Factor Analysis (FA)

FA is a multivariate method used for data reduction purposes. The basic idea is to represent a set of variables by a smaller number of variables. & these variables are called / known as factors.

Technical Definition.

FA is a method for modelling observed variables & their covariance structure, in terms of a smaller number of underlying unobservable (latent) "factors".

What is the aim or purpose of FA in form simple words.

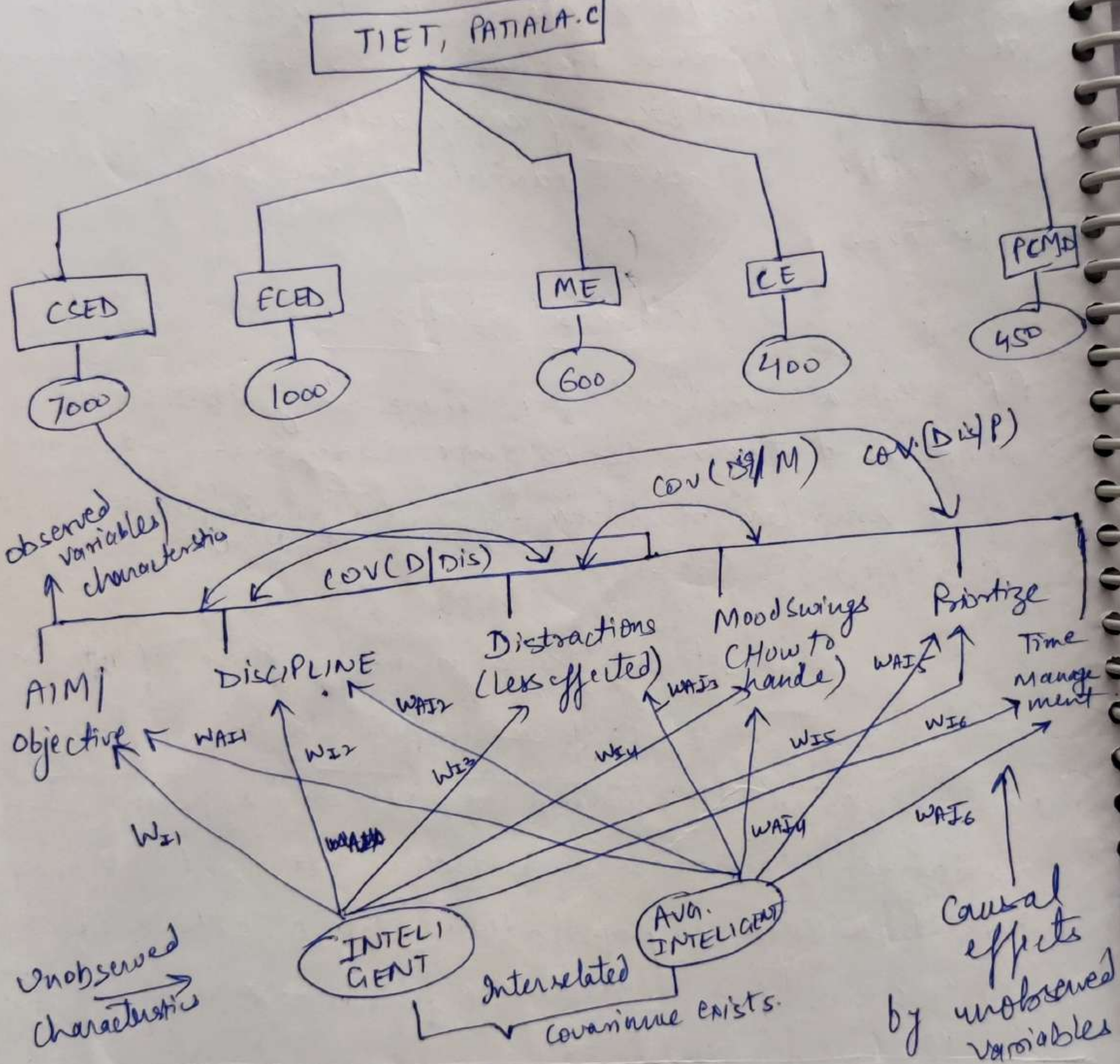
i.e To estimate a model which explains variance / covariance between a set of observed variables (in a population) by a set of (fewer or less no. of) unobserved factors & weightings.

Some of the Examples.

