ПОДХОД К ФОРМАЛИЗАЦИИ ВЛОЖЕННОЙ КУСОЧНО-ЛИНЕЙНОЙ РЕГРЕССИИ

С.И. Носков, *д-р техн. наук, профессор*
Иркутский государственный университет путей сообщения
(Россия, г. Иркутск)

DOI:10.24412/2500-1000-2023-1-2-218-220

Аннотация. *В работе предложен способ формализации вложенной кусочно-линейной регрессии двух типов, являющихся комбинациями традиционной кусочно-линейной регрессии и функции риска. Оценивание параметров вложенной кусочно-линейной регрессии может быть сведено к решению задачи линейно-булевого программирования.*

Ключевые слова: *производственная функция, кусочно-линейная регрессия, функция риска, вложенная кусочно-линейная регрессия, оценивание параметров, метод наименьших модулей.*

При анализе сложных объектов различных объектов широко применяются методы регрессионного анализа (см., например, [1, 2]). Особенно они востребованы в экономике, где с их помощью разрабатываются эффективные эконометрические модели.

Пусть при анализе некоторого объекта исследователь, исходя из соображений содержательного и (или) формального характера, считает, что поведение зависимой переменной y определяется значениями независимых переменных x_1, x_2,…, x_m, то есть предполагает наличие регрессионной зависимости y от x_1, x_2,…, x_m:

$$y_k = F(a; x_{k1}, x_{k2}, \dots, x_{km}) + \varepsilon_k, k = \overline{1,n}, \quad (1)$$

где k – номер наблюдения, n – их количество, a – вектор оцениваемых параметров, ε_k – ошибки аппроксимации, F – аппроксимирующая функция.

Для случая, когда переменная y представляет собой выпуск продукции, а x_1,x_2,…,x_m - ресурсные факторы, в рамках математической экономики разработаны различные типы производственных функций. К числу наиболее известных, наряду с простейшей, линейной, относятся, в частности, следующие [3-5].

1. Функция Кобба-Дугласа

$$y = a_0 \prod_{i=1}^{m} x_i^{a_i} + \varepsilon, a_i > 0, i = \overline{1,m}.$$

2. Функция с постоянной эластичностью замещения

$$y = \left(\sum_{i=1}^{m} a_i x_i^{-\rho}\right)^{-\gamma/\rho} + \varepsilon.$$

3. Многорежимная функция с различными параметрами крутизны

$$y = \left(\sum_{i=1}^{m} a_i x_i^{-\rho_1}\right)^{-\gamma_1/\rho_1} \left(\sum_{i=1}^{m} \beta_i x_i^{-\rho_2}\right)^{-\gamma_2/\rho_2} + \varepsilon.$$

4. Функция Солоу

$$y = \left(\sum_{i=1}^{m} a_i x_i^{\beta_i}\right)^{\gamma} + \varepsilon.$$

5. Функция Аллена

$y = \sum_{i>j} a_{ij} x_i x_j - \sum_{s=1}^{m} \beta_s x_s^2 + \varepsilon.$

6. Функция Сато

$y = a_0 \prod_{i=1}^{m} x_i^{a_i} \left(\sum_{j=1}^{m} \beta_j x_j^{-\rho}\right)^{-\gamma/\rho} + \varepsilon.$

7. Функция Кокса-Бокса

$\frac{y^{\lambda}-1}{\lambda} = a_0 + \sum_{i=1}^{m} a_i \frac{x_i^{\lambda}-1}{\lambda} + \varepsilon.$

В работе [4] рассмотрена возможность комбинирования приведенных выше классических производственных функций посредством использования, в частности, мультипликативных (а, как следствие, и аддитивных тоже) конструкций вида:

$$F(\cdot) = F_1(\cdot) F_2(\cdot). \quad (2)$$

Несколько менее часто по сравнению с перечисленными выше производственными функциями в эконометрике используется производственная функция с нулевой эластичностью замены ресурсов, или с постоянными пропорциями [3,4,6], в математическом отношении представляющая собой кусочно-линейную регрессию:

$$y_k = min\{a_1 x_{k1}, a_2 x_{k2}, \dots, a_m x_{km}\} + \varepsilon_k. \quad (3)$$

Построение функции (3) целесообразно тогда, когда объем выпуска продукции в исследуемой системе определяется «узким местом», а именно, количеством ресурса, обеспечивающего лишь наименьший возможный выпуск. При этом любое увеличение количества других потребляемых ресурсов не может компенсировать дефицит лимитирующего фактора. В работе [6] описаны особенности оценивания параметров регрессии (3) методом наименьших модулей [7], сводящимся к решению задачи линейно-булевого программирования.

