RAD HRVATSKE AKADEMIJE ZNANOSTI I UMJETNOSTI
MATEMATIČKE ZNANOSTI

D. Marković

Parameter estimation problem in the Box-Cox simple linear model

Manuscript accepted for publication

PARAMETER ESTIMATION PROBLEM IN THE BOX-COX SIMPLE LINEAR MODEL

DARIJA MARKOVIĆ

ABSTRACT. Given the data (x_i, y_i), $i = 1, \dots, n$, such that $y_i > 0$ for all $i = 1, \dots, n$, we consider the parameter estimation problem in a simple linear model with the Box-Cox transformation of the dependent variable. Maximum likelihood estimation of its parameter reduces to one nonlinear least squares problem. As a main result, we obtained three theorems in which we give necessary and sufficient conditions which guarantee the existence of the least squares estimate. In the most interesting case when at least three x_i's are different, it is shown that the least squares estimate exists.

1. INTRODUCTION

Suppose we are given the data (x_i, y_i), $i = 1, \dots, n$, such that $y_i > 0$ for all $i = 1, \dots, n$. The Box-Cox simple linear model has the form

$$y_i^{(\lambda)} = ax_i + b + \varepsilon_i, \quad i = 1, \dots, n,$$

where

$$y_i^{(\lambda)} = \begin{cases} \frac{y_i^\lambda - 1}{\lambda}, & \text{for } \lambda \neq 0 \\ \ln y_i, & \text{for } \lambda = 0, \end{cases} \tag{1.1}$$

and where it is assumed that errors ε_i are independent and normally distributed with zero mean and some unknown constant variance $\sigma^2 > 0$ (see [2]). The Box-Cox transformation (1.1) was proposed as a modification of the power transformation introduced by Turkey in [11] in order to avoid discontinuity at $\lambda = 0$. The theoretical properties and a variety of applications of the Box-Cox transformation (1.1) as well as other transformations can be found

2010 *Mathematics Subject Classification.* 65C20, 62J02.

Key words and phrases. Box-Cox regression model, maximum likelihood estimate, nonlinear least squares, least squares estimate, existence problem.

in [3]. A review and different extensions of the Box-Cox transformation with corresponding applications are given in the recent paper [1].

The maximum likelihood (ML) method was one of the techniques used in [2] to estimate the parameters of their model. The ML method provides the option to simultaneously estimate the transformation parameter λ and all regression parameters. It is well known that ML estimation of the unknown vector parameter (λ, a, b) reduces to the following nonlinear least squares (NLS) problem:

$$\min_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b), \qquad \text{where } F(\lambda, a, b) := \sum_{i=1}^{n} \Big(\frac{ax_i + b - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\Big)^2, \tag{1.2}$$

and

$$\dot{y} := \Big(\prod_{i=1}^{n} y_i\Big)^{\frac{1}{n}}$$

is the geometric mean of the y_i's (see, e.g., [2, 7]). If there exists a point $(\lambda_0, a_0, b_0) \in \mathbb{R}^3$ such that $F(\lambda_0, a_0, b_0) \le F(\lambda, a, b)$ for all $(\lambda, a, b) \in \mathbb{R}^3$, i.e. such that $F(\lambda_0, a_0, b_0) = \inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b)$, then it is called a global minimizer of F. It is also called a least squares estimate (LSE) of (λ, a, b) with respect to the problem (1.2) and the function F.

Numerical methods for solving the NLS problem are described in [6] and [8]. Prior to iterative minimization of the sum of squares it is still necessary to ask whether an LSE exists. For NLS problems, this question is difficult to answer (see, e.g., [4, 5, 9, 10]).

In the next section, after presenting some notations and preliminary results, we establish three theorems in which we give necessary and sufficient conditions which guarantee the existence of the LSE for problem (1.2).

2. Existence theorems for NLS problem (1.2)

Necessary and sufficient conditions for the existence of the LSE for problem (1.2) are given in theorems 2.3, 2.4 and 2.5. Before that, we need some notations and technical results which will be used in proofs of our results.

