Multivariate Verfahren 2

factor analysis

Helmut Waldl

April 16th and 17th 2012

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

preface

With factor analysis we do not mean one special method, it is rather a general term for many different techniques.

The purpose of factor analysis is the mapping of a (big) set of observable variables to a set of as few as possible **latent variables**, the **factors**.

There are three basic methods of factor analysis

- maximum likelihood factor analysis
- principal component analysis
- principal factor analysis (principal-axis factoring, common factor analysis)

factor analysis

exploratory vs confirmatory factor analysis

Exploratory factor analysis (EFA)

Based on empirical data we try to find a minimal number of factors that may explain the observable variables as good as possible. Factor analysis in this sense is generating hypotheses and models.

Confirmatory factor analysis (CFA)

We start with a model with certain structural assumptions and a given number of factors. Based on a sample we estimate the parameters and test the goodness of fit of the model.

Unlike other multivariate methods like ANOVA or GLM the factors may not be measured or recorded empirically. Factors are latent variables, a result of the factor analysis model.

factor analysis

exploratory vs confirmatory factor analysis

Exploratory factor analysis (EFA)

Based on empirical data we try to find a minimal number of factors that may explain the observable variables as good as possible. Factor analysis in this sense is generating hypotheses and models.

Confirmatory factor analysis (CFA)

We start with a model with certain structural assumptions and a given number of factors. Based on a sample we estimate the parameters and test the goodness of fit of the model.

Unlike other multivariate methods like ANOVA or GLM the factors may not be measured or recorded empirically. Factors are latent variables, a result of the factor analysis model.

factor analysis

exploratory vs confirmatory factor analysis

<u>Exploratory factor analysis (EFA)</u>
Based on empirical data we try to find a minimal number of factors that may explain the observable variables as good as possible. Factor analysis in this sense is generating hypotheses and models.

<u>Confirmatory factor analysis (CFA)</u>
We start with a model with certain structural assumptions and a given number of factors. Based on a sample we estimate the parameters and test the goodness of fit of the model.

Unlike other multivariate methods like ANOVA or GLM the factors may not be measured or recorded empirically. Factors are latent variables, a result of the factor analysis model.

factor analysis

the statistical model

<u>Given:</u> p observable random variates $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ with expectation $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \dots & \mu_p \end{pmatrix}^T$ and covariance matrix Σ

<u>Assumption:</u> There exist $k < p$ non-observable random variates $\mathbf{f} = \begin{pmatrix} f_1 & \dots & f_k \end{pmatrix}^T$, the **common factors** of $\mathbf{y}$,
p other variates $\mathbf{e} = \begin{pmatrix} e_1 & \dots & e_p \end{pmatrix}^T$
plus a $(p \times k)$ matrix of coefficients $L = (l_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,k}}$ such that we have the following relation:

$$
\begin{aligned}
y_1 &= \mu_1 + l_{11} f_1 + l_{12} f_2 + \ldots + l_{1k} f_k + e_1 \\
y_2 &= \mu_2 + l_{21} f_1 + l_{22} f_2 + \ldots + l_{2k} f_k + e_2 \\
&\vdots \\
y_p &= \mu_p + l_{p1} f_1 + l_{p2} f_2 + \ldots + l_{pk} f_k + e_p
\end{aligned}
$$

factor analysis

the statistical model

Given: p observable random variates $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ with expectation $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \dots & \mu_p \end{pmatrix}^T$ and covariance matrix Σ

Assumption: There exist $k < p$ non-observable random variates $\mathbf{f} = \begin{pmatrix} f_1 & \dots & f_k \end{pmatrix}^T$, the **common factors** of $\mathbf{y}$,

p other variates $\mathbf{e} = \begin{pmatrix} e_1 & \dots & e_p \end{pmatrix}^T$

plus a $(p \times k)$ matrix of coefficients $L = (l_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,k}}$ such that we have the following relation:

$$
\begin{aligned}
y_1 &= \mu_1 + l_{11} f_1 + l_{12} f_2 + \ldots + l_{1k} f_k + e_1 \\
y_2 &= \mu_2 + l_{21} f_1 + l_{22} f_2 + \ldots + l_{2k} f_k + e_2 \\
&\vdots \\
y_p &= \mu_p + l_{p1} f_1 + l_{p2} f_2 + \ldots + l_{pk} f_k + e_p
\end{aligned}
$$

factor analysis

the statistical model

Given: p observable random variates $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ with expectation $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \dots & \mu_p \end{pmatrix}^T$ and covariance matrix Σ

