

11. 인자분석(Factor Analysis)

11.1 개념 및 목적

주성분분석 : p 개의 주어진 변수들의 일차결합으로 주성분을 구성하고, 변동의 설명에 대한 기여도가 큰 순서대로 $m(<p)$ 개의 주성분을 선택하여, 이들로 전체의 변동을 설명하는 분석법

인자분석(factor analysis) : 주어진 변수들을 가상의 공통인자(변수)들의 일차결합으로 나타내고, 이들 인자들 가운데 중요한 m 개의 인자만 선택하여 전체의 변동을 설명하는 분석법

인자분석이 실제로 이용되는 범위는 매우 넓다.

예 :

- p 개의 변수들을 비슷한 특성을 가진(즉, 상관관계가 높은) 변수들끼리 묶어서 전체의 특성을 파악
- 인자분석의 결과를 회귀분석의 입력변수로 사용하기도 함

예제 11.1 220명의 학생에 대한 6개 과목 성적의 상관행렬

계일어	1.000					
영 어	.439	1.000				
역 사	.410	.351	1.000			
산 수	.288	.354	.164	1.000		
대 수	.329	.320	.190	.595	1.000	
기 하	.248	.329	.181	.470	.464	1.000

<그림 11.1> 시험성적의 표본상관행렬

<표 11.1> $m=2$ 인 인자분석의 결과

과 목	인자1	인자2
계일어	0.66	0.45
영 어	0.69	0.29
역 사	0.52	0.64
산 수	0.74	-0.41
대 수	0.74	-0.38
기 하	0.67	-0.36

인자1 : 일반지능(general intelligence)과 같이 지적능력으로 이해

인자2 : 어학과목의 수리과목에 대한 우세성으로 이해

<표 11.2> Varimax 회전의 결과

과목	인자1	인자2
게일어	0.23	0.76
영 어	0.35	0.66
역 사	-0.00	0.82
산 수	0.83	0.15
대 수	0.81	0.18
기 하	0.75	0.15

인자1 : ‘수리능력’ 으로 이해

인자2 : ‘어학능력’ 으로 이해

11.2 이론적 배경

(1) 모형

다변량정규확률변수 $\mathbf{x}' = (x_1, \dots, x_p)$

평균 : $\boldsymbol{\mu}' = (\mu_1, \dots, \mu_p)$

공분산행렬 : Σ

인자분석 모형 : (인자패턴 , factor pattern)

$$X_1 - \mu_1 = \ell_{11}F_1 + \ell_{12}F_2 + \cdots + \ell_{1m}F_m + \varepsilon_1$$

$$X_2 - \mu_2 = \ell_{21}F_1 + \ell_{22}F_2 + \cdots + \ell_{2m}F_m + \varepsilon_2$$

$$\vdots$$
$$\vdots$$
$$\vdots$$

(11.1)

$$X_p - \mu_p = \ell_{p1}F_1 + \ell_{p2}F_2 + \cdots + \ell_{pm}F_m + \varepsilon_p$$

용어 :

F_j : 공통인자(common factor)

- 전체 변수에 공통으로 사용하는 변수로서 관측이 불가능

ε_i : 유일인자(unique factor) 또는 특정인자(specific factor)

- 유일하게 i 번째 변수에만 기여하는 변수로서 오차항의 개념

ℓ_{ij} : 인자적재(factor loading)

$\psi_i = Var(\varepsilon_i)$: 유일분산 또는 특정분산

$h_i^2 = \ell_{i1}^2 + \dots + \ell_{im}^2$: 공통인자분산 또는 공통성(communality).

F 와 ε 에 대한 가정 :

1. F_j 들은 서로 독립이고, 평균이 0, 분산 1이다.
2. ε_i 들은 서로 독립이고, 평균이 0, 분산 ψ_i 이다.
3. F 와 ε 은 서로 독립이다.

참고

① F 와 ε 은 서로 독립이므로 $Var(x_i) = h_i^2 + \psi_i$

② $h_i^2 = \ell_{i1}^2 + \dots + \ell_{im}^2$ 은 x_i 의 분산 가운데

m 개의 인자 F_1, \dots, F_m 이 설명하는 부분

(2) 인자추출법

주성분분석 (principal component analysis)

주인자분석 (principal factor analysis)

최우인자분석 (maximum-likelihood factor analysis)

① 주성분분석

SAS의 FACTOR 절차에서 디폴트(default) 방법
공분산행렬의 주성분을 이용

공분산행렬 Σ 의 고유근-고유벡터 : $(\lambda_i, e_i), i=1, \dots, p$

기호 :

$$L = \begin{bmatrix} l_{11} & \cdots & l_{1m} \\ \vdots & \ddots & \vdots \\ l_{p1} & \cdots & l_{pm} \end{bmatrix}, \quad f = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix}, \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_p \end{bmatrix}, \quad \Psi = \begin{bmatrix} \psi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \psi_p \end{bmatrix}$$

(11.1)식의 인자모형 행렬 :

$$x - \mu = Lf + \varepsilon \quad (11.2)$$

F_i 들은 평균 0, 분산 1이고 서로 독립이므로

$$\Sigma = LL' + \Psi \quad (11.3)$$

또한 Σ 를 m 처음개의 주성분으로 나타내면

$$\Sigma = (\sqrt{\lambda_1}e_1 \ \cdots \ \sqrt{\lambda_m}e_m) \begin{pmatrix} \sqrt{\lambda_1}e_1 \\ \vdots \\ \sqrt{\lambda_m}e_m \end{pmatrix} + \begin{pmatrix} \psi_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \psi_p \end{pmatrix} \quad (11.4)$$

$$= LL' + \Psi$$

(11.3)과 (11.4)에서 인자행렬 L 은 Σ 의 주성분에서 얻어짐.

실제로는 Σ 대신 S 또는 R 을 사용

인자적재 ℓ_{ij} 는 공통인자의 수 m 이 변해도 고유벡터의 직교성
때문에 변하지 않음.