2.1. *Basic notations and preliminaries.* Let

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y}_\lambda := \frac{1}{n}\sum_{i=1}^{n} y_i^{(\lambda)},$$

and let continuous functions $\alpha, \beta, S : \mathbb{R} \to \mathbb{R}$ be defined by the formulae:

$$\alpha(\lambda) := \begin{cases} \frac{\sum_{i=1}^n (x_i-\bar{x})(y_i^{(\lambda)}-\bar{y}_\lambda)}{\sum_{i=1}^n (x_i-\bar{x})^2}, & \text{if } \sum_{i=1}^n (x_i-\bar{x})^2 \neq 0 \\ 0, & \text{otherwise,} \end{cases}$$

$$\beta(\lambda) := \bar{y}_\lambda - \alpha(\lambda)\bar{x},$$

$$S(\lambda) := F(\lambda, \alpha(\lambda), \beta(\lambda)).$$

By using a well-known fact that the quadratic function $t \mapsto \sum_{i=1}^{r}(t-u_i)^2$ attains its minimum $\sum_{i=1}^{r}(\tau_1 - u_i)^2$ at point $\tau_1 = \frac{1}{r}\sum_{i=1}^{r} u_i$, as well as the fact that the quadratic function $t \mapsto \sum_{i=1}^{r}(tv_i - u_i)^2$ attains its minimum $\sum_{i=1}^{r}(\tau_2 v_i - u_i)^2$ at point $\tau_2 = \sum_{i=1}^{r} u_i v_i / \sum_{i=1}^{r} v_i^2$, we obtain

$$\begin{aligned} F(\lambda, a, b) &= \sum_{i=1}^{n}\Big(\frac{ax_i + b - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\Big)^2 \\ &\geq \sum_{i=1}^{n}\Big(\frac{a(x_i - \bar{x}) + \bar{y}_\lambda - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\Big)^2 \\ &= F(\lambda, a, \bar{y}_\lambda - a\bar{x}) \\ &\geq F(\lambda, \alpha(\lambda), \bar{y}_\lambda - \alpha(\lambda)\bar{x}) \\ &= F(\lambda, \alpha(\lambda), \beta(\lambda)) \\ &= S(\lambda). \end{aligned} \tag{2.3}$$

Furthermore, it is easy to verify that if $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$, then

$$S(\lambda) = \frac{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2 \sum\limits_{i=1}^{n}(y_i^{(\lambda)} - \bar{y}_\lambda)^2 - \Big(\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i^{(\lambda)} - \bar{y}_\lambda)\Big)^2}{\dot{y}^{2(\lambda-1)} \sum\limits_{i=1}^{n}(x_i - \bar{x})^2}, \tag{2.4}$$

whereas if $\sum_{i=1}^{n}(x_i - \bar{x})^2 = 0$, then

$$S(\lambda) = \sum_{i=1}^{n}\Big(\frac{y_i^{(\lambda)} - \bar{y}_\lambda}{\dot{y}^{\lambda-1}}\Big)^2. \tag{2.5}$$

The next lemma will be used to prove our theorems 2.3, 2.4 and 2.5.

LEMMA 2.1. *With the notations as above, we have:*

(i) $\inf\limits_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = \inf\limits_{\lambda\in\mathbb{R}} S(\lambda)$.

(ii) *If a point (λ_0, a_0, b_0) is a global minimizer of F, then λ_0 is a global minimizer of S.*

(iii) *If λ_0 is a global minimizer of S, then $(\lambda_0, \alpha(\lambda_0), \beta(\lambda_0))$ is a global minimizer of F.*

(iv) *If $F(\lambda, a, b) \geq F(\lambda, a_0, b_0)$ for all $a, b \in \mathbb{R}$, then $F(\lambda, a_0, b_0) = S(\lambda)$.*

