Factor Analysis

The essential purpose of factor analysis is to describe, if possible, the covariance relationships among many variables in terms of a few underlying, but unobservable, random quantities called *factors*. Basically, the factor model is motivated by the following argument: Suppose variables can be grouped by their correlations. That is, suppose all variables within a particular group are highly correlated among themselves, but have relatively small correlations with variables in a different group. Then it is conceivable that each group of variables represents a single underlying construct, or 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 seeks to confirm.

Factor analysis can be considered an extension of principal component analysis. Both can be viewed as attempts to approximate the covariance matrix $\mathbf{\Sigma}$. 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.

- Measurement is necessary especially when it is unobservable
- Any behaviour pattern should be measured numerically
- Charles Spearman – Measuring of human intelligence using Factor analysis
- IQ level of Persons? - human intelligence is not visible- unobservable- how do you measure such unobservable?

Few examples – cannot be measured easily

- A psychologist is interested to measure mental ability of a person
- A safety measure manager is interested to measure safety environment of his workplace
- A supply chain analyst is interested to measure supply chain coordination
- A marketing manager is interested to measure the purchase intension of customers
- A welfare officer is interested to measure the ability of worklife in factories

There are certain things which are hidden unobservables and are not directly measurable but those things are very important. Since they causes something to happen. Because of the presence of unobservable things, these manifested to different symptoms.

- Hence the hidden concepts/ constructs – not directly measured-
- concepts/ constructs - manifested into different symptoms
- Example : Construct / factor– Risk taking in nature causing

  Symptoms- rule breaker

Construct factor (F) → quantify F

Risk taking nature

to measure

Causing

Symptoms | Rule Breaking

manifest variables, X ← observing

Observing numbers

$$X_{p\times 1} = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix},$$

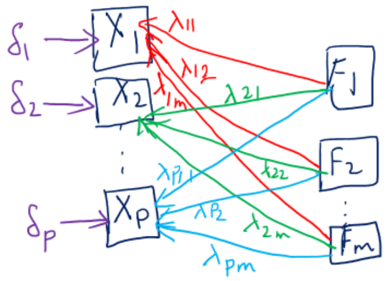
P manifest variables from population

mean vector of X, $\mu_{p\times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}$

$$Cov(X) = \Sigma = \begin{array}{c|ccc} & x_1 & x_2 & \cdots x_p \\ \hline x_1 & \sigma_{11} & \sigma_{12} & \cdots \sigma_{1p} \\ x_2 & \sigma_{12} & \sigma_{22} & \cdots \sigma_{2p} \\ x_3 & \vdots & & \\ x_p & \sigma_{1p} & \sigma_{2p} & \cdots \sigma_{pp} \end{array} \quad p\times p$$

Consider m factors

$$F_{m\times 1} = \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix} \qquad m < p$$



m factors influencing

X, is measured through λ

whatever you observed in x, because of the causal factor F.

- The F_m factors collectively cannot explain everything about X, So add error term

No relationship, between Fs. They are independent Orthogonal

$$X_1 = \mu_1 + \lambda_{11} F_1 + \lambda_{12} F_2 + \dots \lambda_{1m} F_m + \delta_1$$

$$X_2 = \mu_2 + \lambda_{21} F_1 + \lambda_{22} F_2 + \dots \lambda_{2m} F_m + \delta_2$$

$$\vdots$$

$$X_j = \mu_j + \lambda_{j1} F_1 + \lambda_{j2} F_2 + \dots \lambda_{jm} F_m + \delta_j$$

$$\vdots$$

$$X_p = \mu_p + \lambda_{p1} F_1 + \lambda_{p2} F_2 + \dots \lambda_{pm} F_m + \delta_p$$

$$\underbrace{\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_j \\ \vdots \\ X_p \end{bmatrix}_{p\times 1}}_{X} = \underbrace{\begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_j \\ \vdots \\ \mu_p \end{bmatrix}_{p\times 1}}_{\mu} + \underbrace{\begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & & & \\ \lambda_{j1} & \lambda_{j2} & \cdots & \lambda_{jm} \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{bmatrix}_{p\times m}}_{\lambda} \underbrace{\begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \\ \vdots \\ F_m \end{bmatrix}_{m\times 1}}_{F} + \underbrace{\begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_j \\ \vdots \\ \delta_p \end{bmatrix}_{p\times 1}}_{\delta}$$

$$\underset{(p\times 1)}{X} = \underset{(p\times 1)}{\mu} + \underset{(p\times m)}{\lambda} \underset{(m\times 1)}{F} + \delta_{(p\times 1)}$$

$$X - \mu = \lambda F + \delta \quad \text{— (1) Factor model}$$

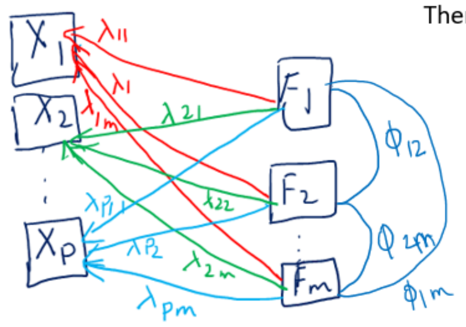
Assumptions:

$E(x) = \mu \qquad E(F) = 0 \qquad E(\delta) = 0$

$Cov(x) = \Sigma \qquad Cov(F) = E(FF^T) = I$

$Cov(F\delta) = 0 \qquad Cov(\delta) = \Psi = \begin{bmatrix} \psi_{11} & 0 & \cdots & 0 \\ 0 & \psi_{22} & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \cdots & \psi_{pp} \end{bmatrix}$

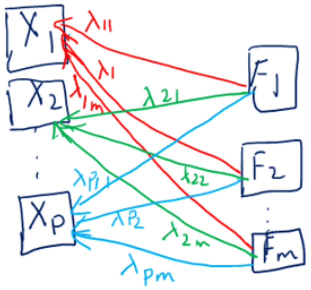
If the factor model hold all these assumptions, then it is called Exploratory orthogonal Factor model



Oblique factor model- When you allow the covariance between factor
Then, it is called Oblique FM

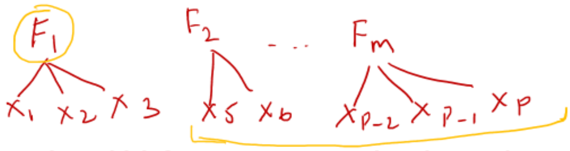
Factors are orthogonal, since they are independent

- Types of Factor Model: **1.Exploratory Factor Model** $X-\mu = \lambda F + \delta$



- **Each X is dependent on all F.**
- **Which are the X variables are basically coming out of F ?**
- **So allow all the full sets, find the factors that are linked to what type of variables**

2. **Confirmatory Factor Model**



Which of the manifest var representing which factor. X1,X2,X3 related F1 and so on -> need to confirm whether the connections are correct or wrong. : Hypothesis.
Through confirmatory FM we are going to prove hypothesis.

