***Concept of Factor Analysis***

**Population, principal concept**

Suppose the random variables *X*1, *X*2 and *X*3 have the covariance matrix



It may be verified that the eigen value – eigen vector pairs are

*λ*1 = 5.83 *e*1′ = [0.383, -0.924, 0]

*λ*2 = 2.00 *e*2′ = [0, 0, 1]

*λ*3 = 0.17 *e*3′ = [0.924, 0.383, 0]

Results: Let Σ be the covariance matrix associated with the random vector *X′* = [*X*1, *X*2,…, *Xp*]. Let Σ have the eigen value – eigen vector pairs (*λ*1, *e*1), (*λ*2, *e*2),…, (*λp*, *ep*) where *λ*1 ≥ *λ*1 ≥ … ≥ *λp* ≥ *c*. Then the *i*th principal component is given by

*Yi* = *ei*′ *X* = *ei*1 *X*1 + *ei*2 *X*2 +…+ *eipXp*, *i* = 1, 2,…, *p*

with these choices, var(*Yi*) = *ei*′ Σ*ei* = *λi*, *i* = 1 (1) *p*

cov(*Yi*, *Yk*) = *ei*′ Σ*ek* = 0, *i* ± *k*

If some *λi* are equal, the choices of the corresponding coefficient vectors *ei* and hence *Yi* are not unique.

Total population variance = σ11+ σ22+…+ σ*pp*

= *λ*1 + *λ*1 + … + *λp*

Proportion of total population variance due to *k*th principal component = 

Results: If *Y*1 = *e*1′ *X*, *Y*2 = *e*2′ *X*, …, *Yp* = *ep*′ *X* are the principal components obtained from the covariance matrix Σ, then



Are the correlation coefficients between the components *Yi* and the variables *Xk*. Here, (*λ*1, *e*1), (*λ*2, *e*2),…, (*λp*, *ep*) are the eigen value – eigen vector pairs for Σ.

Therefore the principal components become

*Y*1 = *e*1′ *X* = 0.383*X*1 – 0.924*X*2

*Y*2 = *e*2′ *X* = *X*3

*Y*3 = *e*3′ *X* = 0.924*X*1 + 0.383*X*2

The variable *X*3 is one of the principal components, because it is uncorrelated with the other two variables.

v(*Y*1) = 5.83 = *λ*1

cov(*Y*1, *Y*2) = cov((0.383 *X*1 – 0.924 *X*2) *X*3)

= 0.383 cov(*X*1 *X*2) – 0.924 cov(*X*1 *X*2)

= 0

It is also readily apparent that

σ11+ σ22+ σ33 = 1+5+3 = *λ*1 + *λ*1 + *λ*3 = 8

The proportion of total variance accounted by the first principal component is 

Further, the first two components account for a proportion (5.83+2)/8 = 0.98 of the population variable.

In this case, the components *Y*1 and *Y*2 could replace the original three variables with little loss of information.

Using relation (\*), we obtain,



Notice here that the variable *X*2 with coefficient –0.924 receives the greatest weight or the component *Y*1. It also has the largest correlation (i.e., absolute value) with *Y*1. The correlation of *X*1 with *Y*1 is almost as large as that for *X*2, indicating that the variables are about equally important to the first principal component. The relative sizes of the coefficients of *X*1 and *X*2 suggest, however, that *X*2 contributes more to the determination of *Y*1 than does *X*1. Since, in this case, both coefficients are reasonably large and they have opposite signs. We xxxxx agree that both xxxxxxxxxxxx in the independent of *Y*1.

Finally,

(as it shows)

The remaining correlations can be neglected since the third component is unimportant.

I. xxxxxx of correlation, in place of *X*, we take *Z*

Proportion of (standardized) population variance due to *k*th principal component =  where *λk*‘s are the eigen values of *p.*

Principal components obtained from covariance and correlation matrices are different.

Consider the covariance matrix  and the derived correlation matrix 

The eigen value – eigen vector pairs from Σ are

*λ*1 = 100.16 *e*1′ = [0.040, 0.999]

*λ*2 = 0.84 *e*2′ = [0.999, - 0.040]

Similarly, the eigen value – eigen vector pairs from *ρ* are

*λ*1 = 1+ *ρ* = 1.4 *e*1′ = [0.707, 0.707]

*λ*2 = 1- *ρ* = 0.6 *e*2′ = [0.707, - 0.707]

The respective principal components become

Σ: *Y*1 = 0.040*X*1 + 0.999*X*2

*Y*1 = 0.999*X*1 – 0.040*X*2

and *ρ*: *Y*1 = 0.707 *Z*1 +0.707 *Z*2

= 0.707 () + 0.707 ()

= 0.707 () + 0.0707 ()

*Y*2 = 0.707 () - 0.0707 ()

Because of its large variance, *X*2 completely dominates the first principal component determined from Σ. Moreover, the first principal component explains a proportion.

of the total population variance.

