UDC 519.233

Asymptotic Explicit Optimal Estimators of an Unknown Parameter in One Power Regression Problem

E. N. Savinkina*, A. I. Sakhanenko†

* *Laboratory of Probability Theory and Mathematical Statistics, Sobolev Institute of Mathematics of SB RAS, Koptyug avenue 4, Novosibirsk, 630090, Russia*

† *Laboratory of Applied Probability, Novosibirsk State University, Pirogova str. 2, Novosibirsk, 630090, Russia*

Abstract. In this work, we consider the task of estimating an unknown parameter in a special nonlinear regression problem. Said problem is a standard example of a nonlinear regression where finding classical ordinary least squares estimators meets considerable computational difficulties. We construct the explicit second-stage estimator of the unknown parameter and also give some examples of the explicit first-stage estimators. Finally, we prove both first-stage and second-stage estimators to be asymptotically normal under wide assumptions and show that under some natural assumptions these estimators have the same precision as the OLS estimator.

Keywords: unknown parameter, explicit asymptotically normal estimator, nonlinear regression.

1. Introduction

Assume we observe random variables $\{Y_i\}$ which can be represented in the following form:

$$Y_i = \sqrt{1+\alpha x_i} + \varepsilon_i, \quad i = 1, 2, ..., \tag{1}$$

where $\{x_i > 0\}$ is a known numerical sequence and $\{\varepsilon_i\}$ are some unobserved random errors, which we assume to be identically distributed, and

$$\mathbb{E}\varepsilon_1 = 0, \quad \mathbb{E}\varepsilon_1^4 < \infty, \quad \sigma^2 = \mathbb{D}\varepsilon_1 > 0. \tag{2}$$

Our aim is to estimate an unknown positive parameter α.

The traditional way of solving this problem is to use the least squares method, according to which we find an estimator for α as such $\tilde{\alpha}_n$ that minimizes the sum of the squares of the residuals, i.e.

$$\tilde{\alpha}_n = \arg\min_{\alpha} H_n(\alpha), \quad \text{where} \quad H_n(\alpha) = \sum_{i \le n} (Y_i - \sqrt{1+\alpha x_i})^2. \tag{3}$$

However, as noted in the monograph by E.Z. Demidenko [1], the task of finding the minimum point $\tilde{\alpha}_n$ in (1) presents serious computational difficulties (due to the fact that the number of local minima in the right-hand side of (1) can tend to infinity).

Equation (1) is the third in series of nonlinear regression equations, for which A.I. Sakhanenko and his students have found explicit asymptotically normal estimators of the unknown parameters (see [2]- [5]).

2. Main section

In [2] we introduced a whole class of simple explicit estimators of the parameter α (see the details in Example 1 below). In particular, we can take the following sample statistic as our estimator:

$$\alpha_n^* = \sum_{i\le n}(x_i - \bar{x}_n)Y_i^2 \Big/ \sum_{i\le n}(x_i - \bar{x}_n)^2, \quad \text{where} \quad \bar{x}_n = \frac{1}{n}\sum_{i\le n} x_i. \tag{4}$$

However, all the estimators in [2] have one common drawback: they are less precise than the OLS estimator (1). To solve this problem it seems natural to use Ronald Fisher's idea [6] for improving these estimators.

Suppose we have already obtained a preliminary estimator α_n^*. To find the second-stage estimator Newton's method is typically used:

$$\tilde{\alpha}_n^* = \alpha_n^* - \frac{H_n'(\alpha_n^*)}{H_n''(\alpha_n^*)}, \tag{5}$$

where

$$H_n'(\alpha) = \sum_{i\le n}\left(x_i - \frac{Y_i x_i}{\sqrt{1+\alpha x_i}}\right), \quad H_n^{''}(\alpha) = \sum_{i\le n}\frac{Y_i x_i^2}{2\sqrt{(1+\alpha x_i)^3}}.$$

But it turns out that the following estimator (2) is better in comparison with the estimator (2):

$$\alpha_n^{**} = \alpha_n^* - \sum_{i\le n}\Big(x_i - \frac{Y_i x_i}{\sqrt{1+\alpha_n^* x_i}}\Big)\Big/\Big(\frac{3}{4}\sum_{i\le n}\frac{x_i^2}{1+\alpha_n^* x_i} - \frac{1}{4}\sum_{i\le n}\frac{x_i^2 Y_i}{\sqrt{(1+\alpha_n^* x_i)^3}}\Big). \tag{6}$$

Now we can present our main statement.