Assumption: There exist $k < p$ non-observable random variates $\mathbf{f} = \begin{pmatrix} f_1 & \dots & f_k \end{pmatrix}^T$, the **common factors** of $\mathbf{y}$,
p other variates $\mathbf{e} = \begin{pmatrix} e_1 & \dots & e_p \end{pmatrix}^T$
plus a $(p \times k)$ matrix of coefficients $L = (l_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,k}}$ such that we have the following relation:

$$
\begin{aligned}
y_1 &= \mu_1 + l_{11} f_1 + l_{12} f_2 + \dots + l_{1k} f_k + e_1 \\
y_2 &= \mu_2 + l_{21} f_1 + l_{22} f_2 + \dots + l_{2k} f_k + e_2 \\
&\vdots \\
y_p &= \mu_p + l_{p1} f_1 + l_{p2} f_2 + \dots + l_{pk} f_k + e_p
\end{aligned}
$$

factor analysis

the statistical model

Given: p observable random variates $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ with expectation $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \dots & \mu_p \end{pmatrix}^T$ and covariance matrix Σ

Assumption: There exist $k < p$ non-observable random variates $\mathbf{f} = \begin{pmatrix} f_1 & \dots & f_k \end{pmatrix}^T$, the **common factors** of $\mathbf{y}$,
p other variates $\mathbf{e} = \begin{pmatrix} e_1 & \dots & e_p \end{pmatrix}^T$
plus a $(p \times k)$ matrix of coefficients $L = (l_{ij})_{\substack{i=1,\dots,p \\ j=1,\dots,k}}$ such that we have the following relation:

$$
\begin{aligned}
y_1 &= \mu_1 + l_{11} f_1 + l_{12} f_2 + \ldots + l_{1k} f_k + e_1 \\
y_2 &= \mu_2 + l_{21} f_1 + l_{22} f_2 + \ldots + l_{2k} f_k + e_2 \\
&\vdots \\
y_p &= \mu_p + l_{p1} f_1 + l_{p2} f_2 + \ldots + l_{pk} f_k + e_p
\end{aligned}
$$

factor analysis

the statistical model

Given: p observable random variates $\mathbf{y} = \begin{pmatrix} y_1 & \ldots & y_p \end{pmatrix}^T$ with expectation $\boldsymbol{\mu} = \begin{pmatrix} \mu_1 & \ldots & \mu_p \end{pmatrix}^T$ and covariance matrix Σ

Assumption: There exist $k < p$ non-observable random variates $\mathbf{f} = \begin{pmatrix} f_1 & \ldots & f_k \end{pmatrix}^T$, the **common factors** of $\mathbf{y}$,
p other variates $\mathbf{e} = \begin{pmatrix} e_1 & \ldots & e_p \end{pmatrix}^T$
plus a $(p \times k)$ matrix of coefficients $L = (l_{ij})_{\substack{i=1,\ldots,p \\ j=1,\ldots,k}}$ such that we have the following relation:

$$
\begin{aligned}
y_1 &= \mu_1 + l_{11} f_1 + l_{12} f_2 + \ldots + l_{1k} f_k + e_1 \\
y_2 &= \mu_2 + l_{21} f_1 + l_{22} f_2 + \ldots + l_{2k} f_k + e_2 \\
&\vdots \\
y_p &= \mu_p + l_{p1} f_1 + l_{p2} f_2 + \ldots + l_{pk} f_k + e_p
\end{aligned}
$$

factor analysis

the statistical model

In compact matrix and vector notation we may write the above system of equations as

$$\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$$

The coefficients l_{ij} are denoted as **loadings** of variate y_i on factor f_j, $L = (l_{ij})$ is called the **loading matrix**.

In the variates e_i all effects are summarized that affect only variate y_i, i.e. **unique factors** associated with the original variate y_i and measurement errors.

Furthermore we assume w.l.o.g. $E(\mathbf{f}) = E(\mathbf{e}) = \mathbf{o}$ and $Cov(\mathbf{e}) = E(\mathbf{e} \cdot \mathbf{e}^T) = V = \mathrm{diag}(v_1^2, \ldots, v_p^2)$, i.e. the unique factors e_i are uncorrelated.
V is also denoted as (individual) **error covariance matrix**

factor analysis

the statistical model

In compact matrix and vector notation we may write the above system of equations as

$$\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$$

The coefficients l_{ij} are denoted as **loadings** of variate y_i on factor f_j, $L = (l_{ij})$ is called the **loading matrix**.

In the variates e_i all effects are summarized that affect only variate y_i, i.e. **unique factors** associated with the original variate y_i and measurement errors.

Furthermore we assume w.l.o.g. $E(\mathbf{f}) = E(\mathbf{e}) = \mathbf{o}$ and $Cov(\mathbf{e}) = E(\mathbf{e} \cdot \mathbf{e}^T) = V = \text{diag}(v_1^2, \ldots, v_p^2)$, i.e. the unique factors e_i are uncorrelated.
V is also denoted as (individual) **error covariance matrix**

factor analysis

the statistical model

In compact matrix and vector notation we may write the above system of equations as

$$\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$$

The coefficients l_{ij} are denoted as **loadings** of variate y_i on factor f_j, $L = (l_{ij})$ is called the **loading matrix**.

In the variates e_i all effects are summarized that affect only variate y_i, i.e. **unique factors** associated with the original variate y_i and measurement errors.

Furthermore we assume w.l.o.g. $E(\mathbf{f}) = E(\mathbf{e}) = \mathbf{o}$ and $Cov(\mathbf{e}) = E(\mathbf{e} \cdot \mathbf{e}^T) = V = \text{diag}(v_1^2, \ldots, v_p^2)$, i.e. the unique factors e_i are uncorrelated.
V is also denoted as (individual) **error covariance matrix**

factor analysis

the statistical model

In compact matrix and vector notation we may write the above system of equations as

$$\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$$

The coefficients l_{ij} are denoted as **loadings** of variate y_i on factor f_j, $L = (l_{ij})$ is called the **loading matrix**.

In the variates e_i all effects are summarized that affect only variate y_i, i.e. **unique factors** associated with the original variate y_i and measurement errors.

Furthermore we assume w.l.o.g. $E(\mathbf{f}) = E(\mathbf{e}) = \mathbf{o}$ and $Cov(\mathbf{e}) = E(\mathbf{e} \cdot \mathbf{e}^T) = V = \text{diag}(v_1^2, \ldots, v_p^2)$, i.e. the unique factors e_i are uncorrelated.

V is also denoted as (individual) **error covariance matrix**

factor analysis

the statistical model

In compact matrix and vector notation we may write the above system of equations as

$$\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$$

The coefficients l_{ij} are denoted as **loadings** of variate y_i on factor f_j, $L = (l_{ij})$ is called the **loading matrix**.

In the variates e_i all effects are summarized that affect only variate y_i, i.e. **unique factors** associated with the original variate y_i and measurement errors.

Furthermore we assume w.l.o.g. $E(\mathbf{f}) = E(\mathbf{e}) = \mathbf{o}$ and $Cov(\mathbf{e}) = E(\mathbf{e} \cdot \mathbf{e}^T) = V = \text{diag}(v_1^2, \ldots, v_p^2)$, i.e. the unique factors e_i are uncorrelated.
V is also denoted as (individual) **error covariance matrix**

factor analysis

the statistical model

Regarding the factors we assume them to be standardized ($Var(f_i) = 1$) and not correlating with the unique factors: $Cov(\mathbf{f}, \mathbf{e}) = E(\mathbf{f} \cdot \mathbf{e}^T) = \mathbf{0}$

In an **orthogonal factor model** (contrary to an **oblique factor model**) we further assume uncorrelated factors:

$$Cov(\mathbf{f}) = E(\mathbf{f} \cdot \mathbf{f}^T) = \Phi = I$$

In the above model all dependencies between the observable variables are explained via the common factors.

factor analysis

the statistical model

Regarding the factors we assume them to be standardized ($Var(f_i) = 1$) and not correlating with the unique factors: $Cov(\mathbf{f}, \mathbf{e}) = E(\mathbf{f} \cdot \mathbf{e}^T) = \mathbf{0}$

In an **orthogonal factor model** (contrary to an **oblique factor model**) we further assume uncorrelated factors:

$$Cov(\mathbf{f}) = E(\mathbf{f} \cdot \mathbf{f}^T) = \Phi = I$$

In the above model all dependencies between the observable variables are explained via the common factors.

factor analysis

the statistical model

Regarding the factors we assume them to be standardized ($Var(f_i) = 1$) and not correlating with the unique factors: $Cov(\mathbf{f}, \mathbf{e}) = E(\mathbf{f} \cdot \mathbf{e}^T) = \mathbf{0}$

In an **orthogonal factor model** (contrary to an **oblique factor model**) we further assume uncorrelated factors:

$$Cov(\mathbf{f}) = E(\mathbf{f} \cdot \mathbf{f}^T) = \Phi = I$$

In the above model all dependencies between the observable variables are explained via the common factors.

factor analysis

data model

We have n independent observations $\mathbf{y}_i = \begin{pmatrix} y_{i1} & \dots & y_{ip} \end{pmatrix}^T \quad i = 1, \dots, n$ of $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ and combine them to the $(n \times p)$-data matrix $\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 & \dots & \mathbf{y}_n \end{pmatrix}^T$:

$$\mathbf{Y} = \begin{pmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{np} \end{pmatrix} = \begin{pmatrix} Y_1 & \cdots & Y_p \end{pmatrix}$$