В работе [8] предложена противоположная по смыслу модели (2) функция риска

$$y_k = max\{a_1 x_{k1}, a_2 x_{k2}, \dots, a_m x_{km}\} + \varepsilon_k. \quad (4)$$

Здесь зависимая переменная имеет негативный по отношению к объекту характер, например, риск, уязвимость, угроза и т.д., а независимые факторы являются частными индикаторами этого агрегирующего показателя.

Возможным способом комбинирования (обобщения) кусочно-линейных регрессий (3) и (4), наряду с формой (2), является введение двух типов **вложенных кусочно-линейных регрессий** следующим образом.

1. Вложенная кусочно-линейная регрессия первого типа:

$$y_k = min\{ min_{i\in I^1}\{ a_i^1 x_{ki}\}, \dots, min_{i\in I^G}\{ a_i^G x_{ki}\}, max_{i\in J^1}\{ \beta_i^1 x_{ki}\}, \dots, max_{i\in J^H}\{ \beta_i^H x_{ki}\}\} + \varepsilon_k. \quad (5)$$

2. Вложенная кусочно-линейная регрессия второго типа:

$$y_k = max\{ min_{i\in I^1}\{ a_i^1 x_{ki}\}, \dots, min_{i\in I^G}\{ a_i^G x_{ki}\}, max_{i\in J^1}\{ \beta_i^1 x_{ki}\}, \dots, max_{i\in J^H}\{ \beta_i^H x_{ki}\}\} + \varepsilon_k. \quad (6)$$

Здесь индексные множества $I^i, i = \overline{1,G}$, $J^j, j = \overline{1,H}$ являются подмножествами множества $\{1,2,\dots,m\}$ и могут иметь непустые попарные пересечения.

Способ идентификации параметров вложенных кусочно-линейных регрессий (5) и (6) может быть аналогичен тому, который используется при оценивании параметров регрессий (3) и (4), и сведен, таким образом, к решению специальным образом сформированных задач линейно-булевого программирования.

Отметим, что в формулах (5) и (6) отражен первый порядок вложенности. Он может также вторым, третьим и т.д.

Автор намерен заняться дальнейшим исследованием вложенных кусочно-линейных регрессий.

Библиографический список

1. Носков С.И., Кириллова Т.К. Регрессионная модель оценки влияния рекреационной деятельности на социально-экономическое развитие территории // Вестник Иркутского государственного технического университета. – 2013. – № 9 (80). – С. 24-28.
2. Носков С.И., Оленцевич В.А., Базилевский М.П. Математическая модель оценки безопасности перевозочного процесса на региональном уровне // Транспортная инфраструктура Сибирского региона. – 2014. – Т. 1. – С. 537-542.
3. Клейнер Г.Б. Производственные функции. – М.: Финансы и статистика, 1986. – 239 с.
4. Носков С. И. Технология моделирования объектов с нестабильным функционированием и неопределенностью в данных. – Иркутск: Облинформпечать, 1996. – 320 с.
5. Минько Э.В., Минько А.Э. Оптимальное управление коммерческими проектами. – Саратов, 2017.
6. Носков С.И., Хоняков А.А. Программный комплекс построения некоторых типов кусочно-линейных регрессий // Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – № 3 (4). – С. 47-55.
7. Носков С. И. О методе смешанного оценивания параметров линейной регрессии // Информационные технологии и математическое моделирование в управлении сложными системами. – 2019. – № 1. – С. 41-45.
8. Носков С.И. Идентификация параметров кусочно-линейной функции риска // Транспортная инфраструктура Сибирского региона. – 2017. – Т. 1. – С. 417-421.

APPROACH TO FORMALIZING NESTED PIECE-LINEAR REGRESSION

S.I. Noskov, *Professor*
Irkutsk State Transport University
(Russia, Irkutsk)

Abstract. *The paper proposes a method for formalizing nested piecewise linear regression, which is a combination of traditional piecewise linear regression and a risk function. Estimating the parameters of a nested piecewise linear regression can be reduced to solving a linear Boolean programming problem.*

Keywords: *production function, piecewise linear regression, risk function, nested piecewise linear regression, parameter estimation, least absolute deviation method.*