② 주인자분석 (principal factor analysis)

③ 최우인자분석(maximum likelihood factor analysis)

F 와 ε 의 결합분포가 정규분포라는 가정하에

$$X - \mu = LF + \varepsilon \sim \text{다변량 정규분포}$$

\therefore 최대가능도추정법(MLE)으로 L 과 Ψ 를 추정

Heywood case : 공통성의 추정값이 1 이상인 경우

(\therefore 특정분산의 추정값이 음수인 경우)

HEYWOOD 옵션을 사용하여 공통성의 추정값을 1로 고정

(3) 공통인자의 개수 (m)의 선택기준

① 고유값의 크기

㉠ 공분산행렬을 이용할 때 :

$$\frac{\lambda_i}{\lambda_1 + \cdots + \lambda_p} > \bar{\lambda} \text{ (또는 } 0.7\bar{\lambda} \text{)} \text{인 } \lambda_i \text{의 개수}$$

㉡ 상관행렬을 이용할 때 :

$$\lambda_i > 1 \text{ (또는 } 0.7 \text{)} \text{인 } \lambda_i \text{의 개수}$$

② Scree plot (10장 참조)

③ 카이제곱 검정 : 최대가능도 인자분석법에서 가설검정

$H_0 : m$ 이 충분하다, $H_1 : \text{더 많은 공통인자가 필요하다}$

④ AIC기준 : Akaike 통계량에 의한 판정기준

통계량의 값이 가장 작은 공통인자의 개수를 선택

(4) 인자회전(factor rotation)

인자추출법에 의해 구해진 인자들이 뚜렷한 의미를 갖지 못하는 경우에, 인자행렬 L 의 원소들이 0 또는 ± 1 에 가깝게 되도록 회전

대표적인 회전방법 :

직교회전(orthogonal rotation)

사각회전(oblique rotation)

① 직교회전

인자축들이 직교성을 유지하도록 회전해주는 방법

예. $T = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ 에 의한 회전

직교회전 방법 :

VARIMAX

QUARTIMAX

② 사각회전

직교회전으로 해석이 어려울 경우 변수 군(group)을 지나는 축을 선택

- 해석이 용이해지는 장점이 있으나 인자들의 상관성이 문제

사각회전 방법 : PROMAX 방법

11.3 FACTOR 절차

일반형 :

```
PROC FACTOR <options>;  
PRIORS communalities;  
VAR variables;  
PARTIAL variables;  
FREQ variables;  
WEIGHT variables;  
BY variables;
```

① PROC FACTOR <options>;

㉠ Data Set options

DATA=SAS dsn(data set name) : 분석에 사용될 자료명

OUT=SAS dsn : 원래의 자료와 인자점수(factor score)를 저장

인자점수는 인자점수계수와 관측값의 곱에 의해 구해지므로 입력
으로 사용될 자료가 공분산행렬이나 상관행렬이어서는 안됨.

㉞ 인자추출 방법의 선택

METHOD(또는 M)=name : name에 의해 인자추출방법을 지정

name의 예:

PRINCIPAL(또는 PRIN, P) : 주성분분석 (default)

ML(또는 M) : 최대가능도 인자분석

㉔ 공통인자의 개수의 선택기준

NFACTORS(또는 NFACT, N)=m :

추출할 인자의 수를 지정. 디폴트이면 원래 변수의 수

MINEIGEN(또는 MIN)=m :

지정된 값보다 큰 고유값에 해당되는 인자만을 보유

디폴트는 “고유값 ≥ 1 ” 인 인자만을 보유

㉞ 인자회전방법의 선택

ROTATE(또는 R)=name :

name의 예:

VARIMAX(또는 V) : varimax rotation

QUARTIMAX(또는 Q) : quartimax rotation

PROMAX(또는 P) : promax rotation

㉔ Output options

SIMPLE(또는 S) : 평균과 표준편차만 출력

CORR(또는 C) : 상관행렬을 출력

SCREE : 고유값의 스크리 산점도(scree plot)를 출력

EIGENVECTORS(또는 EV) : 고유벡터를 출력

MSA :

Kaiser의 measure of sampling adequacy(MSA)를 출력

0.8 이상이면 변수의 수가 충분하며 0.5 이하이면 더 많은 변수가 필요함을 뜻함.

PREPLOT : 회전전의 인자패턴을 출력

PLOT : 회전후의 인자패턴을 출력

NPLOT=n : 그려줄 인자의 수를 지정

디폴트이면 $n(n-1)/2$ 개의 그림을 그려준다.

SCORE : 인자점수계수(factor scoring coefficient)를 출력

ALL : plot을 제외한 모든 output을 출력

REORDER(또는 RE) :

각 인자별로 인자적재의 절대값이 큰 순서대로 변수를 재배열

② VAR variables;

분석에 사용될 숫자 변수들을 지정

③ FREQ variables;

도수를 나타낼 경우 이 변수를 지정

(관측값이 도수만큼 입력된 것으로 간주)

11.4 예제

예제 11.1(계속) 성적자료를 분석하는 SAS 프로그램

```
/* FACTOR1.SAS : FACTOR ANALYSIS OF EXAM SCORES */
DATA STUDY (TYPE=CORR); ①
TITLE 'FACTOR ANALYSIS; EXAM SCORES OF N=220 MALE STUDENTS';
  _TYPE_='CORR'; INPUT _NAME_ $ X1-X6;
  LABEL
    X1='GAELIC' X2='ENGLISH' X3='HISTORY'
    X4='ARITHMETIC' X5='ALGEBRA' X6='GEOMETRY';
  CARDS;
X1      1.000      .      .      .      .      .
X2      .439      1.000      .      .      .      .
X3      .410      .351      1.000      .      .      .
X4      .288      .354      .164      1.000      .      .
X5      .329      .320      .190      .595      1.000      .
X6      .248      .329      .181      .470      .464      1.000
;
```

RUN;

PROC PRINT;

②

RUN;

PROC FACTOR MSA PREPLOT ROTATE=VARIMAX PLOT; ③ ④ ⑤ ⑥

TITLE2 'FACTOR ANALYSIS; EXAM SCORES OF STUDENTS';

RUN;

QUIT;

① _TYPE_='CORR'은 입력 자료가 상관행렬임을 뜻한다.

② 입력된 상관행렬을 출력하며, 결과는 다음과 같다.

OBS	_TYPE_	_NAME_	X1	X2	X3	X4	X5	X6
1	CORR	X1	1.000
2	CORR	X2	0.439	1.000
3	CORR	X3	0.410	0.351	1.000	.	.	.
4	CORR	X4	0.288	0.354	0.164	1.000	.	.
5	CORR	X5	0.329	0.320	0.190	0.595	1.000	.
6	CORR	X6	0.248	0.329	0.181	0.470	0.464	1

<그림 11.3> PRINT 절차에 의해 출력된 상관행렬

③ ‘METHOD=’ 옵션이 사용되지 않았으므로 주성분분석 사용

인자의 수를 지정해 주지 않았으므로 디폴트로 1보다 큰
고유값의 개수인 2개의 인자를 보유
2개의 인자에 의한 인자공헌도 : 64%

④ MSA 옵션에 의해 KAISER의 측도값이 0.775

분석에 사용된 변수의 수가 적당하다고 생각된다.

Kaiser's Measure of Sampling Adequacy: Overall MSA = 0.77487528

X1	X2	X3	X4	X5	X6
0.77155652	0.81124073	0.75553558	0.74479572	0.74957020	0.83163354

Prior Communalities Estimates: ONE

Eigenvalues of the Correlation Matrix: Total = 6 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	2.73288407	1.60311371	0.4555	0.4555
2	1.12977037	0.51459650	0.1883	0.6438
3	0.61517387	0.01395199	0.1025	0.7463
4	0.60122188	0.07642497	0.1002	0.8465
5	0.52479691	0.12864401	0.0875	0.9340
6	0.39615290		0.0660	1.0000

2 factors will be retained by the MINEIGEN criterion.