The column vectors $Y_j = \begin{pmatrix} y_{1j} \\ \vdots \\ y_{nj} \end{pmatrix}$ matter in factor analysis (contrary to ANOVA and GLM)

factor analysis

data model

We have n independent observations $\mathbf{y}_i = \begin{pmatrix} y_{i1} & \dots & y_{ip} \end{pmatrix}^T \quad i = 1, \dots, n$ of $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ and combine them to the $(n \times p)$-data matrix $\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 & \dots & \mathbf{y}_n \end{pmatrix}^T$:

$$\mathbf{Y} = \begin{pmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{np} \end{pmatrix} = \begin{pmatrix} Y_1 & \cdots & Y_p \end{pmatrix}$$

The column vectors $Y_j = \begin{pmatrix} y_{1j} \\ \vdots \\ y_{nj} \end{pmatrix}$ matter in factor analysis (contrary to ANOVA and GLM)

factor analysis

data model

We have n independent observations
$\mathbf{y}_i = \begin{pmatrix} y_{i1} & \dots & y_{ip} \end{pmatrix}^T \quad i = 1, \dots, n$ of $\mathbf{y} = \begin{pmatrix} y_1 & \dots & y_p \end{pmatrix}^T$ and combine them to the $(n \times p)$-data matrix $\mathbf{Y} = \begin{pmatrix} \mathbf{y}_1 & \dots & \mathbf{y}_n \end{pmatrix}^T$:

$$\mathbf{Y} = \begin{pmatrix} y_{11} & \cdots & y_{1p} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{np} \end{pmatrix} = \begin{pmatrix} Y_1 & \cdots & Y_p \end{pmatrix}$$

The column vectors $Y_j = \begin{pmatrix} y_{1j} \\ \vdots \\ y_{nj} \end{pmatrix}$ matter in factor analysis (contrary to ANOVA and GLM)

factor analysis

data model

With the *i*-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \boldsymbol{\mu} & \dots & \boldsymbol{\mu} \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \mu & \dots & \mu \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \mu & \dots & \mu \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \boldsymbol{\mu} & \dots & \boldsymbol{\mu} \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \boldsymbol{\mu} & \dots & \boldsymbol{\mu} \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \mu & \dots & \mu \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

data model

With the i-th observation we affiliate

k **factor values** $\mathbf{f}_i = \begin{pmatrix} f_{i1} & \dots & f_{ik} \end{pmatrix}^T$ and

p **unique factor values** $\mathbf{e}_i = \begin{pmatrix} e_{i1} & \dots & e_{ip} \end{pmatrix}^T$

We combine them to the $(n \times k)$-**factor matrix**
$F = \begin{pmatrix} \mathbf{f}_1 & \dots & \mathbf{f}_n \end{pmatrix}^T = \begin{pmatrix} F_1 & \dots & F_k \end{pmatrix}$

and the $(n \times p)$-**unique factor matrix**
$E = \begin{pmatrix} \mathbf{e}_1 & \dots & \mathbf{e}_n \end{pmatrix}^T = \begin{pmatrix} E_1 & \dots & E_p \end{pmatrix}$

note again the column vectors $F_1, \dots, F_k$ and $E_1, \dots, E_p$!

With the $(n \times p)$-matrix $M = \begin{pmatrix} \boldsymbol{\mu} & \dots & \boldsymbol{\mu} \end{pmatrix}^T$ we may write the data model as

$$Y - M = F \cdot L^T + E$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables $(\mathbf{e}, \mathbf{f})$. This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \mu)(\mathbf{y} - \mu)^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ $(L \cdot \Phi \cdot L^T)$ and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model $(\Phi = I)$ we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables $(\mathbf{e},\ \mathbf{f})$. This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ $(L \cdot \Phi \cdot L^T)$ and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model $(\Phi = I)$ we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables $(\mathbf{e}, \mathbf{f})$. This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ $(L \cdot \Phi \cdot L^T)$ and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model $(\Phi = I)$ we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables $(\mathbf{e}, \mathbf{f})$. This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ $(L \cdot \Phi \cdot L^T)$ and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model $(\Phi = I)$ we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables $(\mathbf{e}, \mathbf{f})$. This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ $(L \cdot \Phi \cdot L^T)$ and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model $(\Phi = I)$ we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

In the factor analysis model $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we explain p residuals with $k + p$ latent variables ($\mathbf{e}$, $\mathbf{f}$). This model cannot be checked empirically.

