Statistics for the Sciences

Factor Analysis

Xuemao Zhang
East Stroudsburg University

January 18, 2025

Outline

- Purpose of factor analysis
- Factor analysis
  - Model
  - Estimations
  - Factor rotation
- Example
- Lab
- Appendix: Estimating Factor Loadings

Purpose of factor analysis

- Factor Analysis (FA) assumes the covariation structure among a set of variables $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$ can be described via a linear combination of unobservable (**latent**) variables, $\mathbf{F} = (f_1, f_2, \ldots, f_m)'$ $(m < p)$,called **factors**.
  - **F** cannot be observed or measured.
  - Factor analysis is more controversial than other analytic methods because it leaves room for subjectivity and judgment.
  - But it can provide insights into the nature of abstract constructs and allows us to superimpose order on complex phenomena

Purpose of factor analysis

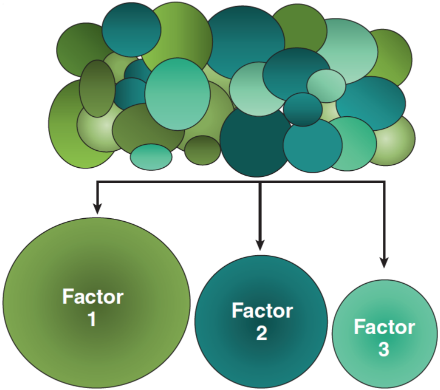


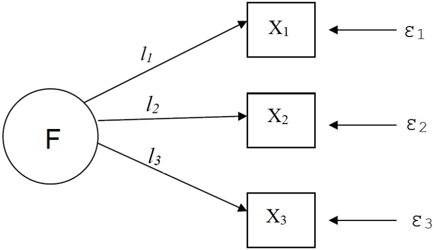
Figure 31–1 Conceptual representation of exploratory factor analysis, showing how multiple "green" items are grouped into three independent factors that explain the structure in a latent trait of "greenness."

Purpose of factor analysis

- FA and PCA have similar themes, i.e., to explain covariation between variables via linear combinations of other variables.
- However, there are distinctions between the two approaches:
  - FA assumes a statistical model that describes covariation in observed variables via **linear combinations of latent variables**.
  - PCA finds uncorrelated **linear combinations of observed variables** that explain maximal variance (no latent variables here).
- FA refers to a statistical model, whereas PCA refers to the eigenvalue decomposition of a covariance (or correlation) matrix.

Factor analysis - Model

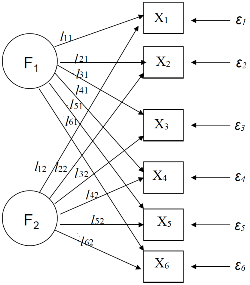
- **Illustration**: One Common Factor Model
  - (X_1, X_2, X_3) are observed, we try to find F



$$
\begin{aligned}
X_1 &= \mu_1 + l_1 F + \varepsilon_1 \\
X_2 &= \mu_2 + l_2 F + \varepsilon_2 \\
X_3 &= \mu_3 + l_3 F + \varepsilon_3
\end{aligned}
$$

Factor analysis - Model

- **Illustration**: Two-Common Factor Model
  - $(X_1, X_2, X_3, X_4, X_5, X_6)$ are observed, we try to find F_1 and F_2.



$$
\begin{aligned}
X_1 &= \mu_1 + l_{11}F_1 + l_{12}F_2 + \varepsilon_1 \\
X_2 &= \mu_2 + l_{21}F_1 + l_{22}F_2 + \varepsilon_2 \\
X_3 &= \mu_3 + l_{31}F_1 + l_{32}F_2 + \varepsilon_3 \\
X_4 &= \mu_4 + l_{41}F_1 + l_{42}F_2 + \varepsilon_4 \\
X_5 &= \mu_5 + l_{51}F_1 + l_{52}F_2 + \varepsilon_5 \\
X_6 &= \mu_6 + l_{61}F_1 + l_{62}F_2 + \varepsilon_6
\end{aligned}
$$

Factor analysis - Model

- $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$ is a random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.
- Statistical **Model:**

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{L}F + \boldsymbol{\varepsilon} \tag{1}$$