Factor Pattern

Factor1

Factor2

X1	Gaelic	0.65782	0.44905
X2	English	0.68842	0.29039
X3	History	0.51737	0.63734
X4	Arithmetic	0.73831	-0.41303
X5	Algebra	0.74388	-0.37545
X6	Geometry	0.67831	-0.35501

Variance Explained by Each Factor

Factor1	Factor2
2.7328841	1.1297704

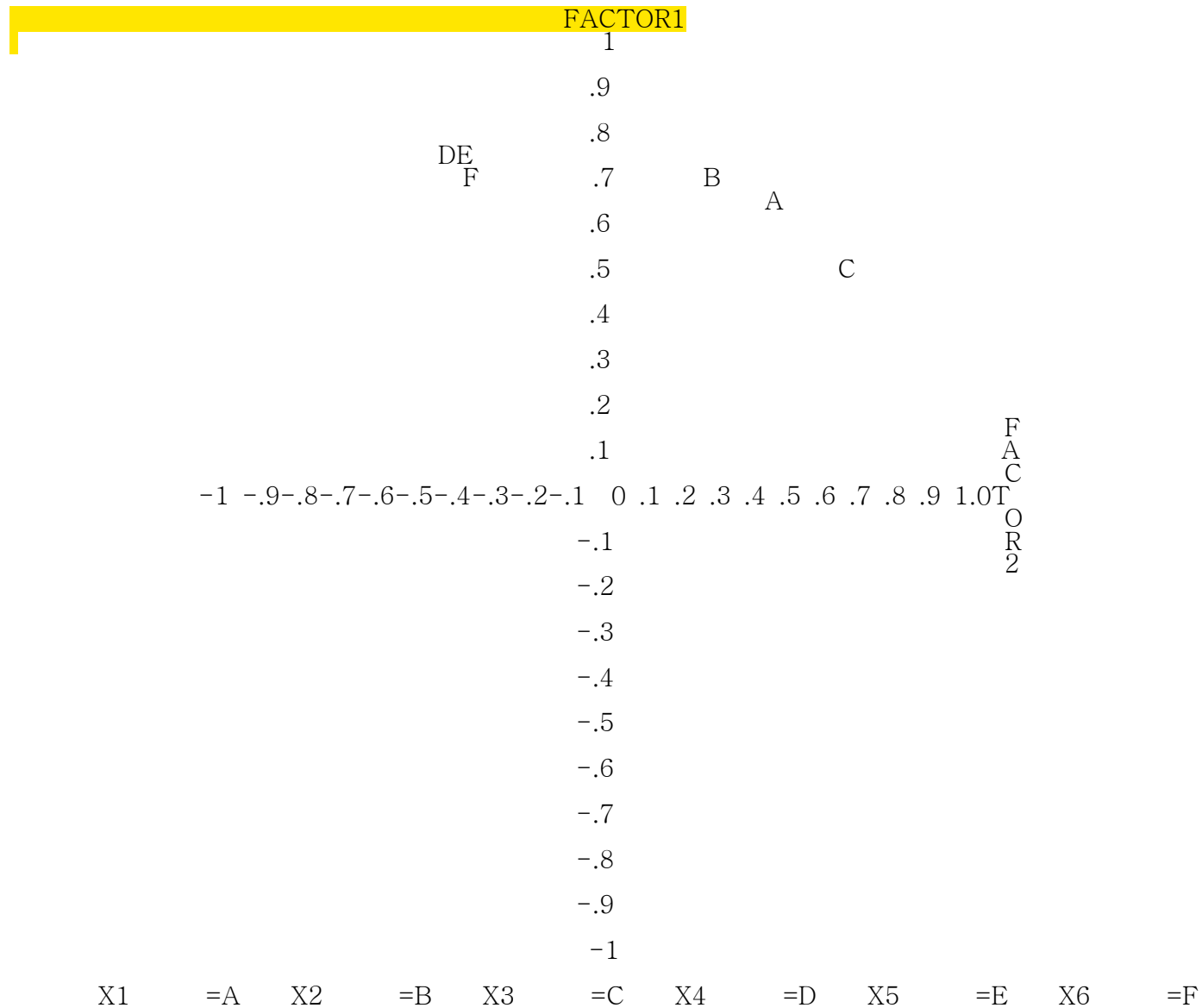
Final Communalities Estimates: Total = 3.862654

X1	X2	X3	X4	X5	X6
0.63437428	0.55824594	0.67387364	0.71570325	0.69431309	0.58614424

<그림 11.3> FACTOR 절차에서 출력된 결과

⑤ PREPLOT에 의해 인자회전을 하기 전의 인자패턴

Plot of Factor Pattern for FACTOR1 and FACTOR2



⑥ 직교회전인 VARIMAX 방법을 사용

<그림 11.6>에서

$$T = \begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} = \begin{bmatrix} 0.7744 & 0.6327 \\ -0.6327 & 0.7744 \end{bmatrix}$$