Key question 1

- What is common in the examples given?
  - The variables to be measured are unobservable or hidden or latent and known as constructs or factors : F
  - They all manifest some sympt oms which can be observed and measured and known as manifest variables : X

Factor analysis quantifies these constructs (factors with the help of the manifest variables)

We are doing this both in EFA, CFA

Key question 2

- The lessor the information contents, the better the design
  - When I have P number of variables, I have p dimensional information and I don’t want this much information. I want only 2 D only.
- FA also reduces dimensions

Purpose:

- The purpose of FA is to describe, if possible, the covariance relationships among many variables in terms of A FEW underlying, but unobservable, random quantities called FACTORS

Variability explanation:

- $Cov(X)=\Sigma$

$$X-\mu = \lambda F + \delta \quad \text{— (1) Factor model}$$

- Know: what is the variability you want to explain using FM?
- Whose variability? – X, what is variability? Cov(X)

$$Cov(X)=E[(X-\mu)(X-\mu)^T]$$

$$= E\left[(\lambda F+\delta)(\lambda F+\delta)^T\right]$$

$$= E\left[\lambda F F^T \lambda^T + \lambda F \delta^T + \delta F^T \lambda^T + \delta \delta^T\right]$$

λ – constant, $= \lambda E(FF^T)\lambda^T + \lambda E(F\delta^T) + E(\delta F^T)\lambda^T + E(\delta\delta^T)$

From assumptions, $= \lambda I \lambda^T + \lambda(0) + (0)\lambda^T + \psi$

$$\Sigma = \lambda\lambda^T + \psi \quad \text{— (2)}$$

if you have λ, you know ψ

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & & & \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{bmatrix}_{p\times m}$$

λ is loading matrix

$$x_j = \lambda_{j1}F_1 + \lambda_{j2}F_2 + \cdots + \underline{\lambda_{jk}F_k} + \cdots \lambda_{jm}F_m + \delta_j$$

$j = 1, 2 \ldots p$

$k = 1, 2 \ldots m$

general loading $= \lambda_{jk}F_k$

loading of k^{th} factor j^{th} x.

$$\lambda\lambda^T = \begin{bmatrix} \lambda_{11} & \lambda_{12} & \cdots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \cdots & \lambda_{2m} \\ \vdots & & & \\ \lambda_{p1} & \lambda_{p2} & \cdots & \lambda_{pm} \end{bmatrix}_{p\times m} \begin{bmatrix} \lambda_{11} & \lambda_{21} & \cdots & \lambda_{p1} \\ \lambda_{21} & \lambda_{22} & & \lambda_{p2} \\ \vdots & \vdots & & \vdots \\ \cancel{\lambda_{p1}}\ \lambda_{1m} & \cancel{\lambda_{p2}}\ \lambda_{2m} & & \lambda_{pm} \end{bmatrix}_{m\times p}$$

$$= \begin{bmatrix} \sum_{k=1}^{m} \lambda_{1k}^2 & \sum_{k=1}^{m} \lambda_{1k}\lambda_{2k} & \cdots & \sum_{k=1}^{m} \lambda_{1k}\lambda_{pk} \\ \sum_{k=1}^{m} \lambda_{1k}\lambda_{2k} & \sum_{k=1}^{m} \lambda_{2k}^2 & \cdots & \sum_{k=1}^{m} \lambda_{2k}\lambda_{pk} \\ \vdots & & & \vdots \\ \sum_{k=1}^{m} \lambda_{1k}\lambda_{pk} & \sum_{k=1}^{m} \lambda_{2k}\lambda_{pk} & \cdots & \sum_{k=1}^{m} \lambda_{pk}^2 \end{bmatrix}$$

Variance?
Covariance?

$$\Sigma = \lambda\lambda^T + \psi = \begin{bmatrix} \sum_{k=1}^{m} \lambda_{1k}^2 & \sum_{k=1}^{m} \lambda_{1k}\lambda_{2k} & \cdots & \sum_{k=1}^{m} \lambda_{1k}\lambda_{pk} \\ \sum_{k=1}^{m} \lambda_{1k}\lambda_{2k} & \sum_{k=1}^{m} \lambda_{2k}^2 & \cdots & \sum_{k=1}^{m} \lambda_{2k}\lambda_{pk} \\ \vdots & & & \\ \sum_{k=1}^{m} \lambda_{1k}\lambda_{pk} & \sum_{k=1}^{m} \lambda_{2k}\lambda_{pk} & \cdots & \sum_{k=1}^{m} \lambda_{pk}^2 \end{bmatrix} + \begin{bmatrix} \psi_{11} & 0 & \cdots & 0 \\ 0 & \psi_{22} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \psi_{pp} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \vdots & & & \vdots \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}_{p\times p}$$

$$\sigma_{11} = \sum_{k=1}^{m} \lambda_{1k}^2 + \psi_{11}$$

$$\sigma_{22} = \sum_{k=1}^{m} \lambda_{2k}^2 + \psi_{22}$$

$$\sigma_{jj} = \sum_{k=1}^{m} \lambda_{jk}^2 + \psi_{jj}$$

$$\sigma_{jj} = \sum_{k=1}^{m} \lambda_{jk}^2 + \psi_{jj}$$

understand the variability part,

total variability

X_j: $\boxed{\sum_{k=1}^{m} \lambda_{jk}^2 \;\Big|\; \psi_{jj}}$

Contributed by the m common factors → Commonality (h_j)

not contributed by the common factors → unique, specific variance of X_j

The common factors are not able to explain this much of variability, this is unexplained

What is the contribution of first factor? λ_{11}^2

$$Cov(xF) = E[(x-\mu)(F-E(F))^T]$$
$$= E[(x-\mu)F^T]$$

We know that
$x - \mu = \lambda F - \delta$

$$(x-\mu)F^T = (\lambda F - \delta)F^T$$
$$= \lambda F F^T - \delta F^T$$

$$\therefore E(x-\mu)F^T = E(\lambda F - \delta)F^T$$
$$= E(\lambda F F^T) - E(\delta F^T)$$
$$= \lambda E(F F^T) + 0$$
$$= \lambda$$

loading matrix.