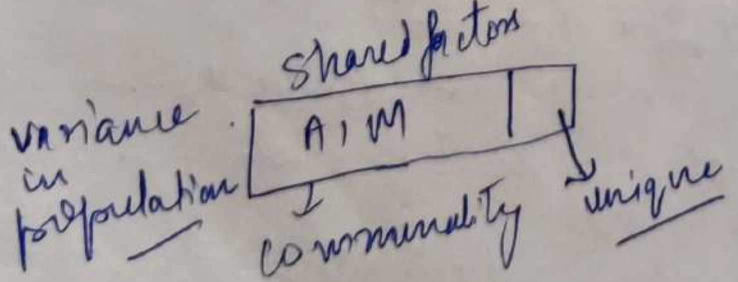
(1) A welfare officer is interested to measure the quality of work life in factories.

(2) A psychologist is interested to measure mental ability of a person.

Raw Example



TIET, PATIALA.C
CSED
7000
ECED
1000
ME
600
CE
400
PCMD
450
observed variables/ characteristics
COV(D|Dis)
COV(Dis/M)
COV(Dis/P)
AIM/ objective
DISCIPLINE
Distractions (Less effected)
Moodswings (How to hande)
Priortize
Time manage ment
WAI1
WAI2
WAI3
WAI4
WAI5
WAI6
WI1
WI2
WI3
WI4
WI5
WI6
INTELI GENT
AVG. INTELIGENT
Unobserved characteristic
Interrelated
covariance exists.
Causal effects by unobserved variables

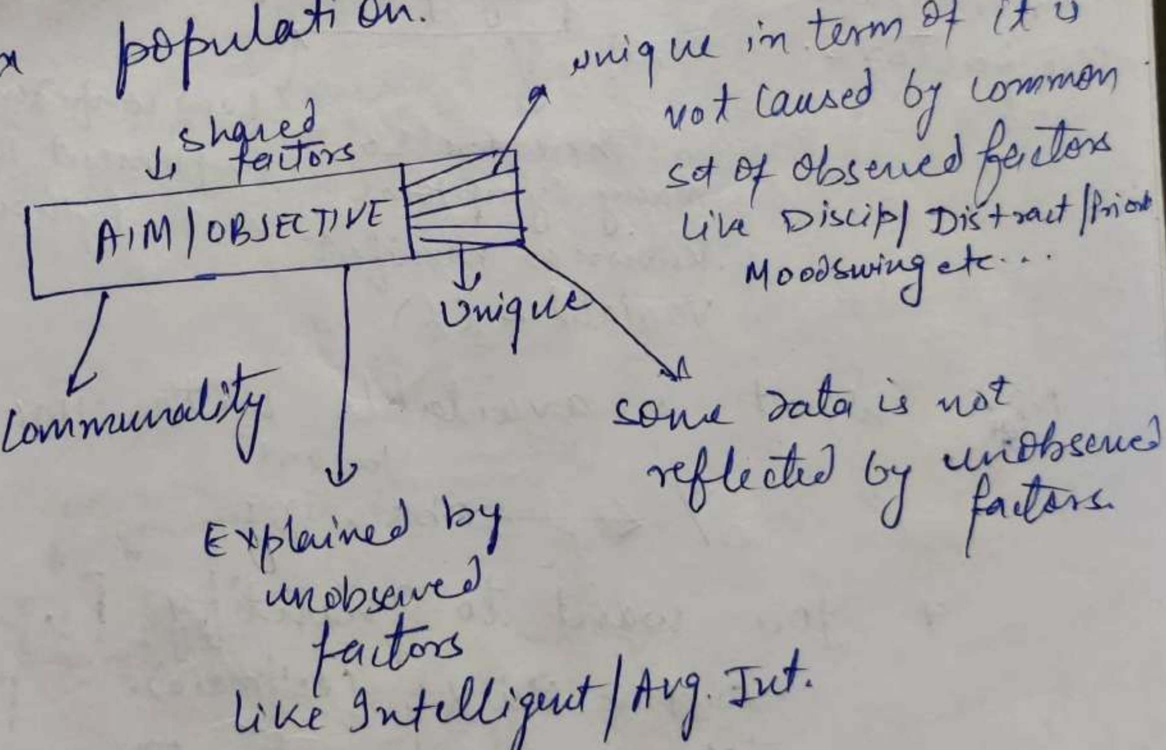


Shared factors
AIM
variance in population
communality
unique

What we conclude from above Example.

· Our Model will explain whether ∃ a variance/ covariance exist b/w observed variables/ characteristics.

for we have to check about the variance of characteristics as an example AIM/OBJECTIVE in a population.