$$\Rightarrow \phi = 43.6^\circ$$

즉, 시계바늘회전방향으로 약 43.6 회전

Rotation Method: Varimax

Orthogonal Transformation Matrix

	1	2
1	0.77441	0.63269
2	-0.63269	0.77441

Rotated Factor Pattern

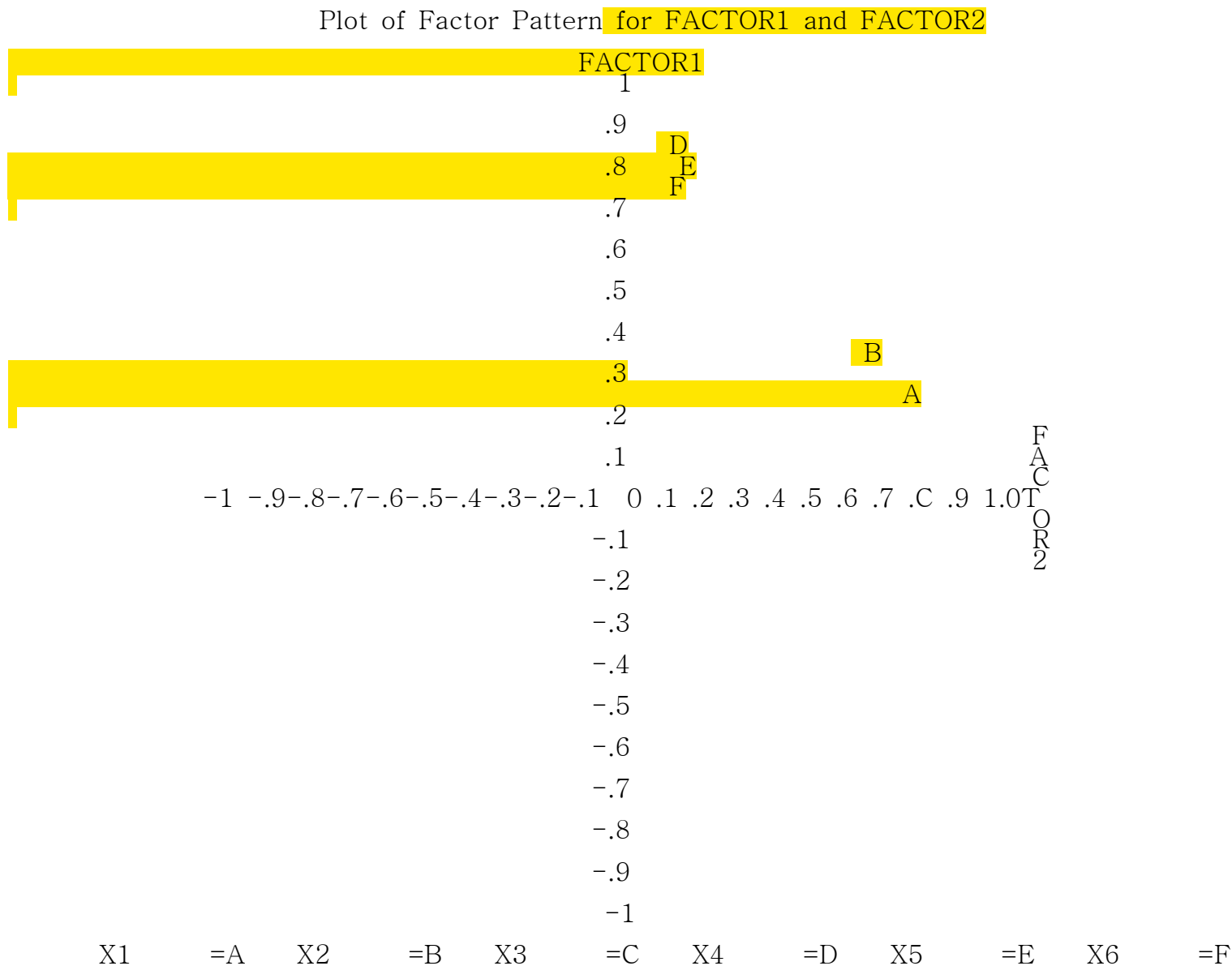
	FACTOR1	FACTOR2	
X1	0.22531	0.76394	Gaelic
X2	0.34939	0.66044	English
X3	-0.00259	0.82089	History
X4	0.83308	0.14727	Arithmetic
X5	0.81360	0.17989	Algebra
X6	0.74990	0.15424	Geometry

Variance explained by each factor

FACTOR1	FACTOR2
2.091164	1.771490

Final Communalities Estimates: Total = 3.862654

X1	X2	X3	X4	X5	X6
0.634374	0.558246	0.673874	0.715703	0.694313	0.586144



<그림 11.7> 인자회전 후의 인자패턴 그림

예제 11.2 한 방진복(防塵服) 생산 회사에서는 **치수체계 조사**

방진복 착용자 240명을 랜덤하게 뽑아 체위조사를 시행

10개의 신체부위에 대한 계측

자료에서 변수

V1: 신장 V2 : 총길이 V3 : 등길이 V4 : 화장

V5 : 소매길이 V6 : 바지길이 V7 : 밑위앞뒤 V8 : 가슴둘레

V9 : 허리둘레 V10 : 엉덩이둘레

<표 11.3> 인체 계측자료 (파일 이름: PHYSICAL.DAT)

– 41 –

```

/* FACTOR2.SAS : FACTOR ANALYSIS FOR 10 PHYSICAL VARIABLES */
DATA PHYSICAL;
    INFILE 'Z:\WPHYSICAL.txt' FIRSTOBS=2 MISSOEVER;
    INPUT ID V1-V10;
    LABEL V1='신장' V2='총길이' V3='등길이' V4='화장' V5='소매길이'
          V6='바지길이' V7='밑위앞뒤' V8='가슴둘레' V9='허리둘레'
          V10='엉덩이둘레';
RUN;
PROC FACTOR DATA=PHYSICAL SCREE PREPLOT          ① ② ③
          ROTATE=VARIMAX REORDER PLOT;          ④ ⑤ ⑥
    VAR V1-V10;
RUN; QUIT;

```

① 1보다 큰 고유값 : 2개 \Rightarrow 디폴트로 2개의 인자를 선택

Initial Factor Method: Principal Components

Prior Communality Estimates: ONE

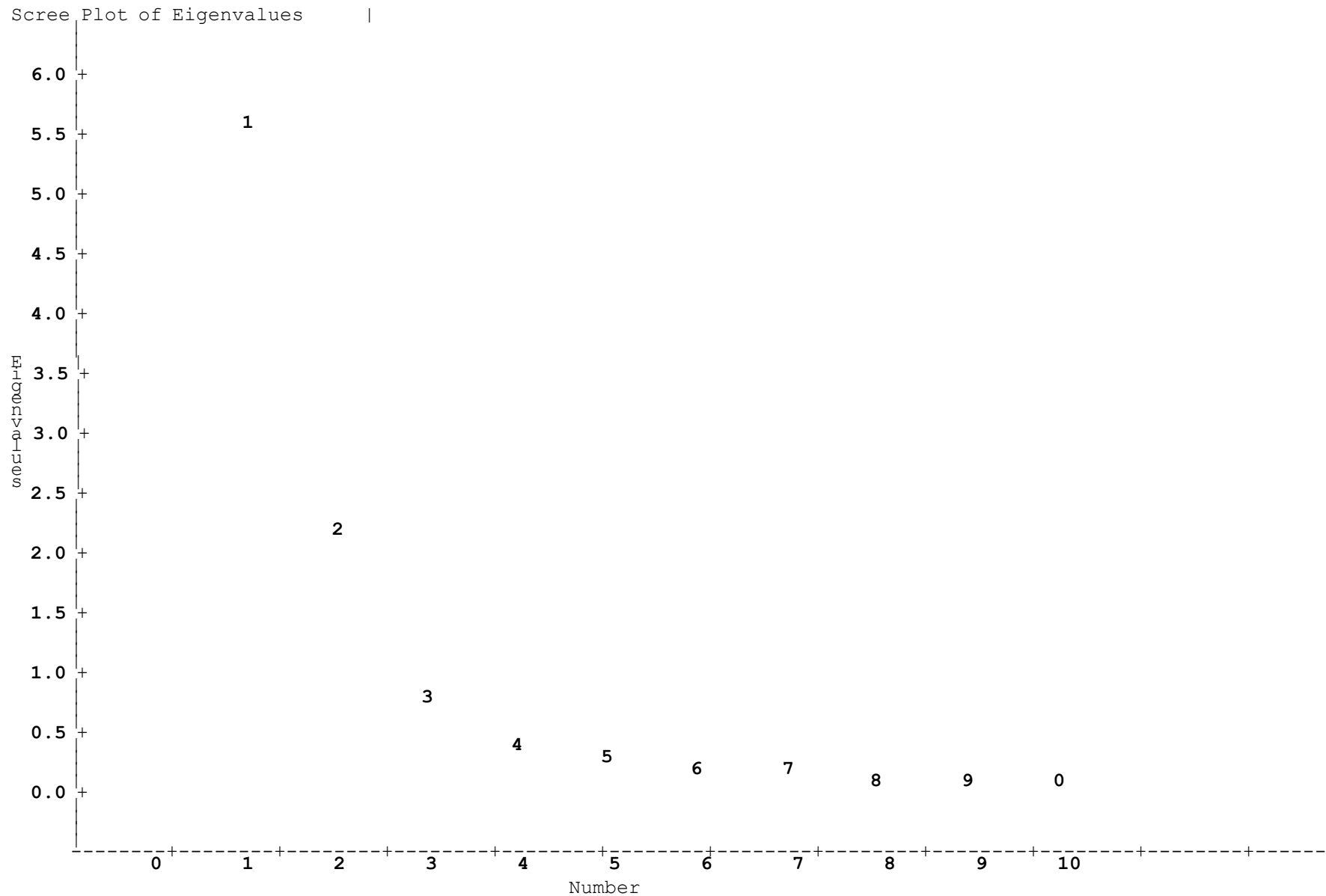
Eigenvalues of the Correlation Matrix: Total = 10 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	5.57573244	3.38921142	0.5576	0.5576
2	2.18652101	1.35294267	0.2187	0.7762
3	0.83357834	0.41375318	0.0834	0.8596
4	0.41982516	0.10546299	0.0420	0.9016
5	0.31436217	0.07252753	0.0314	0.9330
6	0.24183464	0.09019195	0.0242	0.9572
7	0.15164269	0.02917449	0.0152	0.9723
8	0.12246820	0.02401913	0.0122	0.9846
9	0.09844907	0.04286280	0.0098	0.9944
10	0.05558628		0.0056	1.0000