  - $\boldsymbol{L}$ denotes the **pattern matrix** of **factor loadings**
  - **F** denotes the vector of latent **factor scores**. It is a **random** vector variables.
    - They are assumed to be latent factors with mean zero, unit variance, and uncorrelated in Orthogonal Factor Model.
  - $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_p)'$ denotes the vector of latent *error terms*

$$\boldsymbol{L} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{bmatrix}$$

  - **F** and $\boldsymbol{\varepsilon}$ are independent

Factor analysis - Model

- The model is like a series of multiple regressions, predicting each of the observable variables X_i from the values of the unobservable common factors f_i, $i = 1, 2, \ldots, m$:

$$\begin{aligned} X_1 =& \mu_1 + l_{11} f_1 + l_{12} f_2 + \cdots + l_{1m} f_m + \varepsilon_1 \\ X_2 =& \mu_2 + l_{21} f_1 + l_{22} f_2 + \cdots + l_{2m} f_m + \varepsilon_2 \\ &\vdots \\ X_p =& \mu_p + l_{p1} f_1 + l_{p2} f_2 + \cdots + l_{pm} f_m + \varepsilon_p \end{aligned}$$

- We also prefer dimension reduction, so generally we require $m << p$.

Factor analysis - Estimations

- Estimations of the factor loadings $\boldsymbol{L}$ ($m \times p$ parameters) are based on variance-covariance matrix $\boldsymbol{\Sigma}$ of $\boldsymbol{X}$, which is usually estimated by the sample covariance matrix of $\boldsymbol{X}$.
- Methods to estimate $\boldsymbol{L}$
  - PCA method (SPSS)
  - Iterated Principal Axis Factoring Method
  - MLE Method
  - Least Squares
  - Weighted Least Squares
  - Generalized Least Squares
  - Alpha factor analysis
  - Image Factoring
- See Appendix for details of some estimation methods

Criticisms of Factor Analysis

- Labels of factors can be arbitrary or lack scientific basis
- Derived factors often very obvious
  - defense: but we get a quantification
- Correlation matrix is often poor measure of association of input variables.
- "Garbage in, garbage out"
  - a criticism of input variables
  - It is a good idea to check the scatter plot matrix of the variates before conducting FA.

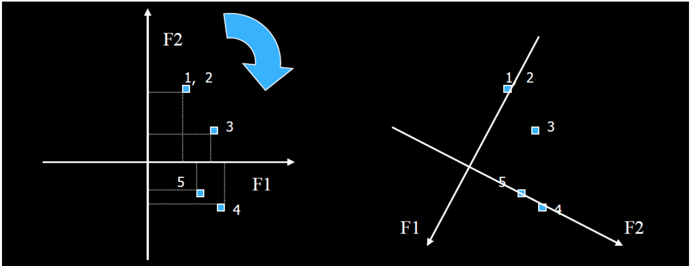
Factor analysis - Estimations

- However, the factor scores $\boldsymbol{F}$ is a random variable, so estimating realizations of $\boldsymbol{F}$ is different from estimating the $\boldsymbol{L}$:
  - Regression Method (SPSS)
  - Bartlett Method (SPSS)
  - Anderson-Rubin Method (SPSS)
  - Thompson's Method
  - Weighted Least Squares
  - Principal Component Method
  - Bayesian Methods
  - Nonlinear Methods

Factor analysis - Factor Rotation

- Goal: simple structure
  - Make factors more easily interpretable
    - While keeping the number of factors and communalities of Xs fixed!
  - **Rotation does NOT improve fit**!
- We "rotate" factors:
  - redefine factors such that `loadings` (or pattern matrix coefficients) on various factors tend to be very high (-1 or 1) or very low (0)
  - intuitively, it makes sharper distinctions in the meanings of the factors

Factor analysis - Factor Rotation



	Factor 1	Factor 2
x1	0.4	0.69
x2	0.4	0.69
x3	0.65	0.32
x4	0.69	-0.4
x5	0.61	-0.35

	Factor 1	Factor 2
x1	-0.8	0
x2	-0.8	0
x3	-0.6	0.4
x4	0	0.8
x5	0	0.7

Factor analysis - Factor Rotation

Orthogonal vs. Oblique Rotation

- Orthogonal: Factors are independent
  - **varimax**: maximize variance of squared loadings *across variables* (sum over factors)
    - Goal: the simplicity of interpretation of factors
  - **quartimax**: maximize variance of squared loadings *across factors* (sum over variables)
    - Goal: the simplicity of interpretation of variables
  - Intuition: from previous picture, there is a **right angle** between axes
- The factor analysis solution is NOT unique

Factor analysis - Factor Rotation

Orthogonal vs. Oblique Rotation

- Oblique: Factors are NOT independent. Change in “angle”.
  - **oblimin**: minimize covariance of squared loadings between factors.
  - **promax**: simplify orthogonal rotation by making small loadings even closer to zero.
  - **Target matrix**: choose “simple structure” a priori.
  - Intuition: from previous picture, angle between axes is not necessarily a right angle.

