Долгих В. Н. О некоторых некорректных рекомендациях по определению параметров нелинейных регрессий / В. Н. Долгих // Проблеми і перспективи розвитку банківської системи України : зб. тез доповідей Х науково-практичної конференції / Державний вищий навчальний заклад "Українська академія банківської справи Національного банку України". – Суми, 2007. –Т. 1. - С. 55-57.

УДК 330.43

О некоторых некорректных рекомендациях по определению параметров нелинейных регрессий

Долгих Владимир Николаевич, к. ф.-м. н., доц. (*Украинская академия банковского дела Национального банка Украины*)

Для описания и прогнозирования процессов в экономике часто применяют регрессионные модели. Однофакторное уравнение регрессии, связывающее условное среднее $\hat{y}$ объясняемой переменной *y* с объясняющей переменной *x*, задаётся в виде: $\hat{y} = f(x, b_0, b_1, ..., b_m)$, где $b_0, \ldots, b_m$ – параметры, подлежащие определению в результате обработки статистической выборки, состоящей из *n* пар чисел (x_i, y_i) (i=1,2,…,n). Обычно оценки параметров b_j получают методом наименьших квадратов (МНК) минимизируя сумму квадратов отклонений фактических значений y_i от вычисленных по уравнению регрессии $\hat{y}_i = f(x_i, b_0, b_1, ..., b_m)$:

$$\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}\left[y_i - f(x_i, b_0, b_1, ..., b_m)\right]^2 = Q(b_0, b_1, ..., b_m) \to \min \qquad (1)$$

Приравнивая нулю частные производные функции Q по параметрам b_j, получаем систему уравнений для их определения. Система будет линейной, если функция *f* линейна относительно параметров. Например, для функции $\hat{y} = a + bx$ имеем:

$$\begin{cases} na + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i \\ a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} x_i y_i \end{cases} \Rightarrow b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2}, \quad a = \frac{1}{n}\left(\sum_{i=1}^{n} y_i - b\sum_{i=1}^{n} x_i\right). \qquad (2)$$

Если же функция регрессии нелинейна относительно параметров, например показательная $\hat{y} = ae^{bx}$ или степенная $\hat{y} = ax^b$, то и система уравнений для их определения также нелинейна и её решение представляет значительные трудности. Нелинейными по параметрам функциями, применяемыми в экономике, являются показательная и степенная производственные функции, функция Кобба-Дугласа, функции Торнквиста, *s*-образные логистическая кривая, кривая Джонсона и др.

Для того, чтобы избежать необходимости решения нелинейных систем, обычно рекомендуют преобразовать исходную функцию регрессии так, чтобы она стала линейной относительно параметров [1-3]. Показательное и степенное уравнения рекомендуют логарифмировать, например показательное уравнение после логарифмирования примет вид: $\hat{y}_* = a_* + bx$, где $\hat{y}_* = \ln \hat{y}$, $a_* = \ln a$. Параметры a_*, b вычисляют по формулам (2), заменив y_i на $\ln y_i$. Функцию $\hat{y} = 1/[a + bx]$ преобразуют к виду: $\hat{y}_* = a + bx$, где $\hat{y}_* = 1/\hat{y}$ и определяют параметры a, b по формулам (2), заменив y_i на $1/y_i$. Однако, при замене в формулах (2) y_i на $\ln y_i$ минимизируется не сумма (1), а сумма $\sum(\ln y_i - \ln \hat{y}_i)^2$. При замене y_i на $1/y_i$ минимизируется сумма $\sum(1/y_i - 1/\hat{y}_i)^2$. В результате искажается принятая в МНК мера близости функции регрессии к экспериментальным точкам, а оценки параметров получаются смещёнными. Рассмотрим примеры. Для каждого значения $x_i = 0{,}4i$ (i= 0,1,…,5) расположим по две “экспериментальные” точки на равных расстояниях $e_i = \pm\ 0{,}9$ от “теоретической” кривой $y=e^x$. Параметры показательного уравнения $\hat{y} = ae^{bx}$, полученные в результате логарифмического преобразования, a=0,5805, b=1,3424 значительно отличаются от “теоретических” a=b=1. Соответствующие кривые приведены на рис. 1. Этот же результат получается и при построении экспоненциальной линии тренда в программе Excel. Аналогичные расчёты для “теоретической” функции $y = 1/(a + bx)$ при a=b=1 дают a = 1,632, b = -1,639. Полученная функция имеет разрыв 2-го рода в точке x≈0,996 (“теретическая” функция разрывна в точке x = -1).

Для получения несмещённых оценок предлагается использовать нелинейный метод наименьших квадратов (НМНК), заключающийся в непосредственной минимизации суммы (1), например при помощи средства *Сервис* ⇒ *Поиск решения* программы Excel. В результате минимизации показательной функции получены оценки: a=1,000003, b=0,999998.

Предлагаемый подход пригоден и для оценки параметров других уравнений регрессий.

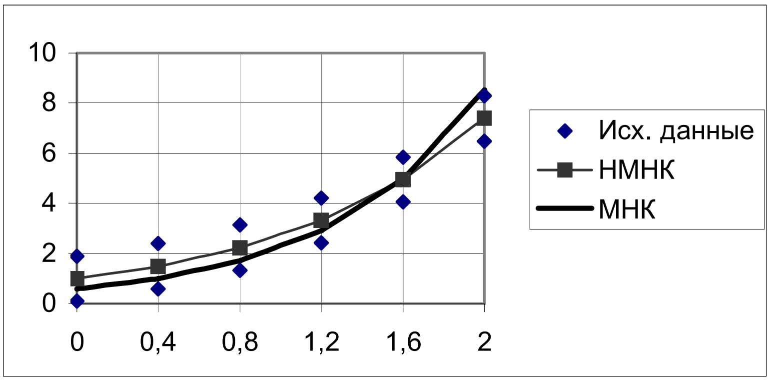

Рис. 1. Экспоненциальные кривые, найденные по МНК и НМНК

СПИСОК ЛИТЕРАТУРЫ

1. Лук’яненко І. Г., Краснікова Л. І. Економетрика: Підручник. – К.: Тов. “Знання”, КОО, 1998. –494 с.
2. Наконечний С. І., Терещенко Т. О., Романюк Т. П. Економетрія: Підручник. – Вид. 2-ге. – К.: КНЕУ 2000. – 296 с.
3. Толбатов Ю. А. Економетрика: Учбовий посібник. – К., Четверта хвиля, 1997. – 320 с.