Lehrstuhl fuer Statistik, Oekonometrie und empirische Wirtschaftsforschung, Universitaet Tuebingen

Faktorenanalyse

Dr. S. Prohl

22. Juni 2007/ SoSe 2007

Faktorenanalyse

Literatur:

- Handl, A.: Multivariate Verfahren: Theorie und Praxis multivariater Verfahren unter besonderer Beruecksichtigung von S-Plus, 2002, Kapitel 9.
- Backhaus, K., Erichson, B., Plinke, W., und Weiber, R.: Multivariate Analysemethoden, 2006, Kapitel 5.

Faktorenanalyse

- Problemstellung der Faktorenanalyse
- Faktorenanalyse: Modell, Schaetzung von Parametern.
- Faktorenanalyse: Hauptkomponentenmethode.
- Faktorenanalyse: Maximum-Likelihood Methode.
- Bestimmung der Anzahl der Faktoren
- Beispiel

Faktorenanalyse

Modell:

$$
\begin{aligned}
Y_1 &= \mu_1 + l_{11}F_1 + \ldots l_{1K}F_K + E_1 \\
&\vdots \\
Y_P &= \mu_P + l_{P1}F_1 + \ldots l_{PK}F_K + E_P \\
Y - \mu &= LF + E \\
Y^T &= (Y_1, \ldots, Y_P) \text{ Zufallsvektor, beobachtbar} \\
\mu &= E(Y) \\
L &= (l_{ij})(pxq)\text{-Matrix der Faktorladungen} \\
F^T &= (F_1, \ldots, F_K) \text{ Faktoren, Zufallsvektor, nicht beobachtbar,} \\
&\quad \text{gemeinsame Faktoren} \\
E^T &= (E_1, \ldots, E_P) \text{ Vektor der spezifischen Faktoren}
\end{aligned}
$$

Faktorenanalyse

Annahmen:

1. $E(F) = 0$
2. $V(E) = \Psi = diag(\Psi_1^2, \ldots, \Psi_P^2)$
3. $Cov(F, E) = 0$

   $V(F) = I$ orthogonales Faktor-Modell

Fundamentalgleichung - orthogonales Faktoren-Modell:

$V(Y) = \Sigma = LL' + \Psi$

$Cov(Y, F) = L$

Rotationsproblem

Nichteindeutigkeit der Loesung:
Die Matrizen L und F sind durch die Fundamentalgleichung nicht eindeutig bestimmt.
Sei T orthogonale Matrix:

$$\begin{aligned} F^* &= T^T F \\ L^* &= LT \\ Y - \mu &= LTT^T F + E = LF \\ \Sigma &= LL^T + \Psi = LTT^T L^T + \Psi = LL^{*T} + \Psi \end{aligned}$$

Kommunalitaetenproblem

:

Waehle Ψ, k, so dass

min rg$(\Sigma - \Psi) = k$

$\Psi > 0$

$\Sigma - \Psi$ positiv definiert

Beachte: Eindeutige Loesung nur fuer k minimal und unter bestimmten Bedingungen.

ML-Faktoren-Analyse

Schaetzung von Parametern:

$$\Sigma = LL^T + \Psi$$

Nebenbedingung: $L\Psi^{-1}L$ diagonal

$$S = \frac{1}{N-1}\sum_{i=1}^{N}(y-\bar{y})(y_i-\bar{y})^T \text{empirische Kovarianz}$$

S besitzt Wishart-Verteilung

$$l(\Sigma;S) = c \cdot |S|^{(N-p-2)/2}|\Sigma|^{-(N-1)/2}$$

$$exp[-((N-1)/2)tr(S\Sigma^{-1})]$$

Man erhaelt ML-Schaetzer von Λ und Ψ durch Maximierung von $l(\Sigma;S)$

$$f(l,\Psi) = ln|\Sigma| + tr(S\Sigma^{-1})$$

Bestimmungsgleichungen

(1) $L = S\hat{\Sigma}^{-1}L \qquad \hat{\Sigma} = \hat{L}\hat{L}^T + \hat{\Psi}$

(2) $diag(\hat{\Sigma}) = diag(S)$

(3) $\hat{L}\hat{\Psi}^{-1}\hat{L}$ diagonal

- Loesung durch Iteratives Verfahren
- Asymptotische Eigenschaften
- $\hat{L}, \hat{\Psi}$ sind konsistent und asymptotisch normalverteilt.

Identifikationsbedingungen

$$
\begin{aligned}
\Sigma &= LL^T + \Psi \\
\Psi^{-1/2}\Sigma\Psi^{-1/2} &= \Sigma^* \\
\Psi^{-1/2}(\Sigma - \Psi)\Psi^{-1/2} &= \Sigma^* - I \\
(\Sigma^* - I) &= \Omega\Delta\Omega^T \\
L &= \Psi^{1/2}\Omega\Delta^{1/2} \\
LL^T &= \Sigma - \Psi \\
\Psi^{-1/2}(\Sigma - \Psi)\Psi^{-1/2} &= (\Psi^{-1/2}L)(\Psi^{-1/2}L)^T
\end{aligned}
$$

Es gilt:

$$L^T\Psi^{-1}L = \Delta^{1/2}\Omega^T\Psi^{1/2}\Psi^{-1}\Psi^{1/2}\Omega\Delta^{1/2} = \Delta$$

Bestimmung der Anzahl der Faktoren

$$\begin{aligned}
H_0: &\quad \Sigma = LL^T + \Psi \qquad \Psi > 0 \\
H_1: &\quad \Sigma \in R^{pp} \qquad \text{positiv definiert} \\
l_0 &= \max_{\Sigma \in H_0} l(\Sigma|S) = l(\hat{L}\hat{L}^T + \hat{\Psi}) \\
l_1 &= l(S) \\
-2ln\frac{l_0}{l_1} &= (N-1)[ln|\hat{\Sigma}| - ln|S| + tr(S\Sigma^{-1}) - p] \\
&= (N-1)\cdot[ln|\hat{\Sigma}| - ln|S|] = E_k \\
&\quad E_k \sim \chi^2(df) \\
df &= p(1+1)|2 - [pk + p - k(k-1)|2]
\end{aligned}$$

df: Zahl der Freiheitsgrade
$k = 0$: Test for Sphericity
$\Sigma = V$=diagonal

Hauptfaktorenanalyse

$$\Sigma = LL^T + \Psi$$

Meist wird die Korrelationsmatrix benutzt
Sei $\hat{\Psi}$ Schaetzung fuer Ψ

$$\begin{aligned} \hat{\Psi} &= diag(1-\hat{h}_1^2, \ldots 1-\hat{h}_p^2) \\ h_j^2 &= 1-\Psi_j^2 \quad \text{Kommunalitaeten} \end{aligned}$$

Betrachte Spektralzerlegung

$$\begin{aligned} (\Sigma - \hat{\Psi}) &= T\Lambda T^T \\ \Lambda &= diag(\lambda_1, \ldots, \lambda_p) \\ \hat{L} &= T\Lambda^{1/2} \\ \hat{L}\hat{L}^T &= T\Lambda^{1/2}\Lambda^{1/2}T^T = (\Sigma - \hat{\Psi}) \end{aligned}$$