- The factor analysis solution is NOT unique

Factor Rotation: Which to use? Choice is generally not critical

- Interpretation with orthogonal (varimax) is “simple” because factors are independent: “Loadings” are correlations.
- Configuration may appear more simple in oblique (promax), but correlation of factors can be difficult to reconcile.

Example

- Consider the example for PCA (`wu.csv`): Wu et al. (2021) took soil samples at 300 sites near the Three Gorges Reservoir in China. Each site was classified into one of three categories of land use: orchard ($n = 75$), dry land ($n = 98$), and paddy field ($n = 127$). They also measured the concentrations of ten metals (Ca, Cd, Cr, Cu, Fe, Mg, Mn, Ni, Pb, and Zn; mg/kg) and five soil characteristics (pH, concentrations of N and P in mg/kg, and % of soil organic carbon [SOC] and K) for each site.
- After removing the last two columns

```
## 'data.frame':    300 obs. of  15 variables:
##  $ ca : num  0.81 0.466 0.932 0.81 0.322 ...
##  $ cr : num  66.2 43.8 70 60.6 46.5 ...
##  $ cu : num  20.39 10.29 25.59 19.9 7.44 ...
##  $ fe : num  3.47 2.04 3.78 3.54 2.09 ...
##  $ mn : num  633 229 569 546 239 ...
##  $ pb : num  21.5 22.4 25.6 25.3 20.9 ...
##  $ zn : num  68.8 36.5 78.1 67.5 28.9 ...
##  $ cd : num  0.175 0.134 0.227 0.175 0.096 0.185 0.35 0.258 0.134 0.34 ...
##  $ mg : num  1.4 0.482 1.59 1.232 0.403 ...
##  $ ni : num  25.9 13.3 31.3 26.2 12.2 ...
##  $ soc: num  0.59 0.84 0.98 0.977 0.63 0.582 0.87 1.17 0.86 0.874 ...
##  $ ph : num  6.82 6.17 6.13 6.33 5.41 6.92 6.82 7.34 5.86 6.67 ...
##  $ n  : num  711 745 1123 1026 648 ...
##  $ k  : num  2.12 1.52 2.61 2.16 1.38 ...
##  $ p  : num  597 305 702 482 374 ...
```