PROOF. *(i)* By (2.3) and the definition of infimum we obtain

$$F(\lambda, a, b) \geq S(\lambda) \geq \inf_{\lambda\in\mathbb{R}} S(\lambda) \quad \text{for all } (\lambda, a, b) \in \mathbb{R}^3,$$

and, consequently, $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) \geq \inf_{\lambda\in\mathbb{R}} S(\lambda)$. On the other hand, since

$$\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) \leq F(\lambda, \alpha(\lambda), \beta(\lambda)) = S(\lambda) \quad \text{for all } \lambda \in \mathbb{R},$$

it follows that $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda,a,b) \le \inf_{\lambda\in\mathbb{R}} S(\lambda)$.
(ii) Assume that (λ_0, a_0, b_0) is a global minimizer of F. Then, by virtue of (2.3), we observe that

$$S(\lambda) = F(\lambda, \alpha(\lambda), \beta(\lambda)) \ge F(\lambda_0, a_0, b_0) \ge S(\lambda_0) \quad \text{for all } \lambda \in \mathbb{R}.$$

Therefore, $\inf_{\lambda\in\mathbb{R}} S(\lambda) = S(\lambda_0)$.
(iii) If λ_0 is a global minimizer of S, by (2.3) and the definition of infimum we obtain that for all $(\lambda, a, b) \in \mathbb{R}^3$,

$$F(\lambda_0, \alpha(\lambda_0), \beta(\lambda_0)) = S(\lambda_0) \le S(\lambda) = F(\lambda, \alpha(\lambda), \beta(\lambda)) \le F(\lambda, a, b),$$

from where there follows a desired assertion.
(iv) By the assumption, $S(\lambda) = F(\lambda, \alpha(\lambda), \beta(\lambda)) \ge F(\lambda, a_0, b_0)$. On the other hand, by (2.3), $F(\lambda, a_0, b_0) \ge S(\lambda)$. □

The next lemma is also used in the proofs of theorems 2.3, 2.4 and 2.5. Its proof is omitted because it follows easily from the definition of infinite limit at infinity and the Extreme Value Theorem, which says that a continuous function from a closed interval attains its minimum value at some point in the closed interval.

Lemma 2.2. *Let $f : \mathbb{R} \to [0, \infty)$ be a continuous function such that*

$$\lim_{\lambda\to-\infty} f(\lambda) = \infty \quad \& \quad \lim_{\lambda\to\infty} f(\lambda) = \infty.$$

Then there exist reals $\lambda_1 < 0$, $\lambda_2 > 0$ and a point $\lambda_0 \in [\lambda_1, \lambda_2]$ such that

$$\inf_{\lambda\in\mathbb{R}} f(\lambda) = \inf_{\lambda\in[\lambda_1,\lambda_2]} f(\lambda) = f(\lambda_0).$$

2.2. *Existence theorems.*

Theorem 2.3. *If the data (x_i, y_i), $i = 1, \dots, n$, $n \ge 3$, are such that*

$$|\{x_1, \dots, x_n\}| \ge 3 \tag{2.6}$$

and $y_i > 0$ for all $i = 1, \dots, n$, then NLS problem (1.2) has a solution.

Proof. If $y_1 = y_2 = \ldots = y_n = \dot{y}$, then $F(\lambda, 0, \dot{y}^{(\lambda)}) = 0$ for each $\lambda \in \mathbb{R}$, and the proof is complete. Therefore, suppose further that

$$\min_{i=1,\dots,n} y_i < \dot{y} < \max_{i=1,\dots,n} y_i.$$

To complete the proof, it is enough to show that

$$\lim_{\lambda\to-\infty} S(\lambda) = \infty \quad \& \quad \lim_{\lambda\to\infty} S(\lambda) = \infty. \tag{2.7}$$

Indeed, by Lemma 2.2 this will mean that there exists a point $\lambda_0 \in \mathbb{R}$ such that $S(\lambda_0) = \inf_{\lambda\in\mathbb{R}} S(\lambda)$, and then according to assertion *(iii)* of Lemma 2.1, we have that $\inf\limits_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = F(\lambda_0, \alpha(\lambda_0), \beta(\lambda_0))$.