An example

(Ref: Lawley and Maxwell, 1971; taken from Johnson and Wichern, Appl Mul Stat Ana, 2002)

- Lawley and Maxwell (1971) studied the general intelligence of 220 students

Student no.	Gaelic (X1)	English (X2)	History (X3)	Arithmatic (X4)	Algebra (X5)	Geometry (X6)
1						
2						
...						
220						

Factor ability criteria- can go for FA or not?

An example (contd.)

- The correlation matrix

	X1	X2	X3	X4	X5	X6
X1	1.00	0.44	0.41	0.29	0.33	0.25
X2		1.00	0.35	0.35	0.32	0.33
X3			1.00	0.16	0.19	0.18
X4				1.00	0.60	0.47
X5					1.00	0.46
X6						1.00

Exploratory Factor Analysis

- Model Estimation
- Model Adequacy Testing

Data : $X_{n \times p}$ on p variable(s)

$\Downarrow$

$S_{p \times p}$

$X - \mu = \lambda F + \delta$

$\underset{(p\times p)}{\Sigma} = \underset{(p\times m)}{\lambda} \underset{(m\times p)}{\lambda^T} + \underset{(p\times p)}{\psi}$ (diagonal)

Estimate $\hat{\lambda}, \hat{\psi}$

What is known to us? We know $\hat{\Sigma} = S_{p \times p}$

If think that n is large sample space.
Appropriate Sampling strategy is applied

$$\hat{\Sigma} = \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} & \cdots & \hat{\sigma}_{1p} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} & \cdots & \hat{\sigma}_{2p} \\ \vdots & & & \\ \hat{\sigma}_{1p} & \hat{\sigma}_{2p} & \cdots & \hat{\sigma}_{pp} \end{bmatrix} = S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} & \cdots & s_{2p} \\ \vdots & & & \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

In general,

$\hat{\sigma}_{jk} = s_{jk}$, $j = 1, 2 \cdots p$; $k = 1, 2 \cdots p$

- Now, we can assume that under root estimation that what will happen ultimately, if our sampling size is appropriate at recreating will have to capture properly, the behaviour the covalent structure.

$\hat{\sigma}_{jk} = s_{jk}$, $j = 1, 2 \cdots p$; $k = 1, 2 \cdots p$

- each of the element in ∑-very close.

$\hat{\Sigma} = \hat{\lambda}\hat{\lambda}^T + \hat{\psi}$

We know S only. So Initialize $\hat{\lambda}, \hat{\psi}$ then find out $\hat{\Sigma}$. Then Compare this value with the individual elements of S. And we find out a situation where the element of these matrices (s, $\hat{\Sigma}$) will be as small as possible with certain threshold limit.

We want to estimate : $\hat{\Sigma} = \hat{\lambda}\hat{\lambda}^T + \hat{\psi}$

S ; $\hat{\lambda}$? ; $\hat{\psi}$?

- so we initialize values for this
- And then we find out this value, then we compare this value with s individual elements of the vector.
- And if we find out a situation where the difference between the element of these two matrices will be as small as possible within certain threshold limit.
- Trace (∑)= $\sum_{j=1}^{p} \sigma_{jj}$; Trace (S)= $\sum_{j=1}^{p} s_{jj}$
- Error – difference between S (sample cov) - ∑ (population cov)
- sum square error should be Minimized

$(S-\Sigma)^2$

minimize $G = tr\left[(S-\Sigma)^T(S-\Sigma)\right]$

This leads to 2nd norm

$= \|(S-\Sigma)^2\|$

$= \|(S-\lambda\lambda^T+\psi)\|$

We want

① Initial value of ψ

② The number of factors — m value

Methods available for estimation

1. Principal Component method
2. Principal factor method
3. Maximum likelihood method

1. PCA: Covariance matrix is decomposed into Q_j : eigen values, e_j – eigenvector pairs.

   $\Sigma = \sum_{j=1}^{p} Q_j e_j e_j^T$: Spectral decomposition

$$\underset{p\times p}{\Sigma} = Q_1 e_1 e_1^T + Q_2 e_2 e_2^T + \dots + Q_p e_p e_p^T$$

$$= \underbrace{\left[\sqrt{Q_1}\,e_1 \;\; \sqrt{Q_2}\,e_2 \;\; \cdots \;\; \sqrt{Q_m}\,e_m \;\; \cdots \;\; \sqrt{Q_p}\,e_p\right]}_{\lambda} \underbrace{\begin{bmatrix} \sqrt{Q_1}\,e_1^T \\ \sqrt{Q_2}\,e_2^T \\ \vdots \\ \sqrt{Q_m}\,e_m^T \\ \vdots \\ \sqrt{Q_p}\,e_p^T \end{bmatrix}}_{\lambda^T}$$

$=$

$\Sigma = \lambda\lambda^T$

No. of factors $m < p$.

$m \subset p$

$$\Sigma = \lambda \lambda^T + \psi$$

$$\lambda = \left[\sqrt{Q_1}\, e_1 \quad \sqrt{Q_2}\, e_2 \; \ldots \; \sqrt{Q_m}\, e_m\right]$$

$$= \left[\sqrt{Q_{m+1}}\, e_{m+1} \; \ldots \; \sqrt{Q_p}\, e_p\right] \begin{bmatrix} \sqrt{Q_{m+1}}\, e_{m+1} \\ \vdots \\ \sqrt{Q_p}\, e_p \end{bmatrix}$$

Loadings

$F_1 : \sqrt{Q_1}\, \hat{e}_1$ $\quad \hat{e}_1 \rightarrow$ eigenvalues $\begin{bmatrix} e_{11} \\ e_{12} \\ e_{1p} \end{bmatrix}_{p \times 1}$

$F_2 : \sqrt{Q_2}\, \hat{e}_2$

$\vdots$

$F_m : \sqrt{Q_m}\, \hat{e}_m$

- Sample covalence matrix $S = \hat{\lambda} \hat{\lambda}^T + \hat{\psi}$ -> Spectral decomposition of S
  $\hat{\psi} = S - \hat{\lambda} \hat{\lambda}^T$

1. Find S
2. Find out eigenvalue, eigenvector of S
3. Select m — ?
4. Obtain $\hat{\lambda}$
5. Obtain $\hat{\psi}$

$\hat{\lambda} \hat{\lambda}^T = S - \psi$ ↓ Initialize

→ & PCA → factors ; How to decide the components?