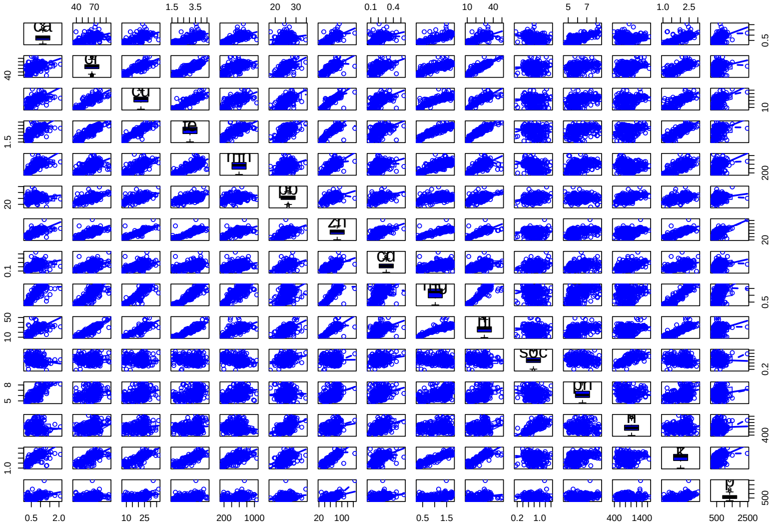
Example

- Correlation Matrix

```
##         ca    cr     cu    fe    mn    pb    zn   cd     mg    ni    soc     ph
## ca   1.000 0.503  0.670  0.69  0.72 0.200 0.737 0.41  0.809  0.68 -0.242  0.686
## cr   0.503 1.000  0.843  0.87  0.59 0.646 0.763 0.48  0.787  0.90  0.024  0.306
## cu   0.670 0.843  1.000  0.90  0.72 0.626 0.893 0.52  0.882  0.89 -0.057  0.346
## fe   0.691 0.869  0.904  1.00  0.78 0.631 0.878 0.46  0.918  0.93 -0.111  0.347
## mn   0.722 0.590  0.724  0.78  1.00 0.338 0.727 0.37  0.794  0.72 -0.340  0.383
## pb   0.200 0.646  0.626  0.63  0.34 1.000 0.589 0.49  0.470  0.61  0.311  0.024
## zn   0.737 0.763  0.893  0.88  0.73 0.589 1.000 0.62  0.887  0.85  0.013  0.393
## cd   0.414 0.483  0.515  0.46  0.37 0.487 0.624 1.00  0.489  0.51  0.246  0.324
## mg   0.809 0.787  0.882  0.92  0.79 0.470 0.887 0.49  1.000  0.92 -0.174  0.477
## ni   0.677 0.897  0.888  0.93  0.72 0.608 0.852 0.51  0.917  1.00 -0.103  0.434
## soc -0.242 0.024 -0.057 -0.11 -0.34 0.311 0.013 0.25 -0.174 -0.10  1.000 -0.098
## ph   0.686 0.306  0.346  0.35  0.38 0.024 0.393 0.32  0.477  0.43 -0.098  1.000
## n   -0.074 0.250  0.175  0.15 -0.18 0.500 0.242 0.43  0.075  0.17  0.722 -0.041
## k    0.712 0.735  0.813  0.84  0.73 0.494 0.841 0.48  0.919  0.87 -0.131  0.449
## p    0.473 0.252  0.466  0.36  0.40 0.202 0.559 0.30  0.425  0.32 -0.042  0.148
##          n     k     p
## ca  -0.074 0.712 0.473
## cr   0.250 0.735 0.252
## cu   0.175 0.813 0.466
## fe   0.151 0.844 0.363
## mn  -0.184 0.726 0.398
## pb   0.500 0.494 0.202
```

Example

- Scatter plot matrix



Example

- Test **significance of the correlations**. If this test is not statistically significant, you should not employ a factor analysis
- A popular test method is Bartlett test of the correlations

```
## $chisq
## [1] 5543
##
## $p.value
## [1] 0
##
## $df
## [1] 105
```

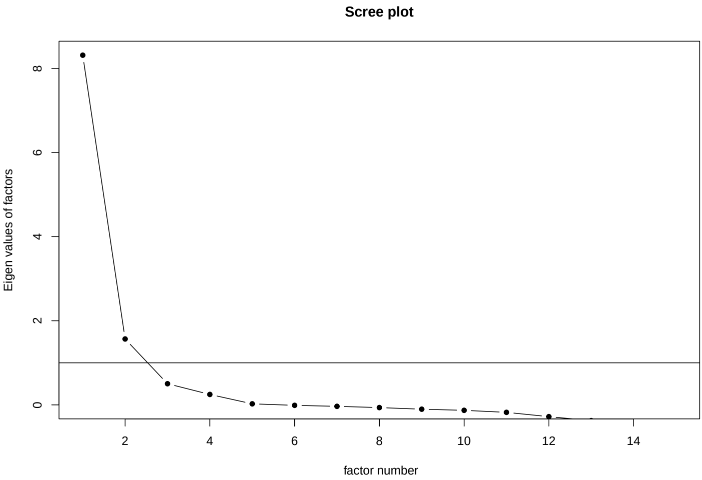
Example

- **KMO**: The Kaiser-Meyer-Olkin (KMO) tests to see if the partial correlations within your data are close enough to zero to suggest that there is at least one latent factor underlying your variables.
  - The minimum acceptable value is 0.6 before undertaking a factor analysis

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = wu)
## Overall MSA =  0.9
## MSA for each item =
##   ca   cr   cu   fe   mn   pb   zn   cd   mg   ni  soc   ph    n    k    p
## 0.87 0.91 0.96 0.91 0.95 0.88 0.93 0.89 0.90 0.93 0.68 0.76 0.69 0.93 0.84
```

Example

- How many factors do we need?



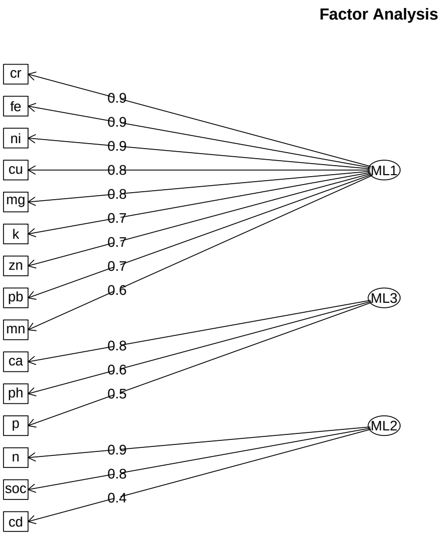
Example

- Factor analysis using ML method
- The Loading Matrix