The model assumptions concerning expectations and covariances imply the **fundamental theorem of factor analysis**:

$$\Sigma = E((\mathbf{y} - \boldsymbol{\mu})(\mathbf{y} - \boldsymbol{\mu})^T) = E((L \cdot \mathbf{f} + \mathbf{e})(L \cdot \mathbf{f} + \mathbf{e})^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} \cdot L^T + L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{e}^T)}_{=\mathbf{0}} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{=\mathbf{0}} \cdot L^T + \underbrace{E(\mathbf{e} \cdot \mathbf{e}^T)}_{=V}$$

i.e. the covariance matrix Σ can be decomposed into a **factor covariance matrix** Φ ($L \cdot \Phi \cdot L^T$) and an **error covariance matrix** V:

$$\Sigma = L \cdot \Phi \cdot L^T + V$$

In the orthogonal factor model ($\Phi = I$) we get

$$\Sigma = L \cdot L^T + V$$

factor analysis

the fundamental theorem

For the variances $Var(y_i) = \sigma_{ii} = \sigma_i^2$ we get

$$\sigma_i^2 = \sum_{j=1}^{k} l_{ij}^2 + v_i^2 = h_i^2 + v_i^2$$

i.e. the variance σ_i^2 of y_i can be decomposed into **communality** h_i^2 and **specific variance** v_i^2

Further we get:

$$Cov(\mathbf{y}, \mathbf{f}) = E((\mathbf{y} - \mu) \cdot \mathbf{f}^T) = E((L \cdot \mathbf{f} + \mathbf{e}) \cdot \mathbf{f}^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{\mathbf{0}} = L \cdot \Phi$$

i.e. in the orthogonal factor model ($\Phi = I$) the loading matrix is also the covariance matrix between observable and latent variables: $Cov(\mathbf{y}, \mathbf{f}) = L$.

factor analysis

the fundamental theorem

For the variances $Var(y_i) = \sigma_{ii} = \sigma_i^2$ we get

$$\sigma_i^2 = \sum_{j=1}^{k} l_{ij}^2 + v_i^2 = h_i^2 + v_i^2$$

i.e. the variance σ_i^2 of y_i can be decomposed into **communality** h_i^2 and **specific variance** v_i^2

Further we get:

$$Cov(\mathbf{y}, \mathbf{f}) = E((\mathbf{y} - \mu) \cdot \mathbf{f}^T) = E((L \cdot \mathbf{f} + \mathbf{e}) \cdot \mathbf{f}^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{\mathbf{0}} = L \cdot \Phi$$

i.e. in the orthogonal factor model ($\Phi = I$) the loading matrix is also the covariance matrix between observable and latent variables: $Cov(\mathbf{y}, \mathbf{f}) = L$.

factor analysis

the fundamental theorem

For the variances $Var(y_i) = \sigma_{ii} = \sigma_i^2$ we get

$$\sigma_i^2 = \sum_{j=1}^{k} l_{ij}^2 + v_i^2 = h_i^2 + v_i^2$$

i.e. the variance σ_i^2 of y_i can be decomposed into **communality** h_i^2 and **specific variance** v_i^2

Further we get:

$$Cov(\mathbf{y}, \mathbf{f}) = E((\mathbf{y} - \boldsymbol{\mu}) \cdot \mathbf{f}^T) = E((L \cdot \mathbf{f} + \mathbf{e}) \cdot \mathbf{f}^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{\mathbf{0}} = L \cdot \Phi$$

i.e. in the orthogonal factor model ($\Phi = I$) the loading matrix is also the covariance matrix between observable and latent variables: $Cov(\mathbf{y}, \mathbf{f}) = L$.

factor analysis

the fundamental theorem

For the variances $Var(y_i) = \sigma_{ii} = \sigma_i^2$ we get

$$\sigma_i^2 = \sum_{j=1}^{k} l_{ij}^2 + v_i^2 = h_i^2 + v_i^2$$

i.e. the variance σ_i^2 of y_i can be decomposed into **communality** h_i^2 and **specific variance** v_i^2

Further we get:

$$Cov(\mathbf{y}, \mathbf{f}) = E((\mathbf{y} - \boldsymbol{\mu}) \cdot \mathbf{f}^T) = E((L \cdot \mathbf{f} + \mathbf{e}) \cdot \mathbf{f}^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{\mathbf{0}} = L \cdot \Phi$$

i.e. in the orthogonal factor model ($\Phi = I$) the loading matrix is also the covariance matrix between observable and latent variables: $Cov(\mathbf{y}, \mathbf{f}) = L$.

factor analysis

the fundamental theorem

For the variances $Var(y_i) = \sigma_{ii} = \sigma_i^2$ we get

$$\sigma_i^2 = \sum_{j=1}^{k} l_{ij}^2 + v_i^2 = h_i^2 + v_i^2$$

i.e. the variance σ_i^2 of y_i can be decomposed into **communality** h_i^2 and **specific variance** v_i^2

Further we get:

$$Cov(\mathbf{y}, \mathbf{f}) = E((\mathbf{y} - \boldsymbol{\mu}) \cdot \mathbf{f}^T) = E((L \cdot \mathbf{f} + \mathbf{e}) \cdot \mathbf{f}^T) =$$

$$= L \cdot \underbrace{E(\mathbf{f} \cdot \mathbf{f}^T)}_{=\Phi} + \underbrace{E(\mathbf{e} \cdot \mathbf{f}^T)}_{\mathbf{0}} = L \cdot \Phi$$

i.e. in the orthogonal factor model ($\Phi = I$) the loading matrix is also the covariance matrix between observable and latent variables: $Cov(\mathbf{y}, \mathbf{f}) = L$.

factor analysis

estimation task of the factor analysis model

The parameters of our model are:

- **the vector of expectations** μ which is relatively irrelevant in factor analysis.