SASTRA

PCA $|S - \lambda I| = 0$

$|S - QI| = 0$

order = no. of roots

① $S = \begin{bmatrix} 100 & 50 & -30 \\ 50 & 64 & -20 \\ -30 & -20 & 49 \end{bmatrix}_{3 \times 3}$

$\hat{Q}_1 = 148.23$

$\hat{Q}_2 = 36.00$

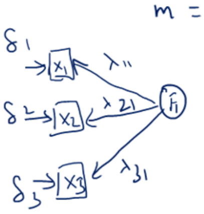
$Q_3 = 28.70$

$\hat{e}_1 : (S - Q_1 I)\, e_1 = 0$, subject to $e_1^T e_1 = I$

$\hat{e}_2 : (S - Q_2 I)\, e_2 = 0$, subject to $e_2^T e_2 = I$

$\hat{e}_3 : (S - Q_3 I)\, e_3 = 0$, subject to $e_3^T e_3 = I$

m = 1 factor model



$S = \hat{\lambda} \hat{\lambda}^T + \hat{\psi}$

$\hat{\lambda}_{p \times m} = \hat{\lambda}_{3 \times 1} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix}$ $\quad \hat{\psi} = \begin{bmatrix} \hat{\psi}_{11} & 0 & 0 \\ 0 & \hat{\psi}_{22} & 0 \\ 0 & 0 & \hat{\psi}_{33} \end{bmatrix}$

② $\hat{e}_1 = \begin{bmatrix} -0.77 \\ -0.54 \\ 0.34 \end{bmatrix} \quad \hat{e}_2 = \begin{bmatrix} 0.21 \\ 0.29 \\ 0.93 \end{bmatrix} \quad \hat{e}_3 = \begin{bmatrix} 0.60 \\ -0.79 \\ 0.11 \end{bmatrix}$

③ $\lambda_1 = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix} \quad \lambda_2 = \begin{bmatrix} \lambda_{12} \\ \lambda_{22} \\ \lambda_{32} \end{bmatrix} \quad \lambda_3 = \begin{bmatrix} \lambda_{13} \\ \lambda_{23} \\ \lambda_{33} \end{bmatrix}$

But we need $m = 1$

④ $\hat{\lambda}_1 = \sqrt{a_1}\,\hat{e}_1 = \sqrt{148.23} \begin{bmatrix} -0.77 \\ -0.54 \\ 0.34 \end{bmatrix} = \begin{bmatrix} -9.37 \\ -6.57 \\ 4.75 \end{bmatrix}$

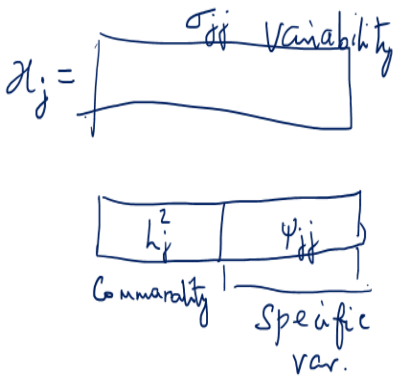
$$\lambda\hat{\lambda}^T = \begin{bmatrix} -9.57 \\ -6.57 \\ 4.75 \end{bmatrix} \begin{bmatrix} -9.57 & -6.57 & 4.75 \end{bmatrix}$$

$$= \begin{bmatrix} 87.80 & 61.56 & -44.51 \\ 61.56 & 43.16 & -31.21 \\ -44.51 & -31.21 & 22.56 \end{bmatrix}$$

Should be 0.

$$\hat{\psi} = S - \lambda\hat{\lambda}^T = \begin{bmatrix} 100 & 50 & -30 \\ 50 & 64 & -20 \\ -30 & -20 & 49 \end{bmatrix} - \begin{bmatrix} 87.80 & 61.56 & -44.57 \\ 61.56 & 43.16 & -31.21 \\ -44.51 & -31.21 & 22.56 \end{bmatrix} = \begin{bmatrix} 12.20 & -11.56 & 14.51 \\ -11.56 & 20.84 & 11.21 \\ 14.51 & 11.21 & 26.44 \end{bmatrix}$$

2. Principal Factor Method



$$\Sigma = \lambda\lambda^T + \psi$$

$$S = \lambda\hat{\lambda}^T + \psi$$

$$\lambda\hat{\lambda}^T = \underline{\underline{S - \psi}}$$

In PCA S is decomposed. But here $S - \psi$ should be decomposed.

$$S = \begin{bmatrix} s_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} & \cdots & s_{2p} \\ \vdots & & & \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix} \quad \psi = \begin{bmatrix} \psi_{11} & 0 & \cdots & 0 \\ 0 & \psi_{22} & \cdots & 0 \\ \vdots & & & \\ 0 & 0 & \cdots & \psi_{pp} \end{bmatrix}$$

$$S - \psi = \begin{bmatrix} s_{11} - \psi_{11} & s_{12} & \cdots & s_{1p} \\ s_{12} & s_{22} - \psi_{22} & \cdots & \\ \vdots & & & \\ s_{1p} & \cdots & & s_{pp} - \psi_{pp} \end{bmatrix}$$

$$= \begin{bmatrix} h_1^2 & s_{12} & \cdots & s_{1p} \\ s_{12} & h_2^2 & \cdots & s_{2p} \\ s_{1p} & s_{2p} & \cdots & s_{pp} \end{bmatrix}$$

$\sigma_{jj} = h_j^2 + \psi_{jj}$

$s_{jj} = h_j^2 + \psi_{jj}$

$s_{jj} - \psi_{jj} = h_j^2$

eigen value, eigen vectors

$\hat{a}_j^* \; e_j^* \quad j = 1, 2 \cdots p$

exactly the same as PCA method.

$$\hat{L}^* = \left[\sqrt{\hat{a}_1^*}\, e_1^* \quad \sqrt{\hat{a}_2^*}\, e_2 \quad \cdots \quad \sqrt{\hat{a}_m^*}\, e_m \right]$$

$p - m \rightarrow$ not contributing in explaining variability

- In this method, $S - \psi$ should be decomposed.

  $\hat{\psi}_0$ – Iterative procedure.

- we initialize, then find estimate, then again you see that what is that the size coming, and then that matrix can be acceptable or not, what is the variance explained, what is not explained, then again you go on repeating these things and -the model is to be accepted.

Maximum likelihood method

- Find the likelihood function- difficult

An example (contd.)

- The correlation matrix

	X1	X2	X3	X4	X5	X6
X1	1.00	0.44	0.41	0.29	0.33	0.25
X2		1.00	0.35	0.35	0.32	0.33
X3			1.00	0.16	0.19	0.18
X4				1.00	0.60	0.47
X5					1.00	0.46
X6						1.00

Variables	Factor-1	Factor-2	Communa-lities
Gaelic	0.553	0.429	0.49
English	0.568	0.288	0.41
History	0.392	0.450	0.36
Arith-matic	0.740	-0.273	0.63
Algebra	0.724	-0.211	0.59
Geometry	0.595	-0.132	0.37
% variance explained	0.37	0.10	

What are these values in the columns?

Square the column values , you will get eigen values. Row wise square and sum =Communality

A **communality** is the extent to which an item correlates with all other items. Higher **communalities** are better. If **communalities** for a particular variable are low (between 0.0-0.4), then that variable may struggle to load significantly on any **factor**.

	F_1	F_2		F_m	
$\to x_1$	λ_{11}	λ_{12}	....	λ_{1m}	$\sum_{k=1}^{m} \lambda_{1k}^2 = h_1^2$ Com
x_2	λ_{21}				
⋮					
x_p	λ_{p1}				
	$\sum_{j=1}^{p} \lambda_{j1}^2$	$\sum_{j=1}^{p} \lambda_{j2}^2$			
	Q_1 eigen	Q_2 eigen			

Maximum likelyhood

$x_i \sim N_p(\mu, \Sigma) = N_p(\mu, \lambda\lambda^T + \psi)$, $i = 1, 2 \ldots n$ Find MLH then log LH

$$L(x \mid \mu, \Sigma) = -\frac{n}{2}\log|2\pi\Sigma| - \frac{n}{2} tr(\Sigma^{-1}S) - \frac{n}{2}(\bar{x}-\mu)\Sigma^{-1}(\bar{x}-\mu)^T$$

Substitute $\hat{\mu} = \bar{x}$

$$L(x \mid \hat{\mu}, \Sigma) = -\frac{n}{2}\log|2\pi\Sigma| - \frac{n}{2} tr(\Sigma^{-1}S)$$

$$\hat{\Sigma} = \hat{\lambda}\hat{\lambda}^T + \hat{\psi}$$

$$L(x \mid \hat{\mu}, \hat{\Sigma}) = -\frac{n}{2}\left[\log|2\pi(\hat{\lambda}\hat{\lambda}^T + \hat{\psi})| + tr\left\{(\hat{\lambda}\hat{\lambda}^T + \hat{\psi})^{-1}S\right\}\right]$$

- Lawley has given some formulation

Giving a normal equation for solving, $S\hat{\psi}^{-1}\hat{\lambda} = \hat{\lambda}(I + \hat{\lambda}^T\hat{\psi}^{-1}\hat{\lambda})$

$$\hat{\psi} = diag(S - \hat{\lambda}\hat{\lambda}^T)$$

- principal components analysis method estimation is easy to understand, because you are just decomposing the matrix using spectral decomposition.
- In case of co principal factor method that ϕ value if it is not known then you cannot go for that method, but it should be initialized.
- The initialized point is that if you can take the diagonal element; assume that also possible take the value element of s initial value than you interpret this process.
- In maximum likelihood method that is told this is preferable one because there are lot of statistical tests, which are possible, but it is the derivation part is little difficult.

- Factor Rotation
- Factor Scaling
- Confirmatory FA
-

The no. of factors - to be retained. How many factors?

→ % cumulative variance explained

→ Eigen value criteria (>1) → Consider at least 1, that one variable's variability will be explained.

→ Scree plot

Set the criteria 90%.
if cum% > 90%, keep 3 factors.

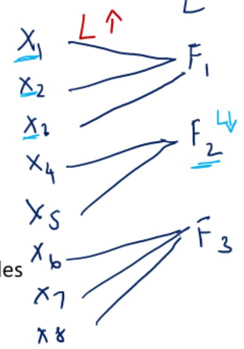
	F_1	F_2	F_3
X_1	λ_{11}	λ_{12}	λ_{13}
X_2	λ_{21}	λ_{22}	λ_{23}
$\vdots$	$\vdots$	$\vdots$	$\vdots$
X_p	λ_{p1}	λ_{p2}	λ_{p3}
	$\sum \lambda_{ij}^2$	$\sum \lambda_{2j}^2$	$\sum \lambda_{3j}^2$
	%	%	%
	cum%	cum%	cum%

Factor Analysis < Dimension reduction / Interpretability
[able to provide name to each of factors]

X_1 X_2 X_3 create F1

And so on..

Name the factors based on the nature of the variables



Exploratory FA: 1.PCA 2.PFM 3.MLM

If not possible then rotate the structure in such a manner that the loading of some of variables, on a particular factor will be maximum, whereas the loading of other variables will be minimum.
Similarly if you want to find some other set of x excluding the ones we have consider for the F1 that which will be highly loaded with F2 and very low loading with other factors. This is possible through Factor Rotation

Results

loadings.

✓ equally loaded
* Higher on F1 low on F2

Variables	Factor-1	Factor-2	Communa-lities
Gaelic	0.553 ✓	0.429 ✓	0.49
English	0.568 *	0.288 *	0.41
History	0.392	0.450	0.36
Arith-matic	0.740	-0.273	0.63
Algebra	0.724	-0.211	0.59
Geometry	0.595	-0.132	0.37
% variance explained	0.37	0.10	

not an issue

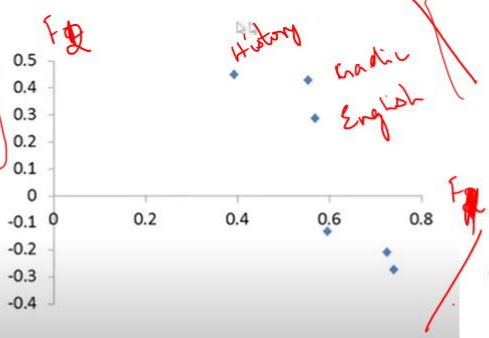
None of the loading is zero.
That means none of the variables are not loaded with two factors

Results (Lawley and Maxwell (1971))

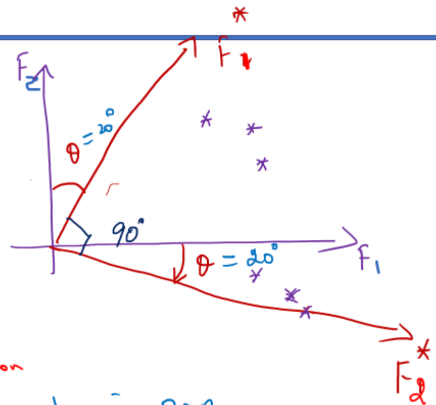
Cant say Highly loaded

Un-rotated

Variables	Factor-1	Factor-2
Gaelic	0.553	0.429
English	0.568	0.288
History	0.392	0.450
Arith-matic	0.740	-0.273
Algebra	0.724	-0.211
Geometry	0.595	-0.132



∴ rotate



if you know θ, then we will be able to find out F_1^* and F_2^* from F_1 & F_2.

