

С.И. Носков¹

¹ Иркутский государственный университет путей сообщения, г. Иркутск, Российская Федерация

ПОСТАНОВКА ЗАДАЧИ ОЦЕНИВАНИЯ ПАРАМЕТРОВ РЕГРЕССИИ ПОСРЕДСТВОМ МИНИМИЗАЦИИ ЗАДАННОЙ НА ГРУППАХ НАБЛЮДЕНИЙ ВЕКТОРНОЙ ФУНКЦИИ ПОТЕРЬ

Аннотация. В работе сформулирована задача оценивания параметров регрессионного уравнения общего вида посредством минимизации векторной функции потерь, компонентами которой являются частные и различающиеся между собой функции потерь, заданные на различных участках обрабатываемой выборки данных.

Ключевые слова: регрессионная модель, методы оценивания параметров, векторная функция потерь, подвыборка, L_v -оценки, множество Парето.

S.I. Noskov¹

¹ Irkutsk state University of railway engineering, Russian Federation.

FORMULATION OF THE PROBLEM OF ESTIMATING REGRESSION PARAMETERS BY MINIMIZING THE VECTOR LOSS FUNCTION SPECIFIED ON OBSERVATION GROUPS

Abstract. The paper formulates the problem of estimating the parameters of a regression equation of a general form by minimizing the vector loss function, the components of which are particular and differing loss functions given in different parts of the processed data sample.

Keywords: regression model, parameter estimation methods, vector loss function, subsample, L_v -estimates, Pareto set.

Рассмотрим регрессионное уравнение общего вида:

$$y_k = F(\alpha; x_{k1}, \dots, x_{km}) + \varepsilon_k, \quad k = \overline{1, n}, \quad (1)$$

где y – зависимая, а $x_i, i = \overline{1, m}$ – независимые переменные, α – вектор идентифицируемых параметров, F – аппроксимирующая вещественная функция, ε_k – ошибки аппроксимации, n – число наблюдений (длина выборки). При этом все переменные и ошибки аппроксимации в модели (1) – детерминированные величины.

В рамках принятого в регрессионном анализе подхода вектор параметров α определяется посредством минимизации заданной функции потерь. Широкий класс таких функций имеет вид [1]:

$$I_v(\alpha) = \sum_{k=1}^n |\varepsilon_k|^v, \quad v \geq 1. \quad (2)$$

Пусть, исходя из некоторых формальных и (или) содержательных соображений исследователь, учитывая характер моделируемого объекта или процесса, может разделить исходную выборку на s непересекающихся подвыборок (групп номеров наблюдений) $P_i, i = \overline{1, s}$:

$$\bigcup_{i=1}^s P_i = \{1, 2, \dots, n\}, \quad P_i \cap P_j = \emptyset, \quad i \neq j.$$

Зададим на каждой из них свою функцию потерь следующим образом:

$$I_{v_i}(\alpha) = \sum_{k \in P_i} |\varepsilon_k|^{v_i}, \quad v_i \geq 1, \quad i = \overline{1, s},$$

потребовав, таким образом, ее минимизации исключительно на соответствующей подвыборке исходной выборки.

Поставим задачу поиска вектора неизвестных параметров α регрессии (1) посредством минимизации векторной функции потерь $I(\alpha)$:

$$I(\alpha) = (I_{v_1}(\alpha), I_{v_2}(\alpha), \dots, I_{v_s}(\alpha)) \rightarrow \min. \quad (3)$$

В регрессионном анализе известны постановки задач (см., например, [2-7]) с векторными функциями потерь, каждая из двух компонент которых предполагает минимизацию заданных *на всей выборке* функций $I_\nu(\alpha)$ со значениями $\nu=1$ (или $\nu=2$) и $\nu \rightarrow \infty$.

Вариантами решения задачи (3) могут быть либо множество паретовских оценок (или какое-либо его подмножество), либо некоторые точечные оценки, наделенные определенными специальными свойствами. В случае, когда функция F линейна, а $\nu_i=1$, $i = 1, s$, задача (3) сводится к многокритериальной задаче линейного программирования, для решения которой в [7] представлены соответствующие способы. Для исследования задачи (3) в различных ее трактовках могут быть успешно использованы приемы, подробно описанные в работах [2-7].