  We are primarily interested in the explanation of the variability of the data.

  We may estimate μ unbiased with $\bar{\mathbf{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i$

- **the covariance matrix** Σ described by k, L and V

factor analysis

estimation task of the factor analysis model

The parameters of our model are:

- **the vector of expectations $\boldsymbol{\mu}$** which is relatively irrelevant in factor analysis.

  We are primarily interested in the explanation of the variability of the data.

  We may estimate $\boldsymbol{\mu}$ unbiased with $\bar{\mathbf{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i$

- **the covariance matrix** Σ described by k, L and V

factor analysis

estimation task of the factor analysis model

The parameters of our model are:

- **the vector of expectations** $\boldsymbol{\mu}$ which is relatively irrelevant in factor analysis.

  We are primarily interested in the explanation of the variability of the data.

  We may estimate $\boldsymbol{\mu}$ unbiased with $\bar{\mathbf{y}} = \frac{1}{n}\sum_{i=1}^{n} y_i$

- **the covariance matrix** Σ described by k, L and V

factor analysis

estimation task of the factor analysis model

The factor analytical approach starts with the fundamental theorem $\Sigma = L \cdot L^T + V$:

1. We have to estimate the unknown parameters k, L and V, this is done by means of ML factor analysis or principal component analysis. The empirical covariance matrix S is our data source.
2. To facilitate the interpretation of factors and the connection with the observable variables we use factor rotation. Rotation serves to make the output more understandable.
3. After the estimation of μ, k, L and V we may estimate the factor values of an arbitrary observation $\mathbf{y}_i$ and combine them to the factor matrix F. Then we may also compute the unique factor matrix E and the partitioning of $Y - M = F \cdot L^T + E$ is realized completely.

The factor analytical approach starts with the fundamental theorem $\Sigma = L \cdot L^T + V$:

1. We have to estimate the unknown parameters k, L and V, this is done by means of ML factor analysis or principal component analysis. The empirical covariance matrix S is our data source.
2. To facilitate the interpretation of factors and the connection with the observable variables we use factor rotation.
   Rotation serves to make the output more understandable.
3. After the estimation of μ, k, L and V we may estimate the factor values of an arbitrary observation $\mathbf{y}_i$ and combine them to the factor matrix F.
   Then we may also compute the unique factor matrix E and the partitioning of $Y - M = F \cdot L^T + E$ is realized completely.

factor analysis

estimation task of the factor analysis model

The factor analytical approach starts with the fundamental theorem $\Sigma = L \cdot L^T + V$:

1. We have to estimate the unknown parameters k, L and V, this is done by means of ML factor analysis or principal component analysis. The empirical covariance matrix S is our data source.
2. To facilitate the interpretation of factors and the connection with the observable variables we use factor rotation.
   Rotation serves to make the output more understandable.
3. After the estimation of μ, k, L and V we may estimate the factor values of an arbitrary observation $\mathbf{y}_i$ and combine them to the factor matrix F.
   Then we may also compute the unique factor matrix E and the partitioning of $Y - M = F \cdot L^T + E$ is realized completely.

factor analysis

estimation task of the factor analysis model

The factor analytical approach starts with the fundamental theorem $\Sigma = L \cdot L^T + V$:

1. We have to estimate the unknown parameters k, L and V, this is done by means of ML factor analysis or principal component analysis. The empirical covariance matrix S is our data source.
2. To facilitate the interpretation of factors and the connection with the observable variables we use factor rotation. Rotation serves to make the output more understandable.
3. After the estimation of $\boldsymbol{\mu}$, k, L and V we may estimate the factor values of an arbitrary observation $\mathbf{y}_i$ and combine them to the factor matrix F. Then we may also compute the unique factor matrix E and the partitioning of $Y - M = F \cdot L^T + E$ is realized completely.

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.

If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**

   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.

   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and

   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

In the fundamental theorem $\Sigma = L \cdot L^T + V$ the covariance matrix Σ is solved for the unknown parameters k, L and V.
If the representation is not unique we don't get consistent estimators.

1. **Existence of L and V**
   From matrix diagonalisation of symmetric matrices we know that for $k = p$ we may represent the covariance matrix according to the fundamental theorem, and that with $V = \mathbf{0}$.
   For a given $k \leq p$ and V the above decomposition exists if the reduced covariance matrix $\Sigma - V$ is positive semidefinite with rank $rk(\Sigma - V) = k$.

