특수성
	인자1	인자2	$h_i^2$	$\psi_i$
$X_1$ (국어)	0.90	0.10	0.82	0.18
$X_2$ (영어)	0.73	0.12	0.55	0.45
$X_3$ (사회)	0.55	0.23	0.36	0.54
$X_4$ (물리)	0.23	0.87	0.32	0.68
$X_5$ (수학)	0.34	0.93	0.81	0.19
$X_6$ (생물)	0.36	0.50	0.38	0.62
$X_7$ (국사)	0.60	0.30	0.45	0.55

인자1은 국어, 영어, 사회, 국사 과목에 비중이 큰 인자로 ‘인문학 학습능력 인자’  
인자2는 수학, 물리, 생물의 과목에서의 비중이 높은 것으로 보아 ‘자연과학 학습능력 인자’

《예제 8.5》 엔지니어 견습생 20명과 비행조종사(pilot) 20명의 여섯 가지의 테스트 결과 비교 자료가 [표 8.3]에 정리 되어있다. 인자분석을 통해 변수에 내재하는 공통인자들을 찾고, 여러 가지 인자적재 추정방법을 사용해 비교하고자 한다.

$X_1$  = 지능(intelligence)

$X_2$  = 상황설명능력(form relations)

$X_3$  = 동력측정검력계(dynamometer)

$X_4$  = 상세 표시 능력(dotting)

$X_5$  = 지각 기구 좌표화 능력(sensory motor coordination)

$X_6$  = 인내력(perseveration)

[표 8.3] 엔지니어와 비행조종사 검사 자료

엔지니어 수습생						파일럿					
$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$	$X_6$
121	22	74	223	54	254	132	17	77	232	50	249
108	30	80	175	40	300	123	32	79	192	64	315
122	49	87	266	41	223	129	31	96	250	55	319
77	37	66	178	80	209	131	23	67	291	48	310
140	35	71	175	38	261	110	24	96	239	42	268
108	37	57	241	59	245	47	22	87	231	40	217
124	39	52	194	72	242	125	32	87	227	30	324
130	34	89	200	85	242	129	29	102	234	58	300
149	55	91	198	50	277	130	26	104	256	58	270
129	38	72	162	47	268	147	47	82	240	30	322
154	37	87	170	60	244	159	37	80	227	58	317
145	33	88	208	51	228	135	41	83	216	39	306
112	40	60	232	29	279	100	35	83	183	57	242
120	39	73	159	39	233	149	37	94	227	30	240
118	21	83	152	88	233	149	38	78	258	42	271
141	42	80	195	36	241	153	27	89	283	66	291
135	49	73	152	42	249	136	31	83	257	31	311
151	37	76	223	74	268	97	36	100	252	30	225
97	46	83	164	31	243	141	37	105	250	27	243
109	42	82	188	57	267	164	32	76	187	30	264

[표 8.4] 인자분석 결과 인자적재값 비교

변수	주성분법		주성분법 (Varimax)		주성분법 (Promax)	
	인자1	인자2	인자1	인자2	인자1	인자2
$X_1$	0.5369	0.4614	-0.0634	0.8344	0.1311	0.6922
$X_2$	-0.1294	0.8696	-0.3578	0.1005	-0.6438	0.6126
$X_3$	0.5135	-0.2538	0.7237	-0.0262	0.5598	0.1099
$X_4$	0.7239	-0.3659	0.7391	0.2947	0.7942	0.1486
$X_5$	-0.4155	-0.4142	-0.4833	-0.0126	-0.0663	-0.5815
$X_6$	0.7145	0.1236	0.2389	0.8001	0.4813	0.5321
분산 설명량	1.7751	1.3544	1.4930	1.4344	1.6113	1.5094

변수	주 축 인 자 법 (SMC)		최대우도법		최대우도법 (Varimax)	
	인자1	인자2	인자1	인자2	인자1	인자2
$X_1$	0.4029	0.3120	0.1886	0.4790	0.3540	0.3737
$X_2$	-0.1062	0.5686	1.0000	0.0000	-0.4239	0.9059
$X_3$	0.3426	-0.1385	-0.1209	0.2697	0.2956	0.0046
$X_4$	0.5586	-0.2467	-0.2664	0.5089	0.5738	-0.0258
$X_5$	-0.2864	-0.2464	-0.2445	-0.2914	-0.1604	-0.3449
$X_6$	0.5557	0.0889	-0.0673	0.6513	0.6185	0.2148
분산 설명량	0.9939	0.5694	1.4930	1.4344	1.1297	1.1261

