Improvement and analysis of statistical estimators of unknown parameter in one class of power regression problems

Ekaterina Savinkina

Applied Probability Workshop, Novosibirsk

August 27, 2020

Problem statement

Suppose we observe random variables $\{Y_i\}$ represented in a way

$$Y_i = (1 + \alpha X_i)^r + \varepsilon_i, i = 1, 2, ..., \tag{1}$$

where $\{X_i > 0\}$ is some known numerical sequence, $r \neq 0$ — a known number and $\{\varepsilon_i\}$ — unobservable random errors[1].

[1] *E. N. Savinkina, A. I. Sakhanenko* Improvement of statistical estimates in one power regression problem // Siberian Electronic Mathematical Reports. – 2019. – Vol. 16. – 1901-1912.

Problem statement

Suppose we observe random variables $\{Y_i\}$ represented in a way

$$Y_i = (1 + \alpha X_i)^r + \varepsilon_i, i = 1, 2, ..., \tag{1}$$

where $\{X_i > 0\}$ is some known numerical sequence, $r \neq 0$ — a known number and $\{\varepsilon_i\}$ — unobservable random errors.

Our aim is to estimate an unknown parameter $\alpha > 0$ using the first n observations.

Improved estimator

Let α_n^* be the previously obtained consistent estimate of α. To improve it we propose using

$$\alpha_n^{**}(c) = \alpha_n^* - \frac{S'_{n,r}(\alpha_n^*)}{cS''_{n,r}(\alpha_n^*) + (1-c)2r^2 E_n^2(\alpha_n^*)}, \qquad (2)$$

where c is a constant chosen by the statistician,

$$S_{n,r}(\alpha) = \sum_{i \le n} (Y_i - (1 + \alpha X_i)^r)^2$$

and

$$E_n^2(t) := \sum_{i \le n} \frac{X_i^2}{(1 + tX_i)^{2-2r}}.$$

Preliminary remarks

Notation:

- $A_{p,n}(t) := \sum_{i \le n} \frac{X_i^p}{(1+tX_i)^{p-2r}}$, $p > 0$
- $\rho_{1,n} := (\alpha_n^* - \alpha)^2 A_{3,n}(\alpha) / E_n(\alpha)$
- $\rho_{2,n} := |\alpha_n^* - \alpha|^3 A_{4,n}(\alpha) / E_n(\alpha)$

Assumption 1

Random errors ε_i are i.i.d. with

$$\mathbb{E}\varepsilon_1 = 0,\ 0 < \mathrm{Var}(\varepsilon_1) = \sigma^2 < \infty.$$

Assumption 2

For all X_i holds

$$\frac{\max_{i \le n} X_i^2 / (1 + \alpha X_i)^{2-2r}}{E_n^2(\alpha)} \to 0.$$

Main results

Theorem 1

Let $\mathrm{Assumptions}\,1\&2$ hold. If $\rho_{1,n} \xrightarrow{p} 0$ and α_n^* is a consistent estimate of α, then for all c the following convergence takes place:

$$\frac{\alpha_n^{**}(c) - \alpha}{d_n} \Rightarrow \mathcal{N}(0,1) \quad \text{where} \quad d_n = \frac{\sigma}{rE_n(\alpha)}. \tag{3}$$

Theorem 2

Let $\mathrm{Assumptions}\,1\&2$ hold. If $\rho_{2,n} \xrightarrow{p} 0$ and α_n^* is a consistent estimate of α, then convergence (3) takes place for $c = -1/2$.

Relationship between the two Theorems

Remark 1

If α_n^* is a consistent estimate of α and $\rho_{1,n} \xrightarrow{p} 0$ then

$$\alpha\rho_{2,n} \leq |\alpha_n^* - \alpha| \cdot \rho_{1,n} \xrightarrow{p} 0,$$

hence conditions of Theorem 2 are weaker than conditions of Theorem 1.

✓ Improved estimator — constructed

✓ Statistical properties — investigated

Thank you for your attention!