```
##
## Loadings:
##     ML1   ML3   ML2
## cr   0.90  0.18  0.14
## cu   0.83  0.42  0.09
## fe   0.89  0.39  0.02
## mn   0.63  0.52 -0.27
## pb   0.68 -0.01  0.46
## zn   0.72  0.58  0.18
## mg   0.77  0.59 -0.05
## ni   0.88  0.39  0.04
## k    0.73  0.54 -0.01
## ca   0.41  0.82 -0.14
## ph   0.14  0.63 -0.07
## p    0.20  0.50  0.07
## soc -0.10 -0.10  0.83
## n    0.15 -0.02  0.87
## cd   0.37  0.41  0.44
##
##                ML1  ML3  ML2
## SS loadings    6.0 3.28 2.02
## Proportion Var 0.4 0.22 0.13
## Cumulative Var 0.4 0.62 0.75
```

Example

- Graph the Loading Matrix
  - Note that it only tells us the largest loading for each item.



Example

- **Communalities**: The communality for each variable is the percentage of variance that can be explained by the retained factors. It's best if the retained factors explain more of the variance in each variable.

```
##        .
## ca  0.86
## cr  0.86
## cu  0.88
## fe  0.94
## mn  0.73
## pb  0.68
## zn  0.89
## cd  0.49
## mg  0.95
## ni  0.93
## soc 0.70
## ph  0.42
## n   0.78
## k   0.83
## p   0.30
```

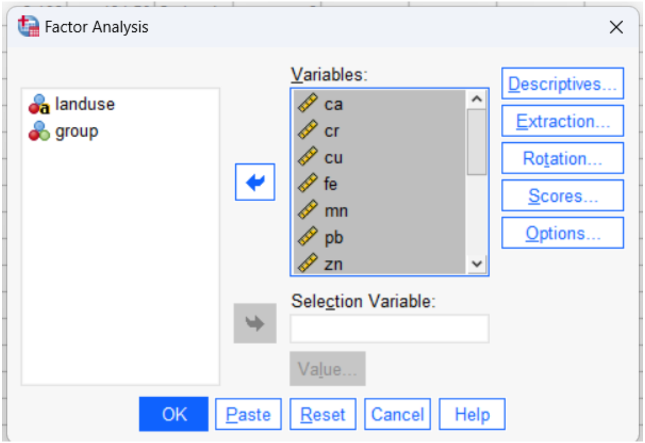
Summary

Factor analysis, despite being a mathematically intensive procedure, is highly subjective

- the type of extraction method to use (principal axis or principal components)
- the number of factors to extract
- the factor rotation method to use when looking for simple structure (and interpretability)
- the interpretation of the factors
- It is simply a method of **parsing large correlation matrices**. If you select several variables that are highly related, you should not be surprised if they group together to form a single factor.

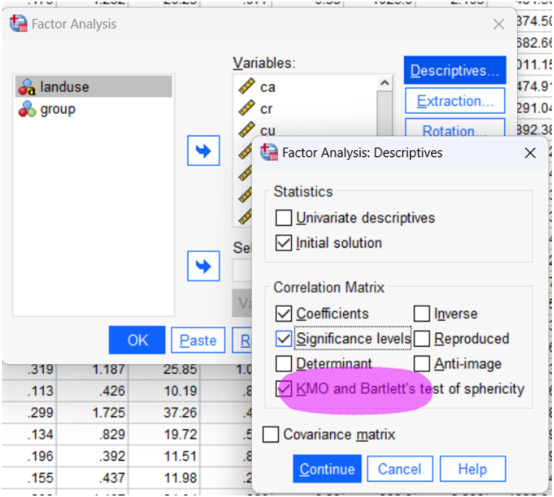
Lab

- After importing data, Click on `Analyze` → `Dimension Reduction` → `Factor`, move variables into the `Variables` box.



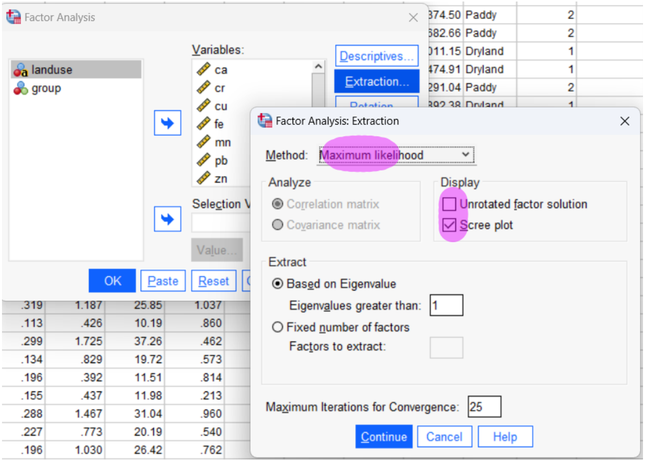
Lab

- Click on `Descriptives`,



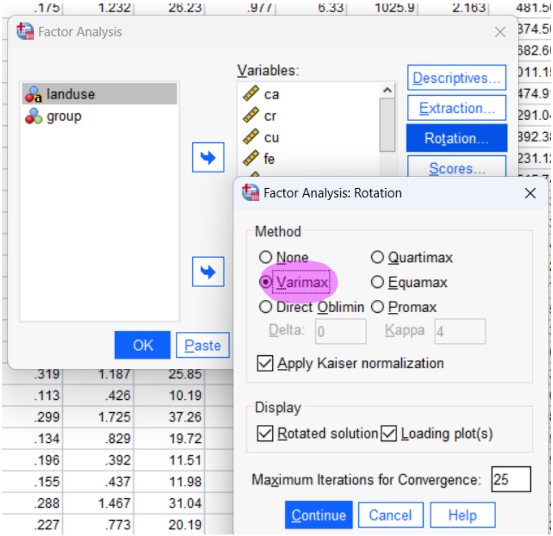
Lab

- Click `Extraction...` and let's select `Maximum Likelihood`, with `Scree plot` checked, and Eigenvalues greater than: `1`.



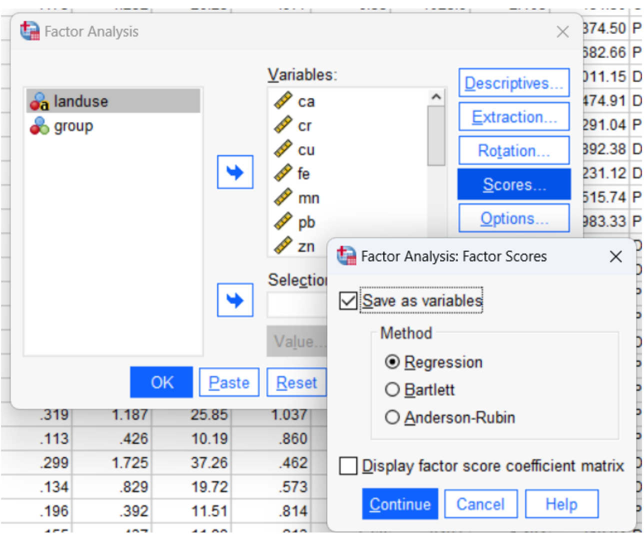
Lab

- Click on `Rotation...` and select `Varimax`



Lab

- Click on Scores...



Lab

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.901
Bartlett's Test of Sphericity	Approx. Chi-Square	5542.933