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Демиденко Е.З. Линейная и нелинейная регрессии.- М.: Финансы и статистика, 1981. 302с.
2. Каменев Г.К. Многокритериальный метод множеств идентификации // Ж. вычисл.матем. и матем. физ.- 2016.- т.56.- №11.- С.1872-1888.
3. Каменев Г.К. Многокритериальный метод идентификации и прогнозирования // Математическое моделирование. – 2017.- №8.- С.29-43.
4. Носков С.И. Компромиссные паретовские оценки параметров линейной регрессии // Математическое моделирование. - 2020. -т.32.- №11.- С. 70–78.
5. Баенхаева А.В., Базилевский М.П., Носков С.И. Программный комплекс множественного оценивания регрессионных моделей // Информационные технологии и проблемы математического моделирования сложных систем.- 2016.- №17.- С.38–44.
6. Носков С.И. L–множество в многокритериальной задаче оценивания параметров регрессионных уравнений // Информационные технологии и проблемы математического моделирования сложных систем.- 2004.- №1.- С.64-71.
7. Носков С.И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облформпечать.- 1996.- 320 с.

REFERENCES

1. Demidenko E.Z. Linear and non-linear regression. - M.: Finance and statistics, 1981. 302p.
2. Kamenev G.K. Multicriteria method of identification sets // Zh. Vychisl.mat. and mat. physics.- 2016.- v.56.- №11.- P.1872-1888.
3. Kamenev G.K. Multi-criteria method of identification and forecasting // Mathematical Modeling. – 2017.- No. 8.- P.29-43.
4. Noskov S.I. Compromise Pareto estimates of linear regression parameters // Mathematical Modeling. - 2020. - v.32.- No. 11.- P. 70–78.
5. Baenkhayeva A.V., Bazilevsky M.P., Noskov S.I. Software complex for multiple estimation of regression models // Information technologies and problems of mathematical modeling of complex systems. - 2016. - No. 17. - P. 38–44.
6. Noskov S.I. L-set in the multicriteria problem of estimating the parameters of regression equations // Information technologies and problems of mathematical modeling of complex systems.- 2004.- №1.- P.64-71.
7. Noskov S.I. Technology for modeling objects with unstable operation and uncertainty in data. - Irkutsk: Oblinformpress. - 1996. - 320 p.

Информация об авторе

Сергей Иванович Носков – д.т.н., профессор, профессор кафедры «Информационные системы и защита информации», Иркутский государственный университет путей сообщения, г. Иркутск, e-mail: sergey.noskov.57@mail.ru

Author

Sergey Ivanovich Noskov, Doctor of Technical Science, Professor, the Subdepartment Information systems and information security, Irkutsk State Transport University, Irkutsk, e-mail: sergey.noskov.57@mail.ru

Для цитирования

Носков С.И. Постановка задачи оценивания параметров регрессии посредством минимизации заданной на группах наблюдений векторной функции потерь // «Информационные технологии и математическое моделирование в управлении сложными системами»: электрон. науч. журн. – 2022. – №3(15). – С.58-60– DOI: 10.26731/2658-3704.2022.3(15).58-60 – Режим доступа: <http://ismm-irgups.ru/toma/315-2022>, свободный. – Загл. с экрана. – Яз. рус., англ. (дата обращения: 15.10.2022)

For citation

Noskov S.I. Formulation of the problem of estimating regression parameters by minimizing the vector loss function specified on observation groups // *Informacionnye tehnologii i matematicheskoe modelirovanie v upravlenii slozhnymi sistemami: ehlektronnyj nauchnyj zhurnal* [Information technology and mathematical modeling in the management of complex systems: electronic scientific journal], 2022. No. 3(15). P. 58-60. DOI: 10.26731/2658-3704.2022.3(15).58-60 [Accessed 15/10/22]