

Факторный анализ

Factor analysis

(1-ая часть)

28 мая 2021

Метод снижения размерности

Постановка задачи

Почему показатели сильно скоррелированы? Какой(-ие) концепт(-ы) объясняет(-ют) совместную изменчивость?

Метод снижения размерности

Постановка задачи

Почему показатели сильно скоррелированы? Какой(-ие) концепт(-ы) объясняет(-ют) совместную изменчивость?

На «выходе» задачи

Перейти к признаковому пространству меньшей размерности, которое задается **латентными факторами**.

Модель

Спецификация

$x_j = \alpha_{j1}F_1 + \alpha_{j2}F_2 + \dots + \alpha_{jk}F_K + \epsilon_j$, где x – наблюдаемый показатель, F – латентный фактор, ϵ – ошибка, j – номер наблюдаемого показателя, K – количество факторов.

Решение является стандартизированным.

Модель

Спецификация

$x_j = \alpha_{j1}F_1 + \alpha_{j2}F_2 + \dots + \alpha_{jk}F_K + \epsilon_j$, где x – наблюдаемый показатель, F – латентный фактор, ϵ – ошибка, j – номер наблюдаемого показателя, K – количество факторов. Решение является стандартизированным.

Допущения

- ❶ $E(\epsilon_j) = 0$
- ❷ $cor(F_k, \epsilon_i) = 0$
- ❸ $cor(\epsilon_j, \epsilon_i) = 0$
- ❹ $cor(F_k, F_i) = 0$
- ❺ $E(F_k) = 0; Var(F_k) = 1$

МГК VS Факторный анализ

Основание	МГК	Факторный анализ
Характер показателей	Наблюдаемые ГК	Латентные факторы
Связь X_j и «концептов»	ГК как линейная комбинация	Факторы как объясняющие переменные
Вариация	Сохранение полной информации	Общность и специфичность
Статистическая инференция	Описание	Инференция. ФА как модель

Разведывательный ФА (EFA)

Описательный метод

Задачи:

- 1 описать структуру данных
- 2 определить количество факторов
- 3 понять, что содержательно отражают факторы
- 4 сделать предварительную диагностику, в каком направлении двигаться дальше в уточнении спецификации модели

Разведывательный ФА (ЕФА)

Описательный метод

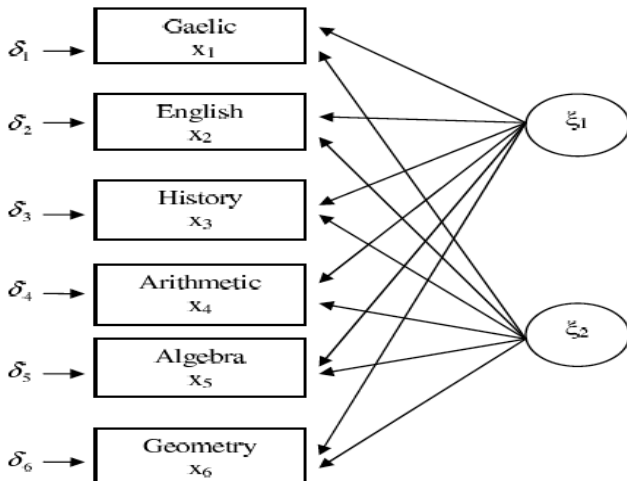
Задачи:

- 1 описать структуру данных
- 2 определить количество факторов
- 3 понять, что содержательно отражают факторы
- 4 сделать предварительную диагностику, в каком направлении двигаться дальше в уточнении спецификации модели

Следующий этап

Тестирование альтернативных спецификаций моделей в рамках конфирматорного (подтверждающего) ФА

Разведывательный ФА: путевая диаграмма



Principal factor method

Разложение вариации в ФА

$Var(X) = h^2 + u^2$, где

h^2 – общность (объясненная факторами вариация),

u^2 – специфичность (необъясненная вариация).

МГК на редуцированной матрице

$$RC_{p,p} = \begin{pmatrix} h_1^2 & \cdots & \text{Cov}(X_1, X_p) \\ \vdots & \ddots & \vdots \\ \text{Cov}(X_p, X_1) & \cdots & h_p^2 \end{pmatrix}, \text{ где}$$

оценка h^2 рассчитывается как квадрат коэффициента множественной корреляции между соответствующим показателем и всеми остальными X .

Этапы ЕФА

Описательный метод

Этапы:

- 1 применим ли факторный анализ?
- 2 определить количество факторов
- 3 оценивание: principal factor, MLE (предположение о нормальности)
- 4 интерпретация и косоугольное вращение факторов, при необходимости
- 5 хорошо ли факторы объясняют совместную изменчивость переменных? Все ли индикаторы нужны?