## 8.8 R을 이용한 인자분석

### [프로그램 8.1] 비행조종사 자료에 대한 인자분석

```
pilot=read.csv("C:/data/pilot_f[REDACTED]header=T)
pilot ; attach(pilot)
x=pilot[,2:7]
m=mean(x) ; m
S=cov(x) ; S
R=cor(x) ; R
fact1=factanal(x, factors=2, rotation="none")          # no rotation
fact2=factanal(x, factors=2, scores = "regression")    # varimax is the default
fact3=factanal(x, factors=2, rotation="promax")       # promax rotation
fact1 ; fact2 ; fact3
```



```

# scree plot
library(graphics)
prin=princomp(x)
screeplot(prin,npcs=6, type="lines",main="scree plot") # 그림 8.7

# plot of factor pattern
namevar=names(fact2$loadings)=c(
plot(fact2$loadings[,1],fact2$loadings[,2],
      xlab="factor1",ylab="factor2", main="factor pattern") # 그림 8.8
      text( x=fact2$loadings[,1], y=fact2$loadings[,2], labels=namevar, adj=0)
      abline(v=0,h=0)

# plot of factor scores
plot(fact2$scores[,1], fact2$scores[,2], pch="*",
      xlab="factor1",ylab="factor2", main="factor scores") # 그림 8.9

```

● R에서는 factanal() 함수를 이용하여 상관행렬을 이용한 인자분석을 할 수 있다. default는 상관행렬로 최대우도법을 이용하고 varimax 회전을 하여 주성분분석을 한다. 주성분법을 이용한 인자분석은 princomp()를 활용한다.

인자분석시 여러 방법으로 인자분석 수행 후 출력결과 중 특히 인자패턴을 비교하며 인자의 의미를 파악하도록 해야 한다.

[결과 8.1]에서 각 방법에 따른 인자적재값을 비교할 수 있다.

$$H_0 : \text{인자 개수 2 개다} \text{ vs } H_0 : \text{not } H_0$$

varimax 회전 인자분석의 카이제곱검정 결과

$$p\text{-값} = 0.333 > 0.05 = \alpha \text{ 이므로}$$

인자 2 개인 모형이 적합하다.

## [결과 8.1] 인자분석 결과

```
> fact1=factanal(x, factors=2, rotation="none", scores = "Bartlett") #no rotation
> fact1
factanal(x = x, factors = 2, scores = "Bartlett", rotation = "none")
Uniquenesses:
      x1      x2      x3      x4      x5      x6
0.734 0.005 0.913 0.671 0.856 0.570
Loadings:
      Factor1 Factor2
x1  0.189    0.480
x2  0.997
x3 -0.121    0.269
x4 -0.267    0.508
x5 -0.245   -0.291
x6          0.652
```

	Factor1	Factor2
SS loadings	1.181	1.071
Proportion Var	0.197	0.178
Cumulative Var	0.197	0.375

Test of the hypothesis that 2 factors are sufficient.  
The chi square statistic is 4.59 on 4 degrees of freedom.  
The p-value is 0.333

```
> fact2=factanal(x, factors=2, scores = "regression") # varimax is the default
> fact2
factanal(x = x, factors = 2, scores = "regression")
Uniquenesses:
      x1      x2      x3      x4      x5      x6
0.734 0.005 0.913 0.671 0.856 0.570
```

Loadings:

	Factor1	Factor2
x1	0.354	0.375
x2	-0.425	0.903
x3	0.295	
x4	0.573	
x5	-0.159	-0.345
x6	0.619	0.217

	Factor1	Factor2
SS loadings	1.130	1.122
Proportion Var	0.188	0.187
Cumulative Var	0.188	0.375

Test of the hypothesis that 2 factors are sufficient.

The chi square statistic is 4.59 on 4 degrees of freedom.

The p-value is 0.333

```
> fact3=factanal(x, factors=2, rotation="promax")  
# promax rotation
```

```
> fact3
```

Call:

```
factanal(x = x, factors = 2, rotation = "promax")
```

Uniquenesses:

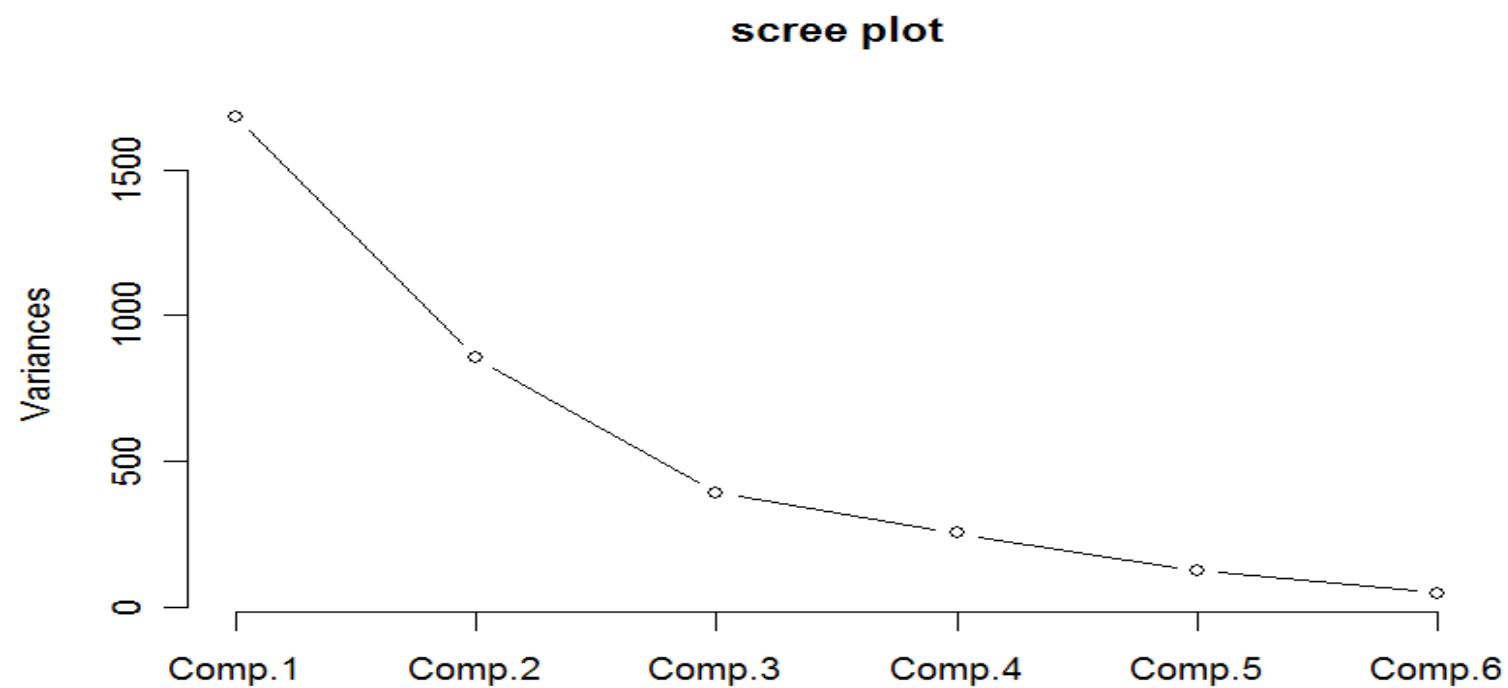
	x1	x2	x3	x4	x5	x6
	0.734	0.005	0.913	0.671	0.856	0.570

Loadings:

	Factor1	Factor2
x1	0.231	0.404
x2	1.032	-0.345
x3	-0.105	0.305
x4	-0.239	0.589
x5	-0.275	-0.200
x6		0.661