   Unfortunately we may not always decompose covariance matrices according to $\Sigma = L \cdot L^T + V$ and
   frequently this decomposition exists mathematically but the result is meaningless from the statistical point of view (e.g. negative specific variances).

factor analysis

identifiability problems

2. **Uniqueness of L with given k and V**

   For $k = 1$ the factor model is identifiable.

   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

   $$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

   $$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

   $$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot {L^*}^T + V$$

factor analysis

identifiability problems

2. **Uniqueness of L with given k and V**

   For $k = 1$ the factor model is identifiable.

   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

   $$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

   $$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

   $$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot {L^*}^T + V$$

2. **Uniqueness of L with given k and V**
   For $k = 1$ the factor model is identifiable.
   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

   $$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

   $$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

   $$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot L^{*T} + V$$

factor analysis

identifiability problems

2. **Uniqueness of L with given k and V**
   For $k = 1$ the factor model is identifiable.
   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \mu = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

   $$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \mu$$

   $$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

   $$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot L^{*T} + V$$

2. **Uniqueness of L with given k and V**

   For $k = 1$ the factor model is identifiable.
   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

$$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

$$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

$$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot {L^*}^T + V$$

2. **Uniqueness of L with given k and V**

   For $k = 1$ the factor model is identifiable.
   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

$$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

$$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

$$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot L^{*T} + V$$

factor analysis

identifiability problems

2. **Uniqueness of L with given k and V**
   For $k = 1$ the factor model is identifiable.
   However it is not for $k > 1$!

   If we may solve $\Sigma = L \cdot L^T + V$ for L with k and V given there always exist infinitely many different solutions:

   Let M be an orthogonal $(k \times k)$-matrix (i.e. $M^{-1} = M^T$). Now in our model equation $\mathbf{y} - \boldsymbol{\mu} = L \cdot \mathbf{f} + \mathbf{e}$ we substitute L for $L^* = L \cdot M$ and $\mathbf{f}$ for $\mathbf{f}^* = M^T \cdot \mathbf{f}$. Then we get

   $$L^* \cdot \mathbf{f}^* + \mathbf{e} = L \cdot M \cdot M^T \cdot \mathbf{f} + \mathbf{e} = L \cdot \mathbf{f} + \mathbf{e} = \mathbf{y} - \boldsymbol{\mu}$$

   $$E(\mathbf{f}^*) = M \cdot E(\mathbf{f}) = \mathbf{o} \qquad Cov(\mathbf{f}^*) = M^T \cdot Cov(\mathbf{f}) \cdot M = M^T \cdot M = I$$

   and L^* decomposes Σ as well as L:

   $$\Sigma = L \cdot L^T + V = L \cdot M \cdot M^T \cdot L^T + V = L^* \cdot {L^*}^T + V$$

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations.
Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ 0 of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.
There are also other criteria (see factor rotation)

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations. Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ 0 of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.
There are also other criteria (see factor rotation)

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations. Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ 0 of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.
There are also other criteria (see factor rotation)

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations. Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k\ 0$ of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.
There are also other criteria (see factor rotation)

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations. Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ 0 of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.

There are also other criteria (see factor rotation)

factor analysis

identifiability problems

Factor loadings are unique except for orthogonal transformations. Multiplications with orthogonal matrices are rotations of the coordinate system from geometrical point of view. $\longrightarrow$ factor rotation.

To get unique loadings L we need additional side conditions. Most frequently one uses:

$$L^T \cdot V^{-1} \cdot L = \Lambda^{\frac{1}{2}} \cdot \Omega^T \cdot V^{\frac{1}{2}} \cdot V^{-1} \cdot V^{\frac{1}{2}} \cdot \Omega \cdot \Lambda^{\frac{1}{2}} \qquad \text{diagonal}$$

where Λ is the diagonal matrix of the descending ordered k positive eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$ 0 of the rescaled covariance matrix $\Sigma^* = V^{-\frac{1}{2}} \cdot \Sigma \cdot V^{-\frac{1}{2}}$ and Ω the matrix of the associated orthonormal eigenvectors.

This condition ($L^T \cdot V^{-1} \cdot L$ diagonal) to get unique loadings is selected merely arbitrarily and not justified by some theory.
There are also other criteria (see factor rotation)

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \text{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

   $rk(\Sigma - V) \leq k$ $\quad$ $\Sigma - V$ positive semidefinite

   V can only be unique if k is minimal.

   The communality problem:

   Find $\min_{V>0}\{rk(\Sigma - V)\} = k$
   with side condition $\Sigma - V$ positive semidefinite
   and decide whether V is unique in this setting.

factor analysis

identifiability problems

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \text{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

   $$rk(\Sigma - V) \leq k \qquad \Sigma - V \text{ positive semidefinite}$$

   V can only be unique if k is minimal.

