

УДК 330.4:519.2

ИССЛЕДОВАНИЕ КАЧЕСТВЕННЫХ ФАКТОРОВ РЕГРЕССИОННОЙ МОДЕЛИ

Самаль С. А., д-р экон. наук, профессор
каф. «Общая математика и информатика»;
Жиркевич А. Б., студент 2 курса

Белорусский государственный университет
г. Минск, Беларусь

Используя корреляционный анализ, достаточно часто возникает необходимость разбиения исходных данных на подмножества, для которых коэффициент корреляции (по крайней мере, для одного) выше, чем для исходного массива данных. Целью такого разбиения исследователи называют выделение «скрытых» совокупностей данных, объединенных некоторой внутренней смысловой зависимостью.

Оптимального алгоритма для разделения объектов на отдельные совокупности с целью определения коэффициентов регрессий не существует. Различная природа данных обуславливает необходимость подходов для разбиения объектов. Однако можно построить обобщенный алгоритм, подбирая параметры которого, мы можем обеспечить достаточную гибкость модели. Интерпретировать такой алгоритм можно используя данные, полученные с помощью наблюдений, например, измерение изменений цен в зависимости от времени.

Алгоритм основан на последовательном добавлении новых наблюдений, рассматриваемых с наивысшим приоритетом относительно остальных, с корректированием коэффициентов регрессии для новых данных. При резком, превосходящем «порог», изменении коэффициентов будем считать, что сформирована отдельная совокупность данных, обладающая в отдельности высоким коэффициентом корреляции относительно общего коэффициента.

Вторым этапом в корреляционном анализе называют выделение новых факторов. Так, один из популярных на западе авторов курса эконометрики Лондонской школы экономики К. Доугерти одним из возможных решений этого вопроса видит оценивание единой регрессии с использованием всей совокупности наблюдений и измерение

степени влияния качественного фактора посредством введения фиктивной переменной (переменные-манекены).

В тоже время, не теряет актуальности и практическая возможность оценивания отдельных регрессий (для каждого из факторов) для указанных категорий с последующим определением степени различия полученных коэффициентов. Самым распространённым приёмом введения новых факторов в настоящее время является искажение (преобразование) пространства значений. Примером является подсчет коэффициентов ранговой корреляции Спирмена и Кендала. Данный метод помогает бороться с «шумовыми объектами» («выбросами»).

Идея построения алгоритма имеет под собой объясняемую природу. Наблюдения в экономике близкие во времени обладают единой моделью поведения. А наша задача - найти совокупности данных с единой моделью поведения, с одинаковыми, влияющими на нее факторами, а значит, с высокой корреляцией в данной отдельной группе.

Практический интерес исследования регрессионных зависимостей было и остается в применении различных методов исследования на реальных выборках данных. Только применение на реальных данных может оценить качество и гибкость предложенных алгоритмов.

Разбиение множества исходных данных на подмножества может приводить к росту групповых коэффициентов корреляции (по крайней мере, одного) и целесообразно лишь в случае четкого функционального объяснения экономической сущности и целесообразности выделения дополнительных, и прежде всего, качественных факторов.