It remains to show (2.7). To do this, we will use the following equality that holds for each $y_r \in \{y_1, \dots, y_n\}$ and for all $\lambda \neq 0$:

$$S(\lambda) = \Big(\frac{y_r^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 \cdot H_r(\lambda), \tag{2.8}$$

where

$$H_r(\lambda) := \frac{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n \Big[\big(\frac{y_i}{y_r}\big)^\lambda - \frac{1}{n}\sum_{j=1}^n \big(\frac{y_j}{y_r}\big)^\lambda\Big]^2 - \Big[\sum_{i=1}^n (x_i - \bar{x})\Big(\big(\frac{y_i}{y_r}\big)^\lambda - \frac{1}{n}\sum_{j=1}^n \big(\frac{y_j}{y_r}\big)^\lambda\Big)\Big]^2}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

The above equality follows easily from (2.4). Let us first show that $\lim_{\lambda\to\infty} S(\lambda) = \infty$. If we take $y_r = \max\{y_i : y_i > \dot{y}\}$ in (2.8), after passing to the limit as $\lambda \to \infty$ we obtain

$$\lim_{\lambda\to\infty} \Big(\frac{y_r^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 = \infty$$

and

$$\begin{aligned}\lim_{\lambda\to\infty} H_r(\lambda) = {} & \sum_{i=1}^n (x_i - \bar{x})^2 \Big(\sum_{\substack{i=1\\ y_i = y_r}}^n \big(1 - \frac{L}{n}\big)^2 + \sum_{\substack{i=1\\ y_i \neq y_r}}^n \big(-\frac{L}{n}\big)^2\Big) \\ & - \Big[\sum_{\substack{i=1\\ y_i = y_r}}^n (x_i - \bar{x})\big(1 - \frac{L}{n}\big) + \sum_{\substack{i=1\\ y_i \neq y_r}}^n (x_i - \bar{x})\big(-\frac{L}{n}\big)\Big]^2 \\ & \geq 0,\end{aligned}$$

where L is the number of y_i's equal to y_r. Thus to prove that $\lim_{\lambda\to\infty} S(\lambda) = \infty$, it suffices to show that $\lim_{\lambda\to\infty} H_r(\lambda) > 0$. Otherwise, if $\lim_{\lambda\to\infty} H_r(\lambda) = 0$, according to the Cauchy-Schwarz inequality, there would exist a constant C such that

$$x_i - \bar{x} = \begin{cases} C\big(1 - \frac{L}{n}\big), & \text{if } y_i = y_r \\ -\frac{CL}{n}, & \text{if } y_i \neq y_r. \end{cases}$$

The latter would mean that only two x_i's are different, which contradicts assumption (2.6). Thus, we have proved that $\lim_{\lambda\to\infty} H_r(\lambda) > 0$. Arguing in a similar way, if we take $y_r = \min\{y_i : y_i < \dot{y}\}$ (2.8), after passing to the limit as $\lambda \to -\infty$, we obtain that $\lim_{\lambda\to-\infty} S(\lambda) = \infty$. This completes the proof of the theorem. □

Theorem 2.4. *If the data (x_i, y_i), $i = 1, \dots, n$, $n \geq 3$, are such that*

$$x_1 = x_2 = \cdots = x_n$$

and $y_i > 0$, for all $i = 1, \dots, n$, then NLS problem (1.2) has a solution.

Proof. If $y_1 = y_2 = \ldots = y_n = \dot{y}$, then $F(\lambda, 0, \dot{y}^{(\lambda)}) = 0$ for each $\lambda \in \mathbb{R}$, and the proof is complete. Therefore, suppose further that

$$\min_{i=1,\ldots,n} y_i < \dot{y} < \max_{i=1,\ldots,n} y_i.$$

Let

$$y_{i_0} := \min_{i=1,\ldots,n} y_i \quad \& \quad y_{i_1} := \max_{i=1,\ldots,n} y_i.$$