Go for result rotation

Remember in PCA

$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$F_1^* = F_1\cos\theta - F_2\sin\theta$$

$$F_2^* = F_1\sin\theta + F_2\cos\theta.$$

$$TT^T = T^T T = TT^{-1} = I$$

This is orthogonal Transformation

Results (Lawley and Maxwell (1971))

Rotated

Variables	Factor-1	Factor-2
Gaelic	0.369	0.594
English	0.433	0.467
History	0.211	0.558
Arith-matic	0.789	0.001
Algebra	0.752	0.054
Geometry	0.604	0.083

verbal

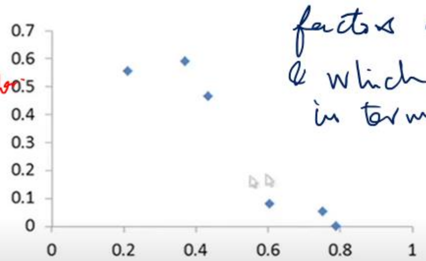
quantitative mathematical

low

higher

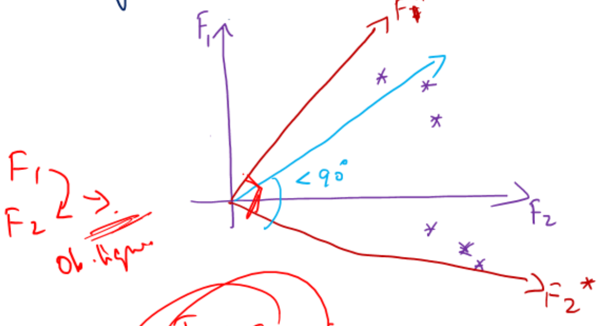
high

low



mathematical & verbal factors are hidden & which are manifested in terms of performance if the rotation is not orthogonal,

Oblique rotation. ie there is Correlation between factors



$$J J^T \neq I$$

Now we need to see under orthogonal rotation, are we losing anything related to explaining the variability of the original variables.

got $\Sigma = \lambda \lambda^T + \psi$

$x - \mu = \lambda F + \delta$ original model

$F^* = T F$

$x - \mu = \lambda F^* + \delta$

orthogonal transformation.

$\therefore \Sigma = \lambda \underbrace{T T^T}_{I} \lambda^T + \psi$

$\therefore \lambda \lambda^T + \psi$. No change in Covariance but only $\circledast$

$\circledast$ Loadings are changing. The variability explained will not be reduced. In case of interpretation it will be improved.

Types of transformations

- Orthogonal→mostly used Varimax, equimax, quartimax
- Oblique

  - Varimax – the variables will be highly loaded with only one factor

	F_1	F_2	...	F_m
x_1	✓	0	...	0
x_2	0	✓	...	0
⋮				
x_p	0	0	✓	0

Factor Scores- Can be used further for regression

i	X_1	X_2	...	X_p	Factor model	F_1	F_2	...	F_m
1	x_{11}	x_{12}	...	x_{1p}		f_{11}	f_{12}	...	f_{1m}
2									
⋮									
i						f_{i1}	f_{i2}	...	f_{im}
⋮									
n	x_{n1}	x_{n2}	...	x_{np}		f_{n1}	f_{n2}	...	f_{nm}

✓ ?? ?

$$x - \mu = \lambda F + \delta$$

is it linear equation?

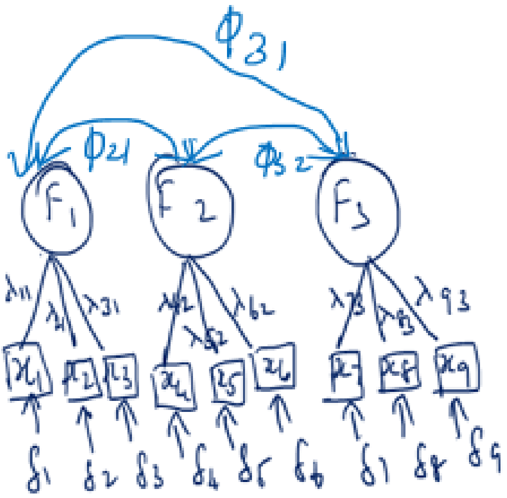
$$y = \beta x + \epsilon$$

✓ ? ✓

This equation is not similar to linear equation, where we know X and Y. But here we don't know lamda, F and delta.

To know scores,
1. Weighted least sq
2. Regression method
3. Maximum likelyhood method.

Confirmatory factor Analysis



	F1	F2	F3
$x_1-\mu_1 = \lambda_{11}F_1+\delta_1$			
$x_2-\mu_2 = \lambda_{21}F_1+\delta_2$			
$x_3-\mu_3 = \lambda_{31}F_1+\delta_3$			
$x_4-\mu_4 =$		$\lambda_{42}F_2+\delta_4$	
$x_5-\mu_5 =$		$=\lambda_{52}F_2+\delta_5$	
$x_6-\mu_6$		$=\lambda_{62}F_2+\delta_6$	
$x_7-\mu_7 =$			$\lambda_{73}F_3+\delta_7$
$x_8-\mu_8 =$			$=\lambda_{83}F_3+\delta_8$
$x_9-\mu_9 =$			$=\lambda_{93}F_3+\delta_9$

Matrix form:

$$\underset{9\times1}{\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ \vdots \\ x_9-\mu_9 \end{bmatrix}} = \underset{9\times3}{\begin{bmatrix} \lambda_{11} & 0 & 0 \\ \lambda_{21} & 0 & 0 \\ \lambda_{31} & 0 & 0 \\ 0 & \lambda_{42} & 0 \\ 0 & \lambda_{52} & 0 \\ 0 & \lambda_{62} & 0 \\ 0 & 0 & \lambda_{73} \\ 0 & 0 & \lambda_{83} \\ 0 & 0 & \lambda_{93} \end{bmatrix}} \underset{3\times1}{\begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix}} + \underset{9\times1}{\begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_9 \end{bmatrix}}$$

Exploratory Factor Analysis

$$x-\mu = \lambda F + \delta$$

$$\Sigma = E\left[(x-\mu)(x-\mu)^T\right]$$

$$EFA = \lambda\lambda^T + \psi$$

$$EFA: \lambda I \lambda^T + \psi$$

$$TT^T$$

Confirmatory factor Analysis

$$CFA: E(FF^T) = \phi$$

$$\therefore CFA = \lambda\phi\lambda^T + \psi$$

$$\boxed{\begin{aligned} EFA&: \lambda\lambda^T + \psi \\ CFA&: \lambda\phi\lambda^T + \psi \end{aligned}}$$

Many loading of CFA will be zero and oblique, that is allowing covariance between factors. So, in CFA, I is not TT^T, instead $E(FF^T)=\varphi$.