   The communality problem:

   Find $\min_{V>0}\{rk(\Sigma - V)\} = k$
   with side condition $\Sigma - V$ positive semidefinite
   and decide whether V is unique in this setting.

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \text{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

$$rk(\Sigma - V) \leq k \qquad \Sigma - V \text{ positive semidefinite}$$

V can only be unique if k is minimal.

The communality problem:

Find $\min_{V>0}\{rk(\Sigma - V)\} = k$
with side condition $\Sigma - V$ positive semidefinite
and decide whether V is unique in this setting.

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \text{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

$$rk(\Sigma - V) \leq k \qquad \Sigma - V \text{ positive semidefinite}$$

V can only be unique if k is minimal.

The communality problem:

Find $\min\limits_{V>0}\{rk(\Sigma - V)\} = k$
with side condition $\Sigma - V$ positive semidefinite
and decide whether V is unique in this setting.

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \operatorname{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

$$rk(\Sigma - V) \leq k \qquad \Sigma - V \text{ positive semidefinite}$$

V can only be unique if k is minimal.

The communality problem:

Find $\min_{V>0}\{rk(\Sigma - V)\} = k$
with side condition $\Sigma - V$ positive semidefinite
and decide whether V is unique in this setting.

factor analysis

identifiability problems

3. **Uniqueness of k and V (communality problem)**
   Necessary and sufficient conditions for the existence of the decomposition $\Sigma = L \cdot L^T + V$ with $(p \times k)$ matrix L and $V = \text{diag}(v_1^2, \ldots, v_p^2) > 0$ and of k and V are

   $$rk(\Sigma - V) \leq k \qquad \Sigma - V \text{ positive semidefinite}$$

   V can only be unique if k is minimal.

   The communality problem:

   Find $\min_{V>0}\{rk(\Sigma - V)\} = k$
   with side condition $\Sigma - V$ positive semidefinite
   and decide whether V is unique in this setting.

Necessary and sufficient conditions for the uniqueness of V have been found only for $k \leq 2$

Central identifiability theorem If $p \geq 2 \cdot k + 1$ And L has the form $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ with nonsingular $(k \times k)$-matrices L_1, L_3 and $(p - 2k \times k)$-matrix L_2
then V is unique.

Remark: the assumption $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ implies that any column of L has at least 2 nonzero elements.

Remark: altogether the conditions for existence and uniqueness of L, k and V are satisfied almost always. Anyway there may be identifiability problems in practise.

factor analysis

identifiability problems

Necessary and sufficient conditions for the uniqueness of V have been found only for $k \leq 2$

Central identifiability theorem If $p \geq 2 \cdot k + 1$ And L has the form $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ with nonsingular $(k \times k)$-matrices L_1, L_3 and $(p - 2k \times k)$-matrix L_2
then V is unique.

Remark: the assumption $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ implies that any column of L has at least 2 nonzero elements.

Remark: altogether the conditions for existence and uniqueness of L, k and V are satisfied almost always. Anyway there may be identifiability problems in practise.

factor analysis

identifiability problems

Necessary and sufficient conditions for the uniqueness of V have been found only for $k \leq 2$

Central identifiability theorem If $p \geq 2 \cdot k + 1$ And L has the form $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ with nonsingular $(k \times k)$-matrices L_1, L_3 and $(p - 2k \times k)$-matrix L_2
then V is unique.

Remark: the assumption $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ implies that any column of L has at least 2 nonzero elements.

Remark: altogether the conditions for existence and uniqueness of L, k and V are satisfied almost always. Anyway there may be identifiability problems in practise.

factor analysis

identifiability problems

Necessary and sufficient conditions for the uniqueness of V have been found only for $k \leq 2$

Central identifiability theorem If $p \geq 2 \cdot k + 1$ And L has the form $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ with nonsingular $(k \times k)$-matrices L_1, L_3 and $(p - 2k \times k)$-matrix L_2
then V is unique.

Remark: the assumption $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ implies that any column of L has at least 2 nonzero elements.

Remark: altogether the conditions for existence and uniqueness of L, k and V are satisfied almost always. Anyway there may be identifiability problems in practise.

factor analysis

identifiability problems

Necessary and sufficient conditions for the uniqueness of V have been found only for $k \leq 2$

Central identifiability theorem If $p \geq 2 \cdot k + 1$ And L has the form

$$L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$$ with nonsingular $(k \times k)$-matrices L_1, L_3 and $(p - 2k \times k)$-matrix L_2

then V is unique.

Remark: the assumption $L = \begin{pmatrix} L_1 \\ L_2 \\ L_3 \end{pmatrix}$ implies that any column of L has at least 2 nonzero elements.

Remark: altogether the conditions for existence and uniqueness of L, k and V are satisfied almost always. Anyway there may be identifiability problems in practise.