AIM
Discipline
Distractions
Moodswings
Priortize
Time Manag.

unobservable
concepts
constructs

Technical Analysis of Factor Analysis

overthinking habbit → some one habbit is there to overthink rather than do a practical work

constructs / factors → (F)

Causing these manifest
[illegible]

from these symptoms we can judge or measure the factors.

$\boxed{\text{Symptoms}}$

There are so many symptoms known as Manifest variables (X)

Low confidence in implement the concepts in R, Python or MatLab.

Now, what is available with you??

$X \leftarrow$ You are observing

& You want to quantify "F". (estimate).

suppose this X have

$$X_{p\times1} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{bmatrix}_{p\times1} \qquad \mu_{p\times1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}_{p\times1}$$

This is known about the population

Values may or may not be known

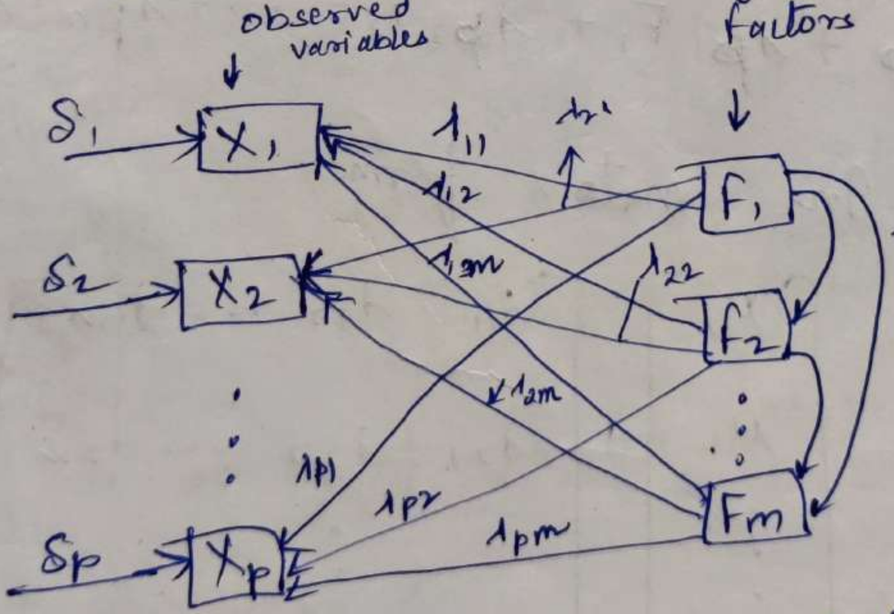
$$\text{cov}(X) = \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1p} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2p} \\ \\ \sigma_{p1} & \sigma_{p2} & \cdots & \sigma_{pp} \end{bmatrix}$$

consider 'm' factors, then

$$F_{m \times 1} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ \vdots \\ F_m \end{bmatrix}$$

"F" is causing X {symptoms}

So; What you observe in X, it is just because of causal factors i.e $F_{m \times 1}$ from that conclusion, we are able to write the linear regression kind of eqn. But first see this diagram, which helps us to write linear regression Eqn.



X_1 is caused by F_1

→ we are not assuming any relationship b/w $F_1, F_2, \ldots, F_m$ i.e they all are orthogonal.

These λ_{pm} are infulencing factors / coefficients / weighting coefficients.

Now, these factors $(F_1, F_2, \dots F_m)$ are not able to explain each & every thing about X_1 and so on. Thus, $\exists$ an error term associated with each $X_1, X_2, \dots, X_p$ variables.

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

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

$$\vdots$$

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

$$\vdots$$

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

In matrix form.

$$\begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}_{p\times 1} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{bmatrix}_{p\times 1} + \begin{bmatrix} \lambda_{11} & \lambda_{12} & \dots & \lambda_{1m} \\ \lambda_{21} & \lambda_{22} & \dots & \lambda_{2m} \\ & \vdots & & \\ \lambda_{p1} & \lambda_{p2} & \dots & \lambda_{pm} \end{bmatrix}_{p\times m} \begin{bmatrix} F_1 \\ F_2 \\ \vdots \\ F_m \end{bmatrix}_{m\times 1} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_p \end{bmatrix}_{p\times 1}$$

$$\therefore \underset{p\times1}{X} = \underset{p\times1}{\mu} + \underset{p\times m}{\Lambda}\underset{m\times1}{F} + \underset{p\times1}{\delta}$$

$$\Rightarrow \underbrace{X-\mu = \Lambda F}{} + \delta \quad \text{——— (1)}$$

Eqn. known as factor Model.