When the variables *X*1, *X*2 are standardized, however, the resulting variables contribute equally to the principal components determined from *ρ*.



In this case, the first principal component explains a proportion  of the first (standardized) population variance.

**Factor Analysis**

The essential purpose of factor analysis is to describe, if possible, the covariance relationships among many variable sin terms of a few underlying, but unobservable random quantities called factors. Basically, the factor model is motivated by the following arguments:-

Suppose variables can be grouped by their correlations i.e., suppose all variables within or particular group are highly correlated among themselves, but have relatively small correlation with variables in a different group. Then it is conceivable that each group of variable represents a single underlying constructor factor that is responsible for the observed correlations. For example, correlations from the group of test scores in Classics, French, English, Mathematics, and Music collected by Spearman suggested an underlying “intelligence” factor. A second group of variables representing physical fitness scores, if available, might correspond to another factor. It is this type of structure that factor analysis xxxxx to confirm.

Factor analysis can be considered an extension of principal component analysis. Both can be viewed as attempts to approximate the covariance matrix Σ. However, the approximation based on the factor analysis model is more elaborate. The primary question in factor analysis is whether the data are consistent with a prescribed structure.

**The orthogonal factor model**

The observable random vector *X* with *p* components has mean *L* and covariance matrix Σ.

The factor model postulates that *X* is linearly dependent upon a few unobservable random variables *F*1, *F*2… *Fm* called common factors and *p* additional sources of variation ε1, ε2 … ε*p* called errors or sometimes specific factors.

In particular, the factor analysis model is,



In matrix notation,

*X* – *μ* = *LF* + ε*p*} → unobservable

The coefficient *lij* = Loading of *i*th variable on the *j*th factor.

**Covariance structure for the orthogonal factor model**

1. Cov (*X*) = *LL*′ + *ψ*



2) Cov(*X*, *F*) = *L*

Or, Cov(*Xi*, *Fj*) = *lij*



The *i*th communality is the sum of squares of the loadings of the *i*th variable on the *m* common factors.

Verifying the relation Σ = *LL*′ + *ψ*



The equality



i.e., Σ = *LL*′ + *ψ*



*p*(diagonal) +….

Factor model assumes that *p* + *p* (*p* – 1)/2 = *p* (*p* + 1)/2. Variance and covariance for *X* …..

When m = p xxxx covariance matrix Σ can be reproduced exactly as *LL*′. So, *ψ* can be the zero matrix.

For example, *X* contains *p* =12 variable and the factor model with *m* = 2 is appropriate then *p* (*p* + 1)/2 = 78 elements of Σ are described in xxxx of the *mp* + *p* = 36 parameters *lij* and *Yi* of the factor model.

Xxxx matrix *S* – (*LL*′ + *ψ*)



Proportion of total sample variance due to the factor =  for a factor analysis of *S*

Consumer preference study =  for a factor analysis of *R*

Example: Step 1 7-point semantic difference scale





Cumulative proportion =  = 93

Table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimated factor loadings  F1 F2 | | Communalities | Specific variance |
| 1 | 0.56 | 0.82 | 0.98 | 0.02 |
| 2 | 0.76 | -0.53 | 0.88 | 0.12 |
| 3 | 0.65 | 0.75 | 0.98 | 0.02 |
| 4 | 0.94 | -0.10 | 0.89 | 0.11 |
| 5 | 0.80 | -0.54 | 0.93 | 0.07 |
| Eigen values | 2.85 | 1.81 |  |  |
| Cumulative proportion of total standardized sample variance |  | 0.931 |  |  |

Now, 

R-LL’-Y’=Residual

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variable | Estimated factor loadings  F1 F2 | | Rotated estimated factor loadings  F1 F2 | Communalities |
| 1 | 0.56 | 0.82 | 0.02 0.99 | 0.98 |
| 2 | 0.76 | -0.53 | 0.94 -0.01 | 0.88 |
| 3 | 0.65 | 0.75 | 0.13 0.98 | 0.98 |
| 4 | 0.94 | -0.10 | 0.84 0.43 | 0.89 |
| 5 | 0.80 | -0.54 | 0.97 -0.02 | 0.93 |
| Cumulative property (standardized sample explained) |  | 0.932 | 0.507 0.932 |  |

Oblique Rotation:

Orthogonal rotations are appropriate for a factor model in which the common factors are xxxx to be independent.

The best will be that whose xxxxx matrix is small.

Bartlett’s Correction

*H*0 : Σ(*p, p*) = *LL*′ + *ψ*

*H*1 : Σ is any other positive definite matrix.

Using Bartlett’s Correction we reject *H*0 at the α-level of significance



must be positive. It follows that .