Note that for each $r \in \{1, \ldots, n\}$ and for all $\lambda \neq 0$, by virtue of (2.5) we have:

$$\begin{aligned} S(\lambda) &= \sum_{i=1}^{n} \Big(\frac{y_i^{(\lambda)} - \bar{y}_\lambda}{\dot{y}^{\lambda-1}} \Big)^2 = \sum_{i=1}^{n} \Big(\frac{y_i^\lambda}{\lambda \dot{y}^{\lambda-1}} - \frac{1}{n} \sum_{i=1}^{n} \frac{y_i^\lambda}{\lambda \dot{y}^{\lambda-1}} \Big)^2 \\ &\geq \Big(\frac{y_r^\lambda}{\lambda \dot{y}^{\lambda-1}} - \frac{1}{n} \sum_{i=1}^{n} \frac{y_i^\lambda}{\lambda \dot{y}^{\lambda-1}} \Big)^2 \\ &= \dot{y}^2 \Big(\frac{1}{\lambda} \big(\frac{y_r}{\dot{y}} \big)^\lambda \Big)^2 \Big(1 - \frac{1}{n} \sum_{i=1}^{n} \big(\frac{y_i}{y_r} \big)^\lambda \Big)^2 \end{aligned} \tag{2.9}$$

Since

$$\frac{y_{i_1}}{\dot{y}} > 1 \quad \& \quad \frac{y_{i_0}}{\dot{y}} < 1,$$

we have

$$\lim_{\lambda \to \infty} \Big(\frac{1}{\lambda} \big(\frac{y_{i_1}}{\dot{y}} \big)^\lambda \Big)^2 = \infty \quad \& \quad \lim_{\lambda \to -\infty} \Big(\frac{1}{\lambda} \big(\frac{y_{i_0}}{\dot{y}} \big)^\lambda \Big)^2 = \infty.$$

Therefore, after putting $r = i_0$ and $r = i_1$ in (2.9), we obtain

$$\lim_{\lambda \to \infty} S(\lambda) = \infty \quad \& \quad \lim_{\lambda \to -\infty} S(\lambda) = \infty.$$

Thus, according to Lemma 2.2, there exists a point $\lambda_0 \in \mathbb{R}$ such that $S(\lambda_0) = \inf_{\lambda \in \mathbb{R}} S(\lambda)$. Now, to complete the proof, note that from assertion *(iii)* of Lemma 2.1 it follows that $F(\lambda_0, \alpha(\lambda_0), \beta(\lambda_0)) = \inf\limits_{(\lambda,a,b) \in \mathbb{R}^3} F(\lambda, a, b)$. $\square$

Theorem 2.5. *Suppose that the data (x_i, y_i), $i = 1, \ldots, n$, $n \geq 3$, are such that*

$$|\{x_1, \ldots, x_n\}| = 2$$

and $y_i > 0$, for all $i = 1, \ldots, n$. Let

$$\xi_1 := \min_{i=1,\ldots,n} x_i, \quad \xi_2 := \max_{i=1,\ldots,n} x_i,$$

$$Y_{\xi_i} := \bigcup_{\substack{j=1 \\ x_j = \xi_i}}^{n} \{y_j\}, \quad i = 1, 2.$$

Then NLS problem (1.2) has no solution if and only if exactly one of the sets Y_{ξ_1} and Y_{ξ_2} is singleton and the second set is contained in $(0, \dot{y}]$ or in $[\dot{y}, \infty)$.

PROOF. Without loss of generality, assume that $|Y_{\xi_2}| = 1$, $|Y_{\xi_1}| \geq 2$ and $Y_{\xi_1} \subseteq [\dot{y}, \infty)$ or $Y_{\xi_1} \subseteq (0, \dot{y}]$. Let $Y_{\xi_2} = \{y_s\}$. Define continuous functions $\tilde{\alpha}, \tilde{\beta} : \mathbb{R}\backslash\{0\} \to \mathbb{R}$ by the formulae:

$$\begin{aligned}\tilde{\alpha}(\lambda) &:= \frac{y_s^\lambda}{\lambda(\xi_2 - \xi_1)}\\ \tilde{\beta}(\lambda) &:= -\tfrac{1}{\lambda} - \tilde{\alpha}(\lambda)\xi_1.\end{aligned}$$