Theorem. *Let the following conditions hold:*

$$|\alpha_n^* - \alpha|^3 \sum_{i\le n}\frac{x_i^4}{(1+\alpha x_i)^3}\Big/\sqrt{\sum_{i\le n}\frac{x_i^2}{1+\alpha x_i}} \xrightarrow{p} 0, \tag{7}$$

$$\frac{\max_{i\le n} x_i^2/(1+\alpha x_i)}{\sum_{i\le n} x_i^2/(1+\alpha x_i)} \to 0. \tag{8}$$

Then statistic α_n^{**} is a consistent and asymptotically normal estimator of the parameter α, i.e.

$$\frac{\alpha_n^{**}-\alpha}{d_n} \Rightarrow \mathcal{N}(0,1), \quad \textit{where} \quad d_n = 2\sigma \Big/ \sqrt{\sum_{i\le n} \frac{x_i^2}{1+\alpha x_i}}. \tag{9}$$

We emphasize that in (2), (2), (2) and further all limits are taken when $n \to \infty$.

Remark. It can be shown that if the OLS estimator $\tilde{\alpha}_n$ is asymptotically normal, then

$$(\tilde{\alpha}_n - \alpha)/d_n \Rightarrow \mathcal{N}(0,1)$$

with d_n defined in (2).

Example 1. When $n \ge n_1 = \min\{i : x_i \ne x_1\}$ we can always find constants $c_{n,1}, ..., c_{n,n}$ such that

$$\sum_{i\le n} c_{n,i} = 0 \quad \text{and} \quad \sum_{i\le n} c_{n,i} x_i \ne 0.$$

It poses no difficulties to prove that in this case statistic

$$\alpha_n^* = \sum_{i\le n} c_{n,i} Y_i^2 \Big/ \sum_{i\le n} c_{n,i} x_i \tag{10}$$

appears to be an unbiased estimator of an unknown parameter α. Moreover, if (1) holds then

$$\mathbb{E}\alpha_n^* = \alpha, \quad \mathbb{D}\alpha_n^* = \sum_{i\le n} c_{n,i}^2 \mathbb{D} Y_i^2 \Big/ (\sum_{i\le n} c_{n,i} x_i)^2.$$

It is easy to make sure that in this case the following simple condition

$$\mathbb{D}\alpha_n^* \left(\sum_{i\le n} \frac{x_i^2}{1+\alpha x_i} \right)^{1/3} \to 0$$

is sufficient for the validity of assumption (2) in the Theorem.

Example 2. Let us consider the simplest situation when $c_{n,i} = x_i - \bar{x}_n$. In this case the estimator (2) turns into the estimator (2).

Now assume in addition that

$$\sup_n x_n < \infty, \quad n^{-1/3} \sum_{i\le n} (x_i - \bar{x}_n)^2 \to \infty. \tag{11}$$

One can make sure that when (2) holds, the estimator α_n^* (see (2)) meets all the conditions in the Theorem.

We shall emphasize that for (2) to be true the validity of the following natural assumptions is enough:

$$\sup_n x_n < \infty, \quad \lim_{n\to\infty} \inf_n n^{-1} \sum_{i\le n} (x_i - \bar{x}_n)^2 > 0.$$

3. Conclusions

We have constructed a new explicit asymptotically normal estimator of an unknown parameter in one power regression problem. Moreover we have shown its precision to be as good as that of OLS estimator.

Acknowledgments

The work is partially supported by RFBR grant No 15-01-07460a.

References

1. *Demidenko E. Z.* Optimizacija i regressija [Optimization and Regression]. — Moscow: Nauka, 1989.
2. *Savinkina E. N., Sakhanenko A. I.* Explicit estimators of an unknown parameter in a power regression problem // Siberian Electronic Mathematical Reports. — 2014. — Vol. 11. — P. 725–733.
3. *Linke Yu. Yu., Sakhanenko A. I.* Asymptotically normal estimation of a parameter in a linear-fractional regression problem // Siberian Mathematical Journal. — 2000. — Vol. 41, no. 1. — P. 150–163.
4. *Yermolenko K. V., Sakhanenko A. I.* Explicit asymptotically normal estimators of an unknown parameter in a partially linear regression problem // Siberian Electronic Mathematical Reports. — 2013. — Vol. 10. — P. 719–726.
5. *Kalenchuk A. A., Sakhanenko A. I.* The existence of explicit asymptotically normal estimators of an unknown parameter in a logarithmic regression problem // Siberian Electronic Mathematical Reports. — 2015. — Vol. 12. — P. 874–883.
6. *Fisher R. A.* Theory of statistical estimation // Proc. Camb. Phil. Soc. — 1925. — Vol. 22. — P. 700–725.