

Explorační faktorová analýza

PSY232/PSY532: R101: PRAKTICKÝ ÚVOD PRO
POUŽÍVÁNÍ STATISTICKÉHO PROGRAMU R

28. LISTOPAD 2016 | HYNEK CÍGLER

Východiska

Faktorové metody slouží k zjednodušení struktury dat.

- Vývoj od začátku 20. století; předchůdcem je například metoda triád aj.

Založené na zjednodušeném popisu korelační, resp. variančně-kovarianční matice. CTT paradigma.

Analýza hlavních komponent (PCA, principal component analysis)

- Snížení množství informace v datech.
- Identifikace sdíleného pozorovaného rozptylu napříč položkami v celé matici.
- Deskriptivní model.

Explorační faktorová analýza (EFA)

- Identifikace latentní faktorové struktury.
- Identifikace sdíleného latentního rozptylu předpokládaných faktorů.
- Strukturní model.

Faktorově-analytické modely jsou vlastně zobecněním lineární regrese.

Klíčové pojmy

Faktorový náboj (λ)

- (Standardizovaný) regresní koeficient pro predikci položky na základě faktoru.

Komunalita (h^2)

- Rozptyl položky vysvětlitelný všemi extrahovanými faktory (u ortogonálního řešení suma čtverců všech nábojů dané položky)
- Jedinečnost ($u^2 = 1-h^2$) – jedinečný rozptyl dané položky
- Heywoodův případ nastává, když $h^2 > 1$.

Vysvětlený rozptyl

- Podíl rozptylu položek vysvětlených daných faktorem k celkovému pozorovanému rozptylu.

Vlastní hodnota (eigenvalue)

- Vysvětlený rozptyl vynásobený počtem položek (vysvětlený rozptyl na škále počtu položek).

Odhad počtu faktorů

Konzervativní:

- Scree-plot
- Kaiserovo pravidlo

Empirické

- **Hornova paralelní analýza**, optimální souřadnice
- Velicerovo MAP
- VSS
- **Comparison Data**
- A další

Nejvýhodnější je PA v kombinaci s vizuální interpretací scree-plotu.

Extrakce

Maximum-likelihood

- Klasická iterativní metoda pomocí maximum-likelihood funkce.
- Předpokladem je pozitivně-definitivní korelační matice (determinant > 0).

Ordinary least square (minres)

- Minimalizuje reziduální korelace správným odhadem vysvětleného rozptylu (na diagonále). Nižší nároky na kvalitu korelační matice.

Weighted least square (gls) a Generalized least square (wls)

- Položky jsou různým způsobem vážené oproti minres, což je výhodné v případě položek s výrazně rozdílnými rozptyly.

Principal axis factoring

- Konzervativní metoda založená na dekompozici vlastních hodnot.
- Malé nároky na rozdíl od ML, ale minres je výhodnější.

A řada dalších

Rotace

Nerotované řešení nejde interpretovat. Proto je faktorová matice rotována tak, aby se zvýraznily rozdíly v nábojů položek na jednotlivých faktorech (a řešení šlo interpretovat).

Rotace je záležitostí matice faktorových nábojů (nic jiného není třeba).

Pro každé řešení existuje nekonečně mnoho matematicky stejně dobrých rotací, interpretovatelnost rozhoduje.

Ortogonální rotace

- Faktory spolu nekorelují.

Šikmá rotace

- Faktory spolu korelují.

Shoda modelu s daty

Analogie k fitu v CFA (viz příště).

Ukazatel vhodnosti našeho faktorového řešení.

- Nezávislé na rotaci.

Test dobré shody (co nejmenší poměr χ^2/df).

RMSEA (co nejmenší), TLI (co největší).