2 factors will be retained by the MINEIGEN criterion.

② SCREE 옵션에 의해 스크리 산점도가 출력

3번째 값부터 감소가 완만해지므로 2개의 인자가 적당



③ PREPLOT 옵션에 의해 회전 전의 인자패턴 및 그림을 출력

(일부 내용 생략)

Initial Factor Method: Principal Components

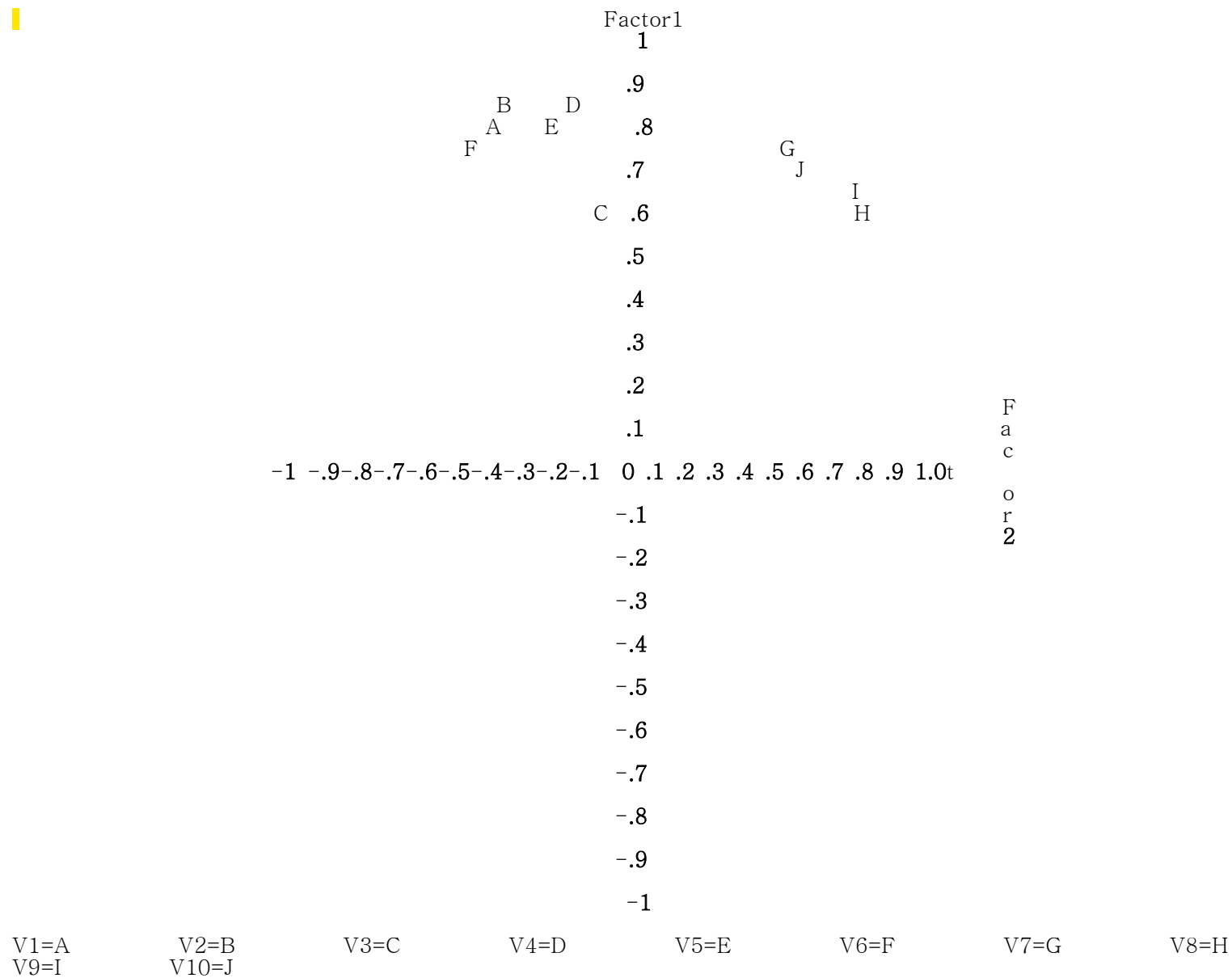
Factor Pattern

		Factor1	Factor2
V4	화장	0.85862	-0.21701
V2	총길이	0.85790	-0.41922
V1	신장	0.82453	-0.44965
V5	소매길이	0.79164	-0.27739
V6	바지길이	0.75009	-0.50627
V7	밑위앞뒤	0.73414	0.46639
V10	엉덩이둘레	0.71687	0.50727
V3	등길이	0.58625	-0.12699
V8	가슴둘레	0.62423	0.69569
V9	허리둘레	0.66867	0.67327

Factor1의 의미 : 전체 크기(size)