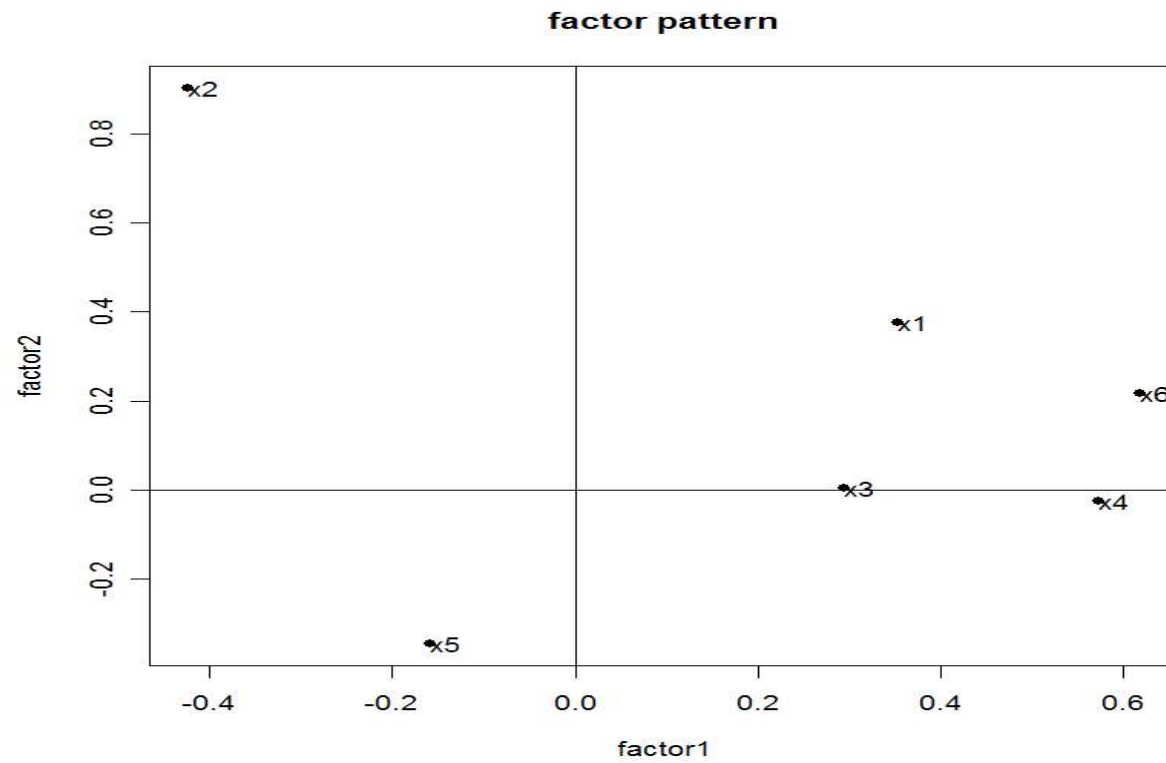
The p-value is 0.333

	Factor1	Factor2
SS loadings	1.262	1.199
Proportion Var	0.210	0.200
Cumulative Var	0.210	0.410
Test of the hypothesis that 2 factors are sufficient.		
The chi square statistic is 4.59 on 4 degrees of freedom.		

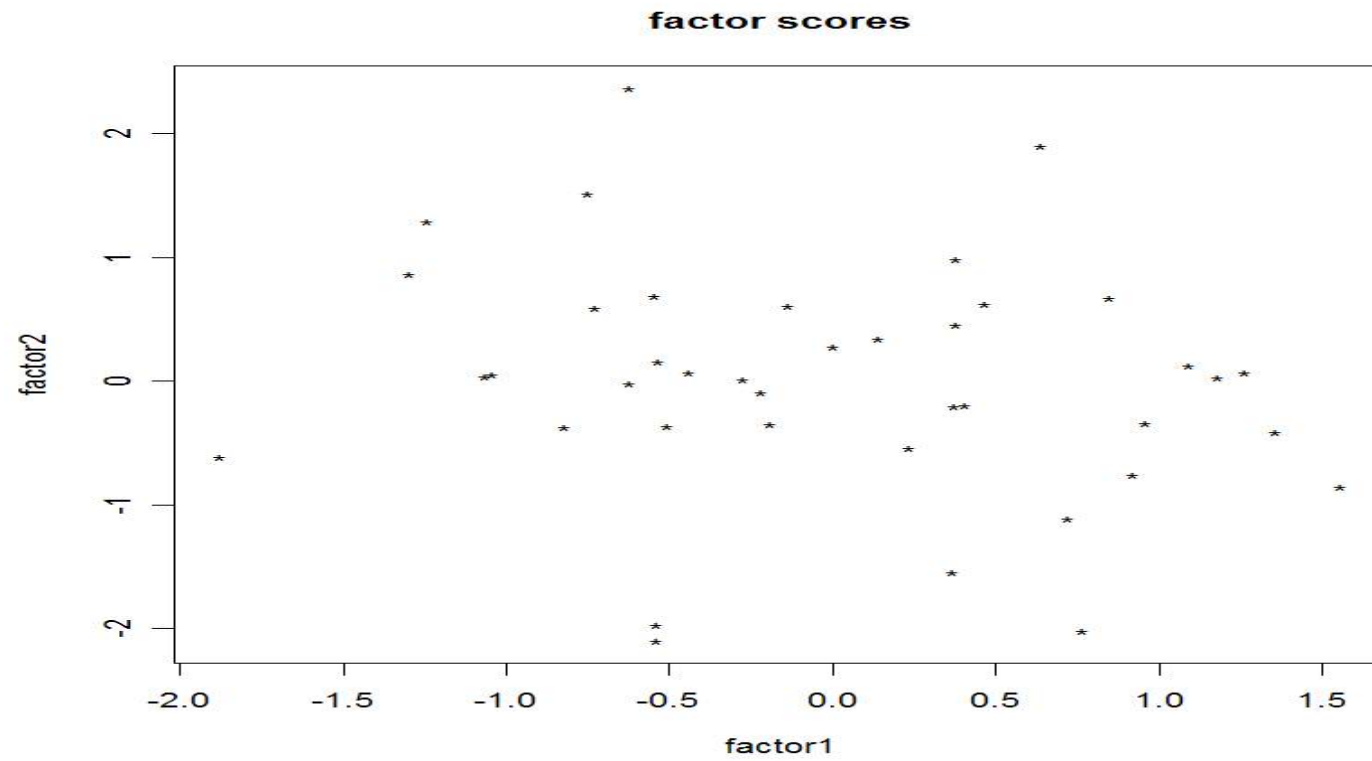


[그림 8.7] 인자 개수에 대한 스크리 그래프





[그림 8.8] Varimax 회전된 인자 패턴



[그림 8.9] Varimax 회전된 인자 점수 그래프