Problem 1:From the given covariance (S) , eigen values $\sqrt{\widehat{Q}_\iota}$and eigen vactors $\overrightarrow{\hat{e}_\iota}$, calculate error matrix(φ) for the number factor m=1.

$$\hat{Q}_1 = 148.23$$

$$S = \begin{bmatrix} 100 & 50 & -30 \\ 50 & 64 & -20 \\ -30 & -20 & 49 \end{bmatrix} \qquad \hat{Q}_2 = 36.00$$

$$\hat{Q}_3 = 28.7$$

$$\hat{e}_1 = \begin{bmatrix} -0.77 \\ -0.54 \\ 0.34 \end{bmatrix} \quad \hat{e}_2 = \begin{bmatrix} 0.21 \\ 0.29 \\ 0.93 \end{bmatrix} \quad \hat{e}_3 = \begin{bmatrix} 0.60 \\ -0.79 \\ 0.11 \end{bmatrix}$$

Solution:

Steps:

find:

1. $\lambda = \sqrt{Q}\,\vec{e}$
2. $\lambda \lambda^T$
3. $\Psi = S - \lambda\lambda^T$ 4. find Ψ

$$S = \begin{array}{c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \\ \begin{bmatrix} 100 & 50 & -30 \\ 50 & 64 & -20 \\ -30 & -20 & 49 \end{bmatrix} \end{array} \qquad \hat{Q}_1 = 148.23$$

$$\hat{Q}_2 = 36.00$$

$$\hat{Q}_3 = 28.7$$

m = 1

$$S = \hat{\lambda}\hat{\lambda}^T + \hat{\varphi}$$

$$\hat{\lambda}_{p \times m} = \hat{\lambda}_{3 \times 1} = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix}$$

$$\hat{e}_1 = \begin{bmatrix} -0.77 \\ -0.54 \\ 0.34 \end{bmatrix} \quad \hat{e}_2 = \begin{bmatrix} 0.21 \\ 0.29 \\ 0.93 \end{bmatrix} \quad \hat{e}_3 = \begin{bmatrix} 0.60 \\ -0.79 \\ 0.11 \end{bmatrix}$$

$$\lambda_1 = \begin{bmatrix} \lambda_{11} \\ \lambda_{21} \\ \lambda_{31} \end{bmatrix} \quad \lambda_2 = \begin{bmatrix} \lambda_{12} \\ \lambda_{22} \\ \lambda_{32} \end{bmatrix} \quad \lambda_3 = \begin{bmatrix} \lambda_{13} \\ \lambda_{23} \\ \lambda_{33} \end{bmatrix}$$

$m = 1.$

$$\hat{\lambda}_1 = \sqrt{Q_1}\, \hat{e}_1 = \sqrt{148.23} \begin{bmatrix} -0.77 \\ -0.54 \\ 0.34 \end{bmatrix} = \begin{bmatrix} -9.37 \\ -6.57 \\ 4.75 \end{bmatrix}$$

$$\hat{\lambda}\hat{\lambda}^T = \begin{bmatrix} -9.37 \\ -6.57 \\ 4.13 \end{bmatrix}_{3\times1} \begin{bmatrix} -9.37 & -6.57 & 4.13 \end{bmatrix}_{1\times3}$$

$$= \begin{bmatrix} 87.8 & 61.56 & -38.7 \\ 61.56 & 43.16 & -27.13 \\ -38.7 & -27.13 & 17.06 \end{bmatrix}_{3\times3}$$

$$S - \lambda\hat{\lambda}'$$

$$\Psi = \begin{bmatrix} 100 & 50 & -30 \\ 50 & 64 & -20 \\ -30 & -20 & 49 \end{bmatrix} - \begin{bmatrix} 87.8 & 61.56 & -38.7 \\ 61.56 & 43.16 & -27.13 \\ -38.7 & -27.13 & 17.06 \end{bmatrix} = \begin{bmatrix} 12.2 & -11.56 & 8.7 \\ -11.56 & 20.84 & 7.13 \\ 8.7 & 7.13 & 22.94 \end{bmatrix} \quad 0$$

$$\Psi = S - \lambda\lambda^T$$

Note: if m=1 is not mentioned in question, find 3 loadings and then calculate error matrix.

Problem 2: From the given Correlation matrix determine the factors.

Example 9.3 (Factor analysis of consumer-preference data) In a consumer-preference study, a random sample of customers were asked to rate several attributes of a new product. The responses, on a 7-point semantic differential scale, were tabulated and the attribute correlation matrix constructed. The correlation matrix is presented next:

Attribute (Variable)		1	2	3	4	5
Taste	1	1.00	.02	.96	.42	.01
Good buy for money	2	.02	1.00	.13	.71	.85
Flavor	3	.96	.13	1.00	.50	.11
Suitable for snack	4	.42	.71	.50	1.00	.79
Provides lots of energy	5	.01	.85	.11	.79	1.00

It is clear from the circled entries in the correlation matrix that variables 1 and 3 and variables 2 and 5 form groups. Variable 4 is "closer" to the (2, 5) group than the (1, 3) group. Given these results and the small number of variables, we might expect that the apparent linear relationships between the variables can be explained in terms of, at most, two or three common factors.

Therefore f1={1,3} and f2={2,4,5}.

Problem 3:

SASTRA DEEMED TO BE UNIVERSITY

3 a : given: eigenvalues & eigenvectors Q, $\vec{e}$

b : given: loadings λ

find Covariance matrix Σ .

additional Step for 3-a.

1. find loadings $\lambda = \sqrt{Q}\,\vec{e}$

both 3.a & 3.b. - Same Steps

3.a. instead of loadings if eigen values and eigenvectors are given calculate loadings using the formula

$$\lambda = \left[\sqrt{Q_1}\, e_1 \;\; \sqrt{Q_2}\, e_2 \;\ldots\; \sqrt{Q_m}\, e_m\right]$$

Else use the given loadings to find $\sum$.

Problem 3.b) Find covariance ∑ from the loadings given below.