	df	105
	Sig.	<.001

Lab

Communalities

	Initial
ca	.848
cr	.867
cu	.889
fe	.942
mn	.762
pb	.699
zn	.912
cd	.545
mg	.955
ni	.937
soc	.613
ph	.608
n	.675
k	.866
p	.483

Extraction Method: Maximum Likelihood.

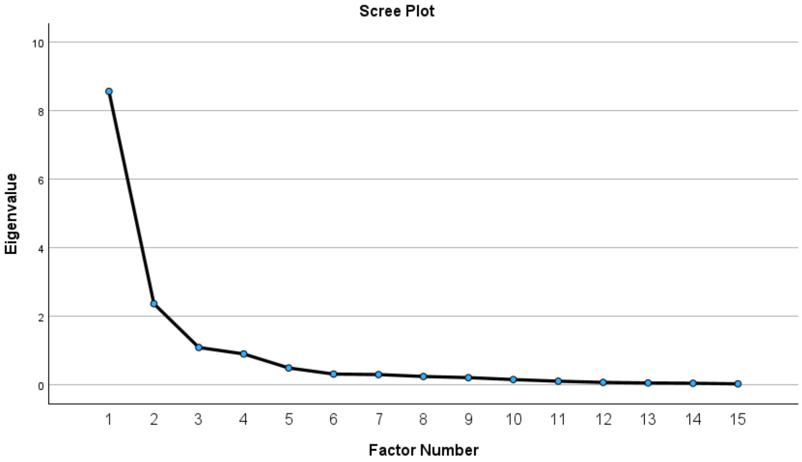
Lab

Total Variance Explained

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Factor	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	8.569	57.125	57.125	5.957	39.714	39.714
2	2.372	15.812	72.936	3.279	21.863	61.577
3	1.094	7.290	80.226	2.017	13.445	75.022
4	.906	6.037	86.264			
5	.496	3.308	89.572			
6	.317	2.116	91.688			
7	.303	2.019	93.707			
8	.247	1.647	95.354			
9	.214	1.427	96.782			
10	.158	1.051	97.833			
11	.110	.732	98.565			
12	.075	.501	99.065			
13	.058	.390	99.455			
14	.050	.331	99.786			
15	.032	.214	100.000			

Extraction Method: Maximum Likelihood.

Lab


Scree Plot
Eigenvalue
10
8
6
4
2
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
Factor Number

Lab

- Factor loadings

Rotated Factor Matrix[a]

	Factor		
	1	2	3
ca	.406	.824	-.145
cr	.897	.183	.141
cu	.829	.424	.089
fe	.886	.394	.019
mn	.627	.517	-.272
pb	.683	-.010	.464
zn	.721	.581	.181
cd	.366	.409	.436
mg	.772	.595	-.050
ni	.882	.388	.037
soc	-.099	-.102	.827
ph	.145	.631	-.066
n	.152	-.018	.871
k	.732	.542	-.013
p	.195	.503	.071

Extraction Method: Maximum Likelihood.
Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 6 iterations.

Appendix: Estimating Factor Loadings

- $\mathbf{X} = (X_1, X_2, \ldots, X_p)'$ is a random vector with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$.
- Statistical **Model:**

$$\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{L}F + \varepsilon \tag{2}$$

  - $\boldsymbol{L}$ denotes the **pattern matrix** of *factor loadings*
  - $\mathbf{F}$ denotes the vector of latent *factor scores*
  - $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_p)'$ denotes the vector of latent *error terms*

$$\boldsymbol{L} = \begin{bmatrix} l_{11} & l_{12} & \cdots & l_{1m} \\ l_{21} & l_{22} & \cdots & l_{2m} \\ \vdots & \vdots & & \vdots \\ l_{p1} & l_{p2} & \cdots & l_{pm} \end{bmatrix}$$

Appendix: Estimating Factor Loadings

Model assumptions (orthogonal FA model)

- $E(f_i) = 0$, $var(f_i) = 1$, and $cov(f_i, f_j) = 0$ if $i \neq j$
- $E(\varepsilon_i) = 0$, $var(\varepsilon_i) = \psi_i$, and $cov(\varepsilon_i, \varepsilon_j) = 0$ if $i \neq j$