Then

$$F(\lambda, \tilde{\alpha}(\lambda), \tilde{\beta}(\lambda)) = \sum_{\substack{i=1\\ x_i=\xi_1}}^{n} \left(\frac{\tilde{\alpha}(\lambda)\xi_1 + \tilde{\beta}(\lambda) - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\right)^2 + \sum_{\substack{i=1\\ x_i=\xi_2}}^{n} \left(\frac{\tilde{\alpha}(\lambda)\xi_2 + \tilde{\beta}(\lambda) - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\right)^2$$

$$= \sum_{\substack{i=1\\ x_i=\xi_1}}^{n} \dot{y}^2\left(\frac{y_i^\lambda}{\lambda\dot{y}^\lambda}\right)^2. \tag{2.10}$$

Let us show that $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = 0$. Indeed, if $Y_{\xi_1} \subseteq [\dot{y}, \infty)$, i.e., equivalently, if $y_i \geq \dot{y}$ for each y_i such that $x_i = \xi_1$, then, by virtue of (2.10), we obtain

$$\lim_{\lambda\to-\infty} F(\lambda, \tilde{\alpha}(\lambda), \tilde{\beta}(\lambda)) = 0,$$

implying that $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = 0$. But if $Y_{\xi_1} \subseteq (0, \dot{y}]$, once again by virtue of (2.10), we also get

$$\lim_{\lambda\to\infty} F(\lambda, \tilde{\alpha}(\lambda), \tilde{\beta}(\lambda)) = 0,$$

again implying that $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = 0$.

Since $\inf_{(\lambda,a,b)\in\mathbb{R}^3} F(\lambda, a, b) = 0$ and $|Y_{\xi_1}| \geq 2$, it follows that for all $(\lambda, a, b) \in \mathbb{R}^3$,

$$F(\lambda, a, b) \geq \sum_{\substack{i=1\\ x_i=\xi_1}}^{n} \left(\frac{a\xi_1 + b - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}}\right)^2 > 0,$$

and hence problem (1.2) has no solution.

Conversely, suppose that problem (1.2) has no solution. By assuming the theorem, $|Y_{\xi_1}| \geq 1$ and $|Y_{\xi_2}| \geq 1$. The proof will be done in three steps. In Step 1, we will show that the two sets Y_{ξ_1} and Y_{ξ_2} cannot be singletons. In Step 2, we will show that one of the sets Y_{ξ_1} or Y_{ξ_2} must be a singleton. The proof that the set which is not a singleton is contained in $(0, \dot{y}]$ or in $[\dot{y}, \infty)$ will be done in Step 3.

Step 1. We prove this by contradiction. Suppose to the contrary that $Y_{\xi_1} = \{y_{i_1}\}$ and $Y_{\xi_2} = \{y_{i_2}\}$. Choose any real number λ_0, and define

$$a_0 := \frac{y_{i_2}^{(\lambda_0)} - y_{i_1}^{(\lambda_0)}}{\xi_2 - \xi_1}, \quad b_0 := y_{i_1}^{(\lambda_0)} - a_0\xi_1.$$

Then $F(\lambda_0, a_0, b_0) = 0$, contradicting the assumption that problem (1.2) has no solution.
Step 2. Suppose that $|Y_{\xi_1}| \geq 2$ i $|Y_{\xi_2}| \geq 2$. Let N_{ξ_1} denote the number of data points with the abscissa ξ_1, and let N_{ξ_2} denote the number of data points with the abscissa ξ_2, i.e.,

$$N_{\xi_1} := |\{i \in \{1, \ldots, n\} : x_i = \xi_1\}|, \quad N_{\xi_2} := |\{i \in \{1, \ldots, n\} : x_i = \xi_2\}|.$$

By using a well-known fact that the quadratic function $t \mapsto \sum_{i=1}^{r} (t - u_i)^2$ attains its minimum $\sum_{i=1}^{r} (\tau_1 - u_i)^2$ at point $\tau_1 = \frac{1}{r} \sum_{i=1}^{r} u_i$, it is easy to verify that

