

М. П. Базилевский, канд. техн. наук, доцент, Иркутский государственный университет путей сообщения, Иркутск, mik2178@yandex.ru

Программный комплекс построения линейно-мультипликативных регрессий

Статья посвящена проблеме спецификации регрессионных моделей, а именно, вопросу отбора наиболее информативных мультипликаций в линейно-мультипликативных регрессиях. Эта задача может быть формализована в виде задачи частично-булева линейного программирования. Поскольку решение таких задач требует привлечения вычислительных способностей современных компьютеров, то для построения линейно-мультипликативных регрессий был разработан универсальный программный комплекс, который может быть использован в технических исследованиях, экономике, бизнесе, социологии, медицине и др. Для демонстрации его работы решена задача моделирования объема пропуска крупнотоннажных контейнеров на железнодорожном пункте пропуска Забайкальск-Маньчжурия. При этом проведено тестирование скорости решения подобных вычислительных задач.

Ключевые слова: линейно-мультипликативная регрессия, метод наименьших модулей, частично-булево линейное программирование, программный комплекс, объемы пропуска крупнотоннажных контейнеров, железнодорожный пункт пропуска Забайкальск-Маньчжурия.

Введение

В результате развития информационных технологий к настоящему времени в различных областях человеческой деятельности, например, в технических исследованиях, экономике, бизнесе, социологии, медицине и т. д., накоплено огромное количество больших массивов статистических данных, растущее от года к году по экспоненциальному закону. Поэтому весьма актуальной на сегодняшний день является задача обработки и анализа таких данных с целью извлечения из них неизвестных ранее полезных знаний. Термин «анализ данных» употребляется наряду с такими понятиями, как «machine learning», «databasemining», «datamining», «knowledgediscoveryin databases», «intelligent data analysis» и др. При этом практически не существует четкой границы, позволяющей отличать эти термины между собой, что естественно на начальном этапе становления новой научной дисциплины.

Одним из основных инструментов анализа данных является регрессионный анализ. При построении регрессионных моделей ключевым этапом является этап спецификации модели, предполагающий выбор состава и математической формы связи между переменными в уравнении регрессии. От того, насколько удачно будет решена эта проблема, решающим образом зависит успех всего регрессионного исследования. На сегодняшний день не существует системы стандартных рекомендаций и методов, которые образовывали бы строгую теоретическую базу для выбора спецификации модели. При этом разработан значительный арсенал ставших уже классическими способов описания взаимосвязей между моделируемыми показателями (см., например, [1–3]). Целью данной работы является разработка и тестирование программного комплекса построения линейно-мультипликативных регрессионных моделей, подробно рассмотренных в работах [4, 5].