Factor2의 의미 : 둘레 vs. 길이

Plot of Factor Pattern for Factor1 and Factor2



④ VARIMAX 인자회전 결과 (REORDER 옵션에 의해서 크기에 따라 재배열)

Rotation Method: Varimax
Orthogonal Transformation Matrix

	1	2
1	0.80354	0.59525
2	-0.59525	0.80354

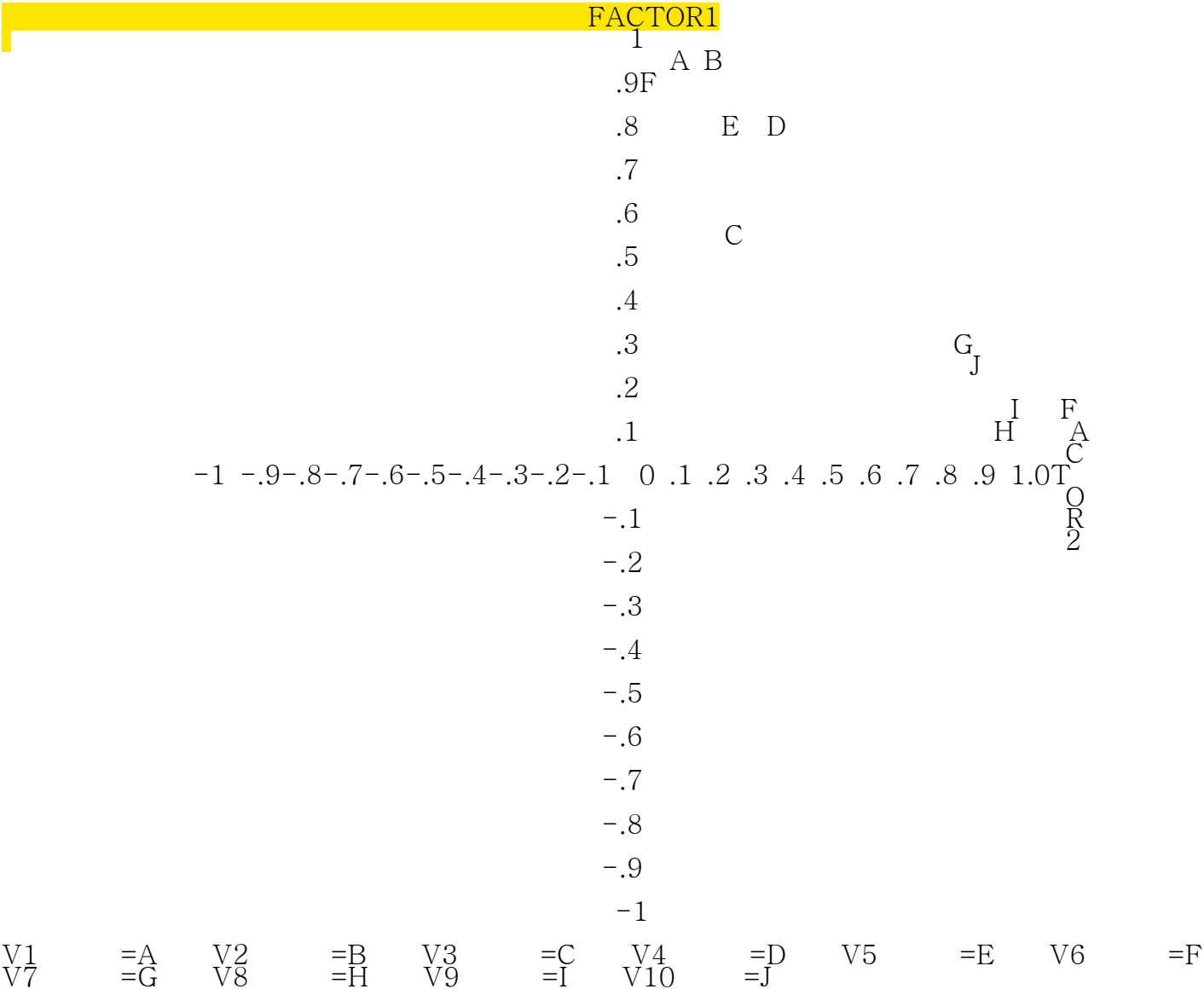
Rotated Factor Pattern

		Factor1	Factor2
V2	총길이	0.93890	0.17381
V1	신장	0.93020	0.12948
V6	바지길이	0.90408	0.03968
V4	화장	0.81911	0.33671
V5	소매길이	0.80123	0.24833
V3	등길이	0.54667	0.24692
V9	허리둘레	0.13654	0.93903
V8	가슴둘레	0.08749	0.93059
V10	엉덩이둘레	0.27409	0.83433
V7	밑위앞뒤	0.31230	0.81176

⑥ PLOT 옵션에 의해 VARIMAX 회전 후의 인자 패턴을
그림으로 그려준다.

Rotation Method: Varimax

Plot of Factor Pattern for FACTOR1 and FACTOR2



<그림 11.13> VARIMAX 회전 후의 인자패턴

인자분석의 일차목표 :

- ① 여러 변수에 공통으로 작용하는 요인들을 찾고,
- ② 원래의 변수들을 비슷한 특성을 가진 변수들끼리 그룹화

예제 11.3

우리나라 임산부들의 브래지어에 대한 불만사항을 파악
서울 소재의 두 병원에서 산부인과 외래환자를 상대로 설문조사
설문내용은 <그림 11.14> 조사결과는 <표 11.4>

인자분석에서는 $n=5$ 개의 인자
VARIMAX 회전을 실시

	매우 그렇다	그렇다	그저 그렇다	그렇지 않다	전혀 그렇지 않다
1. 치수가 잘 안맞다.	()	()	()	()	()
2. 컵이 작다.	()	()	()	()	()
3. 컵아래(밑가슴)부분이 조인다.	()	()	()	()	()
4. 겨드랑이 밑이 조인다.	()	()	()	()	()
5. 원단의 신축성이 없다.	()	()	()	()	()
6. 촉감이 거칠다.	()	()	()	()	()
7. 땀을 잘 흡수하지 못한다.	()	()	()	()	()
8. 오래 착용시 답답하다.	()	()	()	()	()
9. 착용시 앞중심 부분이 들뜬다.	()	()	()	()	()
10. 컵아래 부분이 위로 기어 올라간다.	()	()	()	()	()
11. 어깨끈이 흘러 내린다.	()	()	()	()	()
12. 뒤여밈에 의한 착용시 불편하다.	()	()	()	()	()
13. 앞여밈에 의한 착용시 불편하다.	()	()	()	()	()
14. 가슴을 잘 받쳐주지 못한다.	()	()	()	()	()
15. 가슴을 잘 모아주지 못한다.	()	()	()	()	()
16. 가슴의 흔들림을 잘 고정시키지 못한다.	()	()	()	()	()
17. 디자인이 세련되지 못한다.	()	()	()	()	()
18. 디자인이 다양하지 못하다.	()	()	()	()	()

<그림 11.14> 설문 내용

<표 11.4> 설문조사의 결과(파일이름: QUSTNAIR.DAT)

222444332 344443421	424422425 .24.34433	443444442 444322322	333222214 222332333
252445424 244444451	252222224 224322244	342344322 232.22232	222444422 222211122
322323224 433333333	233312222 223423311	343543255 545522355	554424112 223333333
452434343 551.42443	342232412 143311211	223334212 223233322	232444413 212.22233
224434424 244444434	344223222 242444222	243444234 234332322	222532142 322222244
221344433 242213344	223244223 213323233	442434443 333432233	442244424 424.22444
222334444 323422343	113432425 555522331	332222222 231222233	332223424 422333442
211324224 444323322	344553245 555553444	234444342 224442333	254453455 .11522342
123223212 242321132	222344422 222332333	324555244 233.22444	344444332 323323342
222424222 241422222	352525452 544422224	232424312 233322233	212344344 243342244
222545432 223422332	424434434 424333433	442334433 223433444	223424244 223222243
222434222 243322243	344424212 113322444	232444331 413422223	315255355 155511144
344444324 224433344	232322224 222332243	321444242 224312132	452522222 245.22344
111423422 233233333	341334313 322411222	111233222 453222233	254424322 242422.4.
222434423 234443333	222444433 423.44443	324434423 433433343	124533222 434433322
221333334 222143222	232333213 244.22222	222423424 434332222	322424224 224422222
44344442. 433.33344	212243342 444422222	334544443 334.22344	322333223 223332234
221344234 433122332	324233221 212332344	221123332 454344333	344444324 434422344
231133213 432442322	232224332 323232344	44.54.344 222322344	442244225 252.43242
455524354 245353355	5125422.2 152311144	2.222.222 424422244	443433442 344433344
224434222 444.42442	443445344 333443344	255555255 225222242	244444442 444443344
444433242 231.43321	322344243 433332333	242422222 423322322	332433343 223222233
224522322 242422342	322524424 444442442	222444222 222.2234.	23344.422 223332322
33222433. 2243.3.32	222422222 422322233	555255552 222444242	221334212 424322322
231333322 233332133	233334434 423.33343	322523344 244423433	334444323 342333333
352433424 322322322	242434424 244422333	433444332 444.34344	342333232 24..44433
242423323 233311244	242224422 242244422	424524445 555544442	425544442 242444233
332444423 333443322	112434322 322333343	222423222 244333332	214444422 .44.11133
121333212 233211344	234544322 443422412	344425444 433344444	42425543. 252433322
442434424 442432222	233322232 122322233	222444332 323223323	342434324 422444433
242424444 422441222	344433344 433332244	444442223 242422233	322234334 443333444
323535252 544.33333	224444415 555533333	224422222 222422233	
252442224 444422244	3322323.3 3.2233222	441232222 222332322	

```
/* FACTOR3.SAS : 임신부용 브라지어의 불만조사 */  
DATA BRA;  
    INFILE 'A:\WQUSTNAIR.DAT';  
    INPUT (Q1-Q9) (1.) + 1 (Q10-Q18) (1.);  
RUN;  
PROC FACTOR DATA=BRA ROTATE=VARIMAX REORDER N=5;  
RUN;
```

Eigenvalues of the Correlation Matrix: Total = 18 Average = 1

	Eigenvalue	Difference	Proportion	Cumulative
1	4.01642048	2.03165236	0.2231	0.2231
2	1.98476812	0.18580231	0.1103	0.3334
3	1.79896581	0.31138628	0.0999	0.4333
4	1.48757953	0.18921314	0.0826	0.5160
5	1.29836639	0.31640090	0.0721	0.5881
6	0.98196549	0.03992564	0.0546	0.6427
7	0.94203984	0.16740526	0.0523	0.6950
8	0.77463459	0.00146288	0.0430	0.7380
9	0.77317171	0.10321831	0.0430	0.7810
10	0.66995340	0.08060784	0.0372	0.8182
11	0.58934557	0.07146380	0.0327	0.8510
12	0.51788176	0.03577967	0.0288	0.8797
13	0.48210209	0.05606309	0.0268	0.9065
14	0.42603901	0.03292648	0.0237	0.9302
15	0.39311253	0.07243435	0.0218	0.9520
16	0.32067818	0.02966077	0.0178	0.9698
17	0.29101741	0.03905933	0.0162	0.9860
18	0.25195808		0.0140	1.0000

Rotated Factor Pattern

	FACTOR1	FACTOR2	FACTOR3	FACTOR4	FACTOR5
Q12	0.71906	0.00376	0.23505	-0.01491	0.22604
Q11	0.65522	0.14066	0.04701	-0.13447	-0.02415
Q13	0.64239	0.06521	0.06844	0.42904	-0.15569
Q10	0.63899	0.19655	-0.00907	-0.09028	-0.09349
Q9	0.56617	0.10630	0.04140	0.08853	0.31935
Q4	0.44128	-0.07103	0.22676	0.29025	0.04852
Q15	0.03150	0.85511	0.11644	0.09195	0.04534
Q14	0.09183	0.78340	0.07542	0.17344	0.08229
Q16	0.24664	0.74428	-0.04870	0.02098	0.02279
Q6	0.01398	0.24577	0.80328	0.03763	0.02944
Q5	0.12355	-0.16434	0.72030	0.07697	0.06663
Q8	0.18543	-0.02931	0.56759	0.20292	0.39868
Q7	0.16514	0.35439	0.53507	0.08462	-0.22727
Q2	-0.06004	0.02758	-0.03149	0.76839	0.04330
Q1	-0.06269	0.20231	0.14747	0.75977	0.07074
Q3	0.18342	0.09864	0.17494	0.60623	0.24167
Q17	0.08653	0.12524	0.08600	0.09292	0.85368
Q18	-0.00330	-0.01193	-0.00255	0.11543	0.84583

<그림 11.16> 고유값과 회전된 인자패턴

FACTOR1 (착용감)	Q12	뒤여밈에 의한 착용시 불편하다.
	Q11	어깨끈이 흘러 내린다.
	Q13	앞여밈에 의한 착용시 불편하다.
	Q10	컵아래 부분이 위로 기어 올라간다.
	Q9	착용시 앞중심 부분이 들뜬다.
	Q4	겨드랑이 밑이 조인다.
FACTOR2 (받침성)	Q15	가슴을 잘 모아주지 못한다.
	Q14	가슴을 잘 받쳐주지 못한다.
	Q16	가슴의 흔들림을 잘 고정시키지 못한다.
FACTOR3 (촉감, 신축성)	Q6	촉감이 거칠다.
	Q5	원단의 신축성이 없다.
	Q8	오래 착용시 답답하다.
	Q7	땀을 잘 흡수하지 못한다.
FACTOR4 (치수)	Q2	컵이 작다.
	Q1	치수가 잘 안맞다.
	Q3	컵아래(밑가슴)부분이 조인다.
FACTOR5 (디자인)	Q17	디자인이 세련되지 못한다.
	Q18	디자인이 다양하지 못하다.