$$\begin{aligned}
F(\lambda, a, b) &= \sum_{\substack{i=1 \\ x_i=\xi_1}}^{n} \left(\frac{a\xi_1 + b - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}} \right)^2 + \sum_{\substack{i=1 \\ x_i=\xi_2}}^{n} \left(\frac{a\xi_2 + b - y_i^{(\lambda)}}{\dot{y}^{\lambda-1}} \right)^2 \\
&\geq \sum_{\substack{i=1 \\ x_i=\xi_1}}^{n} \Big(\frac{1}{N_{\xi_1}} \sum_{\substack{j=1 \\ x_j=\xi_1}}^{n} \frac{y_j^{\lambda}}{\lambda \dot{y}^{\lambda-1}} - \frac{y_i^{\lambda}}{\lambda \dot{y}^{\lambda-1}} \Big)^2 \\
&+ \sum_{\substack{i=1 \\ x_i=\xi_2}}^{n} \Big(\frac{1}{N_{\xi_2}} \sum_{\substack{j=1 \\ x_j=\xi_2}}^{n} \frac{y_j^{\lambda}}{\lambda \dot{y}^{\lambda-1}} - \frac{y_i^{\lambda}}{\lambda \dot{y}^{\lambda-1}} \Big)^2 \\
&= F(\lambda, a_0, b_0)
\end{aligned} \tag{2.11}$$

for all $(\lambda, a, b) \in \mathbb{R}^3$, where

$$a_0 := \frac{\frac{1}{N_{\xi_2}} \sum\limits_{\substack{j=1 \\ x_j=\xi_2}}^{n} y_j^{(\lambda)} - \frac{1}{N_{\xi_1}} \sum\limits_{\substack{j=1 \\ x_j=\xi_1}}^{n} y_j^{(\lambda)}}{\xi_2 - \xi_1} \quad \text{and} \quad b_0 := \frac{1}{N_{\xi_1}} \sum_{\substack{j=1 \\ x_j=\xi_1}}^{n} y_j^{(\lambda)} - a_0 \xi_1.$$

According to assertion *(iv)* of Lemma 2.1, we have

$$F(\lambda, a_0, b_0) = S(\lambda).$$

Let

$$y_{min} := \min_{i=1,\ldots,n} y_i, \quad y_{max} := \max_{i=1,\ldots,n} y_i$$

Without loss of generality, assume that $y_{min} \in Y_{\xi_1}$ (the case $y_{min} \in Y_{\xi_2}$ can be handled in a similar way). Then from (2.11) it easily follows that

$$
\begin{aligned}
S(\lambda) &= F(\lambda, a_0, b_0) \\
&\geq \Bigg(\frac{1}{N_{\xi_1}} \sum_{\substack{j=1 \\ x_j=\xi_1}}^{n} \frac{y_j^\lambda}{\lambda \dot{y}^{\lambda-1}} - \frac{y_{min}^\lambda}{\lambda \dot{y}^{\lambda-1}}\Bigg)^2 \\
&= \Big(\frac{y_{min}^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 \Bigg(1 - \frac{1}{N_{\xi_1}} \sum_{\substack{j=1 \\ x_j=\xi_1}}^{n} \Big(\frac{y_j}{y_{min}}\Big)^\lambda\Bigg)^2. \qquad (2.12)
\end{aligned}
$$

Without loss of generality, we may also assume that $y_{max} \in Y_{\xi_1}$ (the case $y_{max} \in Y_{\xi_2}$ can be handled in a similar way). Once again, arguing as above, by virtue of (2.11), we get

$$
S(\lambda) \geq \Big(\frac{y_{max}^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 \Bigg(1 - \frac{1}{N_{\xi_1}} \sum_{\substack{j=1 \\ x_j=\xi_1}}^{n} \Big(\frac{y_j}{y_{max}}\Big)^\lambda\Bigg)^2, \qquad (2.13)
$$

Since

$$
\frac{y_{min}}{\dot{y}} < 1 \quad \& \quad \frac{y_{max}}{\dot{y}} > 1,
$$

we have that

$$
\lim_{\lambda \to -\infty} \Big(\frac{y_{min}^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 = \infty \quad \& \quad \lim_{\lambda \to \infty} \Big(\frac{y_{max}^\lambda}{\lambda \dot{y}^{\lambda-1}}\Big)^2 = \infty,
$$

and therefore from (2.12) and (2.13) we obtain that

$$
\lim_{\lambda \to -\infty} S(\lambda) = \infty \quad \& \quad \lim_{\lambda \to \infty} S(\lambda) = \infty.
$$