One important assumption is there

i.e f_1 f_2 ... f_m

$\underbrace{\qquad\qquad\qquad}$ "No" relationship i.e orthogonal

Usual Assumpt

$E(X) = \mu$

$Cov(X) = \Sigma$

$E(F) = 0$

$E(\delta) = 0$

$Cov(F) = E(FF^T) = I$

$Cov(F\delta) = 0$ $Cov(\delta) = \Psi$

↓

Diagonal matrices

$$\begin{bmatrix} \psi_{11} & 0 & \cdots & 0 \\ 0 & \psi_{22} & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \psi_{pp} \end{bmatrix}$$

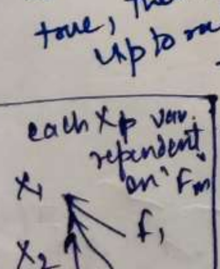
But, if f_1 f_2 ... f_m (ϕ_{21}, ϕ_{m1})

when there is correlation exist b/w them

* If these all assumptions are hold true, then FA is called exploratory/orthogonal

then above factor model becomes "Oblique" factor model.

each X_p var. dependent on F_m



Types of factor Model

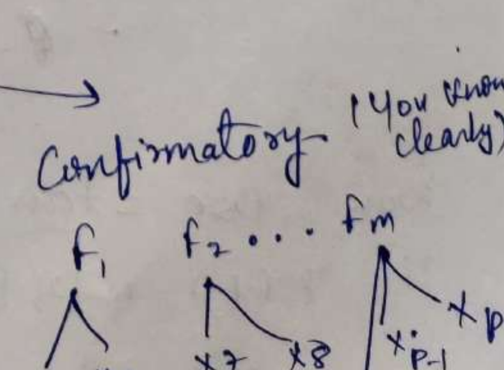
Exploratory

You are exploring the possibilities of hidden factors "F".

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

↑ Not clearly that's why each & every thing connected.

Confirmatory (You know clearly)



from above discussion it is cleared that 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.

i.e $\quad Cov(X) = \Sigma_{p \times p}$ $\quad$ we know that assumptions.

Now $\underbrace{X - \mu = \Lambda F + \delta}$ ← factor Model.

$\uparrow (X)$.

I want to know that what kind of variability you are able to explain by above model.

$$\Sigma = Cov(X) = E\{(X-\mu)(X-\mu)^T\}$$

$$= E\{(\Lambda \underline{F}+\delta)(\Lambda \underline{F}+\delta)^T\}$$

$$= E\{\Lambda F \cdot F^T \Lambda^T + \Lambda F \delta^T + \delta \cdot F^T \cdot \Lambda^T + \delta \delta^T\}$$

$$= \Lambda E(F \cdot F^T) \cdot \Lambda^T + \Lambda E(\underline{F\delta^T}) + E(\delta \cdot F^T) \cdot \Lambda^T + E(\delta \cdot \delta^T)$$

Now, Use standard assumption, Proof by yourself

(i) $E(F) = 0$ (2) $E(\delta) = 0$, (3) $Cov(F) = E(F F^T) = I$

(4) $Cov(\delta) = \Psi$ (5) $Cov(F\delta) = 0$

or $E(\delta\delta^T) = \Psi$

$$= \Lambda I \Lambda^T + \Lambda \cdot 0 + 0 \cdot \Lambda^T + \Psi$$

$$\Sigma = \Lambda\Lambda^T + \Psi = cov(X) \quad \text{---} \quad (2)$$

↑ what is benift of that ??

i.e If I know Λ, Ψ then, I can easily compute variability Σ / cov(X).

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

↓ Loading matrix

In General case:

$$X_j = \lambda_{j1} f_1 + \lambda_{j2} f_2 + \cdots + \lambda_{jk} f_k + \cdots + \lambda_{jm} f_m + \delta_j$$

Loading

λ_{jk} = loading of k^{th} factor on j^{th} X

$j = 1, 2, \cdots p$

$k = 1, 2, \cdots, m$

then $\Lambda \cdot \Lambda^T = \Lambda_{p \times m} \cdot \Lambda^T_{m \times p}$

↙ Symmetric matrix.

cov.

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

↓ var.

So, our data is partioned into two parts

X_j : $\boxed{\underbrace{\lambda \cdot \lambda^T}_{\text{contributions by "m" common factors}} \;\big|\; \Psi}$ — unexplained or Hidden things

Known as Communality (h^2) ∴ $k = 1, 2, \ldots m$

our aim is to optimize this.

Known as Unique or specific variance of X_j

Next topic is to estimate (λ).