<그림 11.17> 인자별 변수의 내용

각 인자별 불만족도 : 각 인자의 평균

인자	평균
인자1 (착용감)	3.13
인자2 (받침성)	2.62
인자3 (촉감·신축성)	3.06
인자4 (치수)	2.74
인자5 (디자인)	3.00

(낮을수록 큰 불만)

받침성과 치수에 대한 불만이 다른 요인보다 높음

이유 : 전용 브래지어의 치수가 맞지 않아서 일반 브래지어 가운데
큰 것으로 대체하고 있기 때문인 것으로 추정

추가 내용 : 인자점수(factor scores)와 그의 응용

인자점수의 이용 :

p 개의 변수 X_1, \dots, X_n 대신에 m ($m < p$) 개의 인자를 사용하여 관측값을 나타내고자 할 때 인자점수를 사용.

즉, 인자점수는 인자에 대한 각 관측값의 수준을 나타내며, 따라서 이들 인자점수를 새로운 관측값으로 간주하고 이들을 설명변수(독립변수)로 사용하여 회귀분석 등을 시행할 수 있다.

예: 만족도 조사

$$(\text{전체 만족도})_i = \beta_0 + \beta_1(\text{인자점수1})_i + \cdots + \beta_m(\text{인자점수}m)_i$$

출력 결과에서 표준화된 베타계수는 각 인자의 중요도에 해당

사례 연구 : 신축 학교 평가에 관한 연구

데이터 : 신축된 2개 초등학교에서 설문조사
(A초등학교 - 79명, B초등학교 - 60명)
(대조그룹 부재)

조사내용 : 신축 학교의 만족도 조사

조사항목 : ① 총체적 만족도 (2문항)

② 개별 만족도 (13분야에서 36개의 개별만족도)

연구내용 : 분야별 만족도와 개선 우선도

설문조사의 예

5점척도 : (매우 만족, 약간 만족, 보통, 약간 불만족, 불만족)

(1) 외부공간에 대한 만족도 조사

- ① 자연학습장에 대한 만족 상태
- ② 놀이 장소와 시설에 대한 만족 상태
- ③ 휴식공간에 대한 만족 상태

(2) 규모에 대한 조사 내용

- ① 일반교실 규모 ② 특별교실 규모 ③ 화장실 규모

(3) 가변성에 대한 만족도

- ① 교실 벽체 활용성 ② 다목적실

(4) 건물 외관

- ① 건물 외부 형태 ② 건물 외부 색채 ③ 건물 외관 재료

(5) 실내 마감재료

- ① 시공 상태 ② 안정성 ③ 내구성 ④ 마감재료

(6) 동선 : 4개항

(7) 적정 치수 : 3개항

- (8) 환경 유지 관리 : 3개항
- (9) 채광 : 2개항
- (10) 소음 : 2개항
- (11) 진동 : 3개항
- (12) 환기 : 2개항
- (13) 설비 : 2개항 (36개 항목)

전체 만족도

- ① 신축 학교에 대한 만족도는?
- ② 상급 학교도 같은 종류의 학교이기를 원하십니까?

개선우선 항목 분석

1. 설문조사 결과에 대한 인자분석 실시 -- 주요 요인을 찾음

결과 -- 10개 인자 (eigen value 1.0 이상)

인자 1 : 체감 실내환경 (냉방, 난방, 환기, 실내마감재)

인자 2 : 외부 공간 (자연학습장, 놀이시설, 휴식공간 등)

인자 3 : 내부 마감재 (실내마감재료의 안정성,

실내마감재료의 내구성, 교실 창, 교실내 진동 등)

인자 4 : 교실규모 및 가변성 (화장실 규모 등)

인자 5 : 진동과 치수 (계단에서 진동, 화장실 치수 등)

인자 6 : 시공 상태

인자 7 : 외관

인자 8 : 복도 공간

인자 9 : 소음

인자 10 : 채광

2. 인자점수의 계산

3. 전체 만족도 산출 : 전체 만족도 조사 항목을 이용

4. 인자별 중요도 추출 : 베타계수

중회귀분석에서 베타계수, 즉

$$(\text{전체 만족도})_i = \beta_0 + \beta_1(\text{인자점수1})_i + \cdots + \beta_m(\text{인자점수}m)_i$$

출력 결과에서 베타계수가 각 인자의 중요도에 해당

예.

	비표준화된 계수 (표준편차: 0.085)	표준화된 계수 (beta)	t	p -값	중요도 순위
상수	3.555		42.04	.000	
체감실내환경	0.232	0.192	2.729	.007	4
외부공간	0.412	0.341	4.850	.000	1
내부마감재	0.324	0.269	3.816	.000	2
교실규모	0.051	0.042	0.603	.547	10
진동	0.202	0.168	2.385	.019	6
시공상태	0.271	0.225	3.198	.002	3

5. 요인별 만족도 추출

각 요인에 포함된 변수들 가운데 요인적재가 0.4 이상인 변수들만 뽑아, 이들의 전체 평균값을 만족도로 간주
예.

인자	체감실내환경	외부공간	내부마감재	진동	시공상태
만족도	3.000	3.273	3.132	2.854	2.640

6. 요인의 중요도 지수

베타계수의 Z-score를 구하고, Z-score의 Z-%값이 지수
예.

인자	체감실내환경	외부공간	내부마감재	진동	시공상태
베타계수	0.192	0.341	0.269	0.169	0.225
Z-score	-0.198	1.925	0.899	-0.540	0.273
Z-%	0.425	0.973	0.814	0.298	0.608

Z-score : 베타계수를 표준화시킨 값

Z-% : $P(Z < Z\text{-score})$ 값

7. 요인의 불만족도 지수

불만족도(“6-만족도”)를 구하고,

이들의 Z-score를 구한 다음, Z-score의 Z-%값이 지수

예.

인자	체감실내환경	외부공간	내부마감재	진동	시공상태
만족도	3.00	3.27	3.13	2.85	2.64
불만족도	3.00	2.73	2.87	3.15	3.36
Z-score	0.426	-0.512	-0.018	0.955	1.695
Z-%	0.665	0.305	0.500	0.830	0.995

6. 개선 우선순위 판별지수

$$\text{개선우선도판별지수} = \frac{2(s \cdot w)}{s + w},$$

다만, s : 요인의 중요도 지수

w : 요인의 불만족도 지수

$$(\text{예 : 체감실내환경} = \frac{2(0.4247 \times 0.6650)}{0.4247 + 0.6650} = 0.5184)$$

결과.

	시공상태	내부마감재	체감실내환경	외부공간	진동	복도공간
개선우선도	0.743	0.620	0.518	0.464	0.439	0.278
순위	1	2	3	4	5	6