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\hat{e}_{ij}$	
	F_1	F_2
1. Taste	.56	.82
2. Good buy for money	.78	−.53
3. Flavor	.65	.75
4. Suitable for snack	.94	−.10
5. Provides lots of energy	.80	−.54

Solution:

$$h_i^2 = \lambda_{i1}^2 + \lambda_{i2}^2 + \dots \lambda_{im}^2$$

$$\psi = 1 - h_i^2$$

SASTRA

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\hat{e}_{ij}$		Communality h_j^2	Specific variance $\psi = (1 - h_j^2)$
	F_1	F_2		
1. Taste	$(.56)^2$	$+(.82)^2$	0.99	0.01
2. Good buy for money	$+(.78)^2$	−.53	0.89	0.11
3. Flavor	$+(.65)^2$	.75	0.99	0.01
4. Suitable for snack	$+(.94)^2$	−.10	0.89	0.11
5. Provides lots of energy	$+(.80)^2$	−.54	0.93	0.07
	Q_1 2.87	Q_2 1.82		

$$\Sigma = \lambda\lambda^T + \psi$$

$$\Sigma = \lambda \lambda^T + \psi$$

λ_1	λ_2
.56	.82
.78	.53
.65	.75
.94	−.10
.80	−.54

$$\begin{bmatrix} .56 & .78 & .65 & .94 & .80 \\ .82 & -.53 & .75 & -.1 & -.54 \end{bmatrix} + \begin{bmatrix} 0.01 & & & & \\ & 0.11 & & & \\ & & 0.01 & & \\ & & & 0.11 & \\ & & & & 0.07 \end{bmatrix}$$

$$\Sigma_{corr} = \begin{bmatrix} 0.99 & 0 & 0.98 & 0.44 & 0.01 \\ & 0.89 & 0.11 & \boxed{0.79} & 0.91 \\ & & 0.99 & 0.54 & 0.12 \\ & & & 0.89 & \boxed{0.81} \\ & & & & 0.93 \end{bmatrix} + \begin{bmatrix} 0.01 & & & & \\ & 0.11 & & & \\ & & 0.01 & & \\ & & & 0.11 & \\ & & & & 0.07 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 1 & 0 & 0.98 & 0.44 & 0.01 \\ & 1 & 0.11 & 0.79 & 0.91 \\ & & 1 & 0.54 & 0.12 \\ & & & 1 & 0.81 \\ & & & & 1 \end{bmatrix}$$

Problem 3.c: Given either 3.a or 3.b find factors.

Normalized covariance matrix is correlation matrix. So as given in 3.b find ∑. And apply problem 2 concept to determine factors.

Problem 3.d)

3.D. given 3.a (or) 3.b + covariance (S) question find residual matrix

get $\Sigma = \lambda \lambda^T + \psi$ $\quad S - \Sigma$

Loading are not given start from first step, else start from second step.

1. find $\lambda = \sqrt{a} \cdot \vec{e}$
2. find $\lambda\lambda^T$
3. find $\Sigma \text{Row}(\lambda^2) \rightarrow h_j^2$
4. find $\Sigma \text{col}(\lambda^2) \rightarrow a$
5. Specific $\varphi = 1 - h_j^2$
6. $\Sigma = \lambda\lambda^T + \varphi$
7. $S - \Sigma$
8. Inferance.

(continuing the previous sum for finding residual matrix)

After 6th step,

Step 7:

S

Attribute (Variable)		1	2	3	4	5
Taste	1	1.00	.02	.96	.42	.01
Good buy for money	2	.02	1.00	.13	.71	.85
Flavor	3	.96	.13	1.00	.50	.11
Suitable for snack	4	.42	.71	.50	1.00	.79
Provides lots of energy	5	.01	.85	.11	.79	1.00

Σ

$$\begin{bmatrix} 1.00 & .01 & .97 & .44 & .00 \\ & 1.00 & .11 & .79 & .91 \\ & & 1.00 & .53 & .11 \\ & & & 1.00 & .81 \\ & & & & 1.00 \end{bmatrix}$$

$$\text{Res} = S - \Sigma = \begin{bmatrix} 0 & 0.01 & -0.01 & -0.02 & 0.01 \\ & 0 & 0.02 & -0.08 & -0.06 \\ & & 0 & -0.03 & 0 \\ & & & 0 & 0.02 \\ & & & & 0 \end{bmatrix} \cong 0$$

The predicted covariance nearly produces the sample covariance. So using the factor a good model fit can be obtained.

Problem 3.E) : Extension of 3.a, to 3.D. or separate problem.

From the given loadings find the PVE of factors.

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\hat{e}_{ij}$ F_1	F_2
1. Taste	$(.56)^2$	$+(.82)^2$
2. Good buy for money	$+(.78)^2$	$-.53$
3. Flavor	$+(.65)^2$	$.75$
4. Suitable for snack	$+(.94)^2$	$-.10$
5. Provides lots of energy	$+(.80)^2$	$-.54$

Q_1 $\quad$ Q_2

2.87 $\quad$ 1.82

$$2.87 \qquad PVE(Q_1) = \frac{2.87}{4.69} \times 100 = 61.2$$

$$1.82$$

$$4.69 \qquad PVE(Q_2) = \frac{1.82}{4.69} \times 100 = 38.8$$

Since first factor explain only 61.2% of variance, we need second factor also since the cumulative factor provides 100% variance.

Problem 3.F : Given 3.a or 3.b.

Find PVE, commonality and specific variance from the given loadings. And give your inference.

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\hat{e}_{ij}$ F_1	F_2
1. Taste	.56	.82
2. Good buy for money	.78	−.53
3. Flavor	.65	.75
4. Suitable for snack	.94	−.10
5. Provides lots of energy	.80	−.54

Solution:

$\psi = (1 - h_j^2)$

Variable	Estimated factor loadings $\tilde{\ell}_{ij} = \sqrt{\hat{\lambda}_i}\hat{e}_{ij}$ F_1	F_2	Commanality h_j^2	Specific variance
1. Taste	$(.56)^2$	$+(.82)^2$	0.99	0.01
2. Good buy for money	$+(.78)^2$	−.53	0.89	0.11
3. Flavor	$+(.65)^2$	.75	0.99	0.01
4. Suitable for snack	$+(.94)^2$	−.10	0.89	0.11
5. Provides lots of energy	$+(.80)^2$	−.54	0.93	0.07

Q_1	Q_2
2.87	1.82

$$\Sigma = \lambda\lambda^T + \psi$$

2.87
1.82
4.69

$$PVE(Q_1) = \frac{2.87}{4.69} \times 100 = 61.2$$

$$PVE(Q_2) = \frac{1.82}{4.69} \times 100 = 38.8$$

Inference:

1. Since first factor explain only 61.2% of variance, we need second factor also since the cumulative factor provides 100% variance.

2. Since all the commonality values are near to 1, all the variables are providing good variance and hence we can create good model using these variables.

3. The above inference also be proved through Specific variance, which is near to zero, so unexplained variance by the variables are very less.