  - We must allow each ε_i have different variances since it shows the residual part of X_i that is not in common with the other variables
- $\mathbf{F}$ and $\boldsymbol{\varepsilon}$ are independent

- The covariance between $\mathbf{X}$ and $\mathbf{F}$, $\mathbf{L\Phi}$ in the following, is called the factor **structure matrix**

$$cov(\mathbf{X}, \mathbf{F}) = E((\mathbf{X} - \boldsymbol{\mu})\mathbf{F}^{'}) = E((\mathbf{LF} + \varepsilon)\mathbf{F}^{'}) = \mathbf{L\Phi},$$

where $\mathbf{\Phi} = E(\mathbf{FF}^{'})$.

Appendix: Estimating Factor Loadings

- Covariance structure for $\mathbf{X}$:

$$var(\mathbf{X}) = E[(\mathbf{X} - \boldsymbol{\mu})(\mathbf{X} - \boldsymbol{\mu})^{'}] = \boldsymbol{L}\boldsymbol{L}^{'} + \boldsymbol{\Psi} \tag{3}$$

where $\boldsymbol{\Psi} = diag(\psi_1, \ldots, \psi_p)$.

- The **portion of variance** of the j-th variable that is explained by the m common factors is called the **communality** of the j-th variable:

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

  - $\sigma_{jj} = var(X_j)$, the j-th diagonal of $\boldsymbol{\Sigma}$
  - $h_j^2 = (\boldsymbol{L}\boldsymbol{L}^{'})_{jj} = l_{j1}^2 + l_{j2}^2 + \ldots + l_{jm}^2$ is the the **communality** of X_j; the sum of squared loadings
  - $\psi_j = var(\varepsilon_j)$

Appendix: Estimating Factor Loadings

- Model parameters:
  - the factor loadings $\boldsymbol{L}$ ($m \times p$ parameters) and
  - specific variances on the diagonal of $\boldsymbol{\Psi} = diag(\psi_1, \ldots, \psi_p)$.
- We need to estimate these parameters from Equation (3) by using the sample information $\mathbf{X}$.

- Let $\mathbf{S}$ be the sample covariance matrix of $\mathbf{X}$
- Denote the estimates of $\boldsymbol{L}$ and $\boldsymbol{\Psi}$ by $\hat{\boldsymbol{L}}$ and $\hat{\boldsymbol{\Psi}}$, respectively.
- We solve

$$\mathbf{S} \cong \hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{'} + \hat{\boldsymbol{\Psi}} \tag{4}$$

Factor analysis - PCA method

Principal Components Solution for Factor Analysis

- In the principal component approach, we neglect $\hat{\mathbf{\Psi}}$ and factor $\mathbf{S}$ into

$$\mathbf{S} = \hat{\boldsymbol{L}}\hat{\boldsymbol{L}}^{'}$$

- PCA Solution:

$$\begin{aligned} \hat{\boldsymbol{L}} = & \quad (\sqrt{\lambda_1}\mathbf{v}_1, \sqrt{\lambda_2}\mathbf{v}_2, \ldots, \sqrt{\lambda_m}\mathbf{v}_m) \\ \psi_j = & \quad s_{jj} - \hat{h}_j^2 = s_{jj} - \sum_{k=1}^{m} \hat{l}_{jk}^2 \end{aligned}$$

  where $\mathbf{\Sigma} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{'}$ is the eigenvalue decomposition of $\mathbf{\Sigma}$: $\lambda_1, \ldots, \lambda_m$ are the m largest eigenvalues of $\mathbf{S}$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ are the corresponding eigenvectors

- s_{jj} is an estimate of σ_{jj}

Factor analysis - PCA method

- Proportion of total sample variance explained by the k-th factor is

$$R_k^2 = \frac{\sum_{j=1}^{p} \hat{l_{jk}}^2}{\sum_{j=1}^{p} s_{jj}} = \frac{(\sqrt{\lambda_k}\mathbf{v}_k)'(\sqrt{\lambda_k}\mathbf{v}_k)}{\sum_{j=1}^{p} s_{jj}} = \frac{\lambda_k}{\sum_{j=1}^{p} s_{jj}}$$

Factor analysis - Iterated Principal Axis Factoring Method

- Denote the sample correlation matrix by $\mathbf{R}$ (standardized $\mathbf{S}$)

Assume we are applying FA to a sample correlation matrix

$$\mathbf{R} = \boldsymbol{L}\boldsymbol{L}^{'} + \boldsymbol{\Psi}$$

and we have some initial estimate of the specific variance $\hat{\psi}_j$.

- We can use $\hat{\psi}_j = 1/r^{jj}$, where r^{jj} is the j-th diagonal of of the inverse of $\mathbf{R}$, $\mathbf{R}^{-1}$.

Factor analysis - Iterated Principal Axis Factoring Method

The iterated principal axis factoring algorithm:

- Form $\tilde{\mathbf{R}} = \mathbf{R} - \hat{\boldsymbol{\Psi}}$ given current $\hat{\psi}_j$ estimates
- Update $\tilde{\boldsymbol{L}} = (\sqrt{\tilde{\lambda}_1}\tilde{\mathbf{v}}_1, \sqrt{\tilde{\lambda}_2}\tilde{\mathbf{v}}_2, \ldots, \sqrt{\tilde{\lambda}_m}\tilde{\mathbf{v}}_m)$,

where $\tilde{\mathbf{R}} = \tilde{\mathbf{V}}\tilde{\boldsymbol{\Lambda}}\tilde{\mathbf{V}}'$ is the eigenvalue decomposition of $\tilde{\mathbf{R}}$.

- Update $\hat{\psi}_j = 1 - \sum_{k=1}^{m} \tilde{l_{jk}}^2$

Factor analysis - MLE Method

Maximum Likelihood Estimation for Factor Analysis

- Assume the p X_i $(i = 1, \ldots, p)$ vectors are independent;
- Assume that X_i vector is a multivariate normal vector with mean $\boldsymbol{\mu}$ and var-covariance matrix $\mathbf{L}\mathbf{L}^{'} + \boldsymbol{\Psi}$.
- The ML estimators are obtained by maximizing the log-likelihood function.

License