Thus, by Lemma 2.2, there exists a point $\lambda_0 \in \mathbb{R}$ such that $S(\lambda_0) = \inf_{\lambda \in \mathbb{R}} S(\lambda)$. Therefore, from assertion *(iii)* of Lemma 2.1 it follows that $\inf\limits_{(\lambda,a,b) \in \mathbb{R}^3} F(\lambda, a, b) = F(\lambda_0, \alpha(\lambda_0), \beta(\lambda_0))$, contradicting the assumption that problem (1.2) has no solution.

Step 3. Since $n \geq 3$, without loss of generality, we assume that $|Y_{\xi_2}| = 1$ and $|Y_{\xi_1}| \geq 2$. To complete the proof, it remains to show that $Y_{\xi_1} \subseteq [\dot{y}, \infty)$ or $Y_{\xi_1} \subseteq (0, \dot{y}]$. Suppose to the contrary that

$$
y_p := \min_{y_i \in Y_{\xi_1}} y_i < \dot{y} < \max_{y_i \in Y_{\xi_1}} y_i =: y_q.
$$

Then, arguing in the same way as in Step 2, whereby it is sufficient to replace y_{min} in (2.12) with y_p and y_{max} in (2.13) with y_q, we would obtain that problem (1.2) has a solution, which is in contradiction to the hypothesis. This completes the proof of the theorem. ∎

References

[1] A.C. Atkinson, M. Riani, A. Corbellini, *The Box-Cox Transformation: Review and Extensions,* Stat. Sci. **36** (2021), 239–255.
[2] G.E.P. Box, D.R. Cox, *An analysis of transformations,* J. Roy. Statist. Soc. Ser. B **26** (1964), 211–252.
[3] R.J. Carroll, D. Ruppert, Transformation and Weighting in Regression, Chapman & Hall, New York, 1988.
[4] E. Demidenko, *Criteria for global minimum of sum of squares in nonlinear regression,* Comput. Statist. Data Anal. **51** (2006), 1739–1753.
[5] E. Demidenko, *Criteria for unconstrained global optimization,* J. Optim. Theory Appl. **136** (2008), 375–395.
[6] J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Philadelphia, 1996.
[7] N.R. Draper, D.R. Cox, *On distributions and their transformation to normality,* J. Roy. Statist. Soc. Ser. B **31** (1969), 472–476.
[8] P.E. Gill, W. Murray, M.H. Wright, Practical Optimization, Academic Press, London, 1981.
[9] D. Jukić, *A necessary and sufficient criterion for the existence of the global minima of a continuous lower bounded function on a noncompact set,* J. Comput. Appl. Math. **375** (2020), 112791
[10] Y. Nievergelt, *On the existence of best Mitscherlich, Verhulst, and West growth curves for generalized least-squares regression,* J. Comput. Appl. Math. **248** (2013), 31–46.
[11] J.W. Tukey, *On the comparative anatomy of transformations,* Ann. Math. Statist. **28** (1957), 602–632.

Problem procjene parametara za jednostavan linearni Box-Cox model

Darija Marković

Sažetak. Za dane podatke (x_i, y_i), $i = 1, \dots, n$, takve da je $y_i > 0$ za sve $i = 1, \dots, n$, razmatramo problem procjene parametara za jednostavan linearni model s Box-Cox-ovom transformacijom zavisne varijable. Procjena njegovih parametara metodom maksimalne vjerodostojnosti svodi se na nelinearan problem najmanjih kvadrata. Kao glavni rezultat, dobili smo tri teorema u kojima su dani nužni i dovoljni uvjeti koji jamče egzistenciju procjenitelja najmanjih kvadrata. U najinteresantnijem slučaju kada su barem tri x_i različiti, pokazano je kako procjenitelj najmanjih kvadrata postoji.

Darija Marković
Department of Mathematics
University of Osijek
31 000 Osijek, Croatia
E-mail: `darija@mathos.hr`