

Практическая работа №6: Исследование методов многомерного шкалирования

Цель работы

Ознакомиться с методами многомерного шкалирования на основе языка R.

Основные теоретические положения

Многомерное шкалирование (МНШ) можно рассматривать как альтернативу факторному анализу. Целью последнего, вообще говоря, является поиск и интерпретация «латентных (т.е. непосредственно не наблюдаемых) переменных», дающих возможность пользователю объяснить сходства между объектами, заданными точками в исходном пространстве признаков. Для определенности и краткости, далее, как правило, будем говорить лишь о сходствах объектов, имея ввиду, что на практике это могут быть различия, расстояния или степени связи между ними. В факторном анализе сходства между объектами (например, переменными) выражаются с помощью матрицы (таблицы) коэффициентов корреляций. В методе МНШ дополнительно к корреляционным матрицам, в качестве исходных данных можно использовать произвольный тип матрицы сходства объектов. Таким образом, на входе всех алгоритмов МНШ используется матрица, элемент которой на пересечении её i -й строки и j -го столбца, содержит сведения о попарном сходстве анализируемых объектов (объекта i и объекта j). На выходе алгоритма МНШ получаются числовые значения координат, которые приписываются каждому объекту в некоторой новой системе координат (во «вспомогательных шкалах», связанных с латентными переменными, откуда и название МНШ), причем размерность нового пространства признаков существенно меньше размерности исходного (за это собственно и идет борьба).

Логику МНШ можно проиллюстрировать на следующем простом примере. Предположим, что имеется матрица попарных расстояний (т.е. сходства некоторых признаков) между крупными американскими городами. Анализируя матрицу, стремятся расположить точки с координатами городов в двумерном пространстве (на плоскости), максимально сохранив реальные расстояния между ними. Полученное размещение точек на плоскости впоследствии можно использовать в качестве приближенной географической карты США.

В общем случае метод МНШ позволяет таким образом расположить «объекты» (города в данном примере) в пространстве некоторой небольшой размерности (в данном случае она равна двум), чтобы достаточно адекватно воспроизвести наблюдаемые расстояния между ними. В результате можно «измерить» эти расстояния в терминах найденных латентных переменных. Так, в данном примере можно объяснить расстояния в терминах пары географических координат Север/Юг и Восток/Запад.

Постановка задачи

Порядок выполнения работы

Содержание отчёта

From:
<https://se.moevm.info/> - МОЭВМ Вики [se.moevm.info]

Permanent link:
https://se.moevm.info/doku.php/courses:data_analysis_and_interpretation:task6

Last update:

