The existence theorem for the solution of a nonlinear least squares problem*

Dragan Jukić†

Abstract. *In this paper we prove a theorem which gives necessary and sufficient conditions which guarantee the existence of the global minimum for a continuous real valued function bounded from below, which is defined on a non-compact set. The use of the theorem is illustrated by an example of the least squares problem.*

Key words: *least squares, existence problem, exponential function*

Sažetak. **Teorem o egzistenciji rješenja nelinearnog problema najmanjih kvadrata**. *U radu je naveden teorem koji daje nužan i dovoljan uvjet za egzistenciju globalnog minimuma neprekidne i odozdo omeđene realne funkcije definirane na skupu koji nije kompaktan. Korisnost teorema ilustrirana je na primjeru problema najmanjih kvadrata.*

Ključne riječi: *problem najmanjih kvadrata, problem egzistencije, eksponencijalna funkcija*

1. The least squares problem

We are given a model-function

$$t \mapsto f(t; \mathbf{a}), \tag{1}$$

and the data (p_i, t_i, f_i), $i = 1, \ldots, m$, where $\mathbf{a} \in \Lambda \subseteq R^n$ is the vector of unknown parameters, $t_1 < t_2 < \ldots < t_m$ are the abscissae and $f_1, \ldots, f_m$ are the data's ordinates. The number $p_i > 0$ is the weight of the i-th datum. Usually we have

*The lecture presented at the Mathematical Colloquium in Osijek organized by Croatian Mathematical Society - Division Osijek, January 19, 1996.

†Faculty of Agriculture, Department of Mathematics, Trg Svetog Trojstva 3, HR-31 000 Osijek, Croatia, e-mail: jukicd@oliver.unios.hr

$m \gg n$, i.e. usually the number of data is considerably bigger than the number of unknown parameters.

In practice, the unknown parameter vector $\mathbf{a}^\star \in \Lambda$ for the function-model (**??**) is usually determined either in the sense of ordinary least squares (c.f. [**?**], [**?**], [**?**], [**?**], [**?**], [**?**]), by finding $\mathbf{a}^\star \in \Lambda$ such that

$$S(\mathbf{a}^\star) = \inf_{\mathbf{a}\in\Lambda} S(\mathbf{a}), \quad S(\mathbf{a}) = \frac{1}{2}\sum_{i=1}^{m} p_i [f_i - f(t_i;\mathbf{a})]^2 \tag{2}$$

(*Figure* **??**.*a*), or in the sense of total least squares (c.f. [**?**], [**?**], [**?**]) by finding $(\mathbf{a}^\star, \boldsymbol{\delta}^\star) \in \Lambda \times R^m$ such that

$$F(\mathbf{a}^\star, \boldsymbol{\delta}^\star) = \inf_{(\mathbf{a},\boldsymbol{\delta})\in\Lambda\times\mathbb{R}^m} F(\mathbf{a}, \boldsymbol{\delta}), \quad F(\mathbf{a}, \boldsymbol{\delta}) = \frac{1}{2}\sum_{i=1}^{m} p_i \left\{ [f_i - f(t_i + \delta_i;\mathbf{a})]^2 + \delta_i^2 \right\}, \tag{3}$$

(*Figure* **??**.*b*), where $\boldsymbol{\delta} = (\delta_1, \ldots, \delta_m)^T \in R^m$.

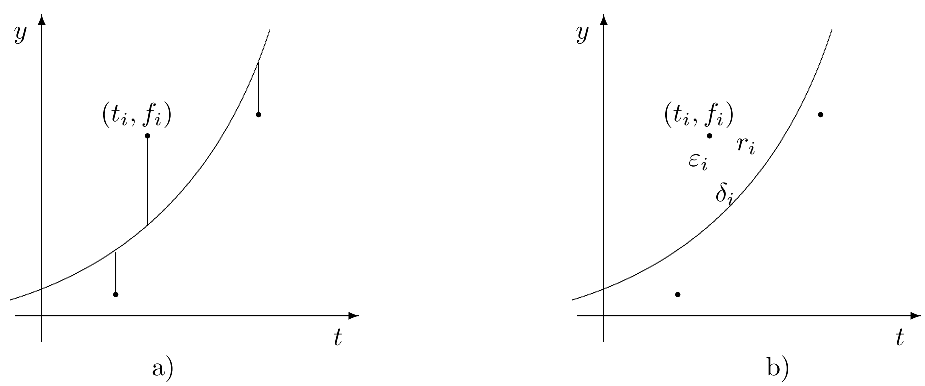


Figure 1:

Hence, in both cases one has an existence problem for a global minimum of a continuous function on some set. In the next section we prove a theorem giving necessary and sufficient conditions for the above problem to have a solution.

2. The existence theorem for the solution of the least squares problem

Let $\Lambda \subseteq R^n$ and let $G : \Lambda \to R$ be a continuous function which is bounded from below, and denote

$$G^\star := \inf_{\lambda \in \Lambda} G(\lambda).$$

One has the following problem:

(P1) *Does there exist a point $\lambda^\star \in \Lambda$, such that $G(\lambda^\star) = G^\star$?*

If the set Λ is compact, then, by continuity of G, the problem (P1) has a solution. Therefore, from now on we assume that the set Λ is not compact. This means that the set Λ is either unbounded or is not closed (or both). In the first case there exists a sequence (x_k) in Λ such that $||x_k|| \to \infty$, and in the second case there exists a sequence (x_k) in Λ converging to a point $x^\star \in \operatorname{Cl}\Lambda \setminus \Lambda$.

Denote by $N(\Lambda)$ the set of all sequences (x_k) in Λ, such that $||x_k|| \to \infty$ or $x_k \to x^\star \in \operatorname{Cl}\Lambda \setminus \Lambda$. Following Demidenko [**?**] (see *Remark* **??**), define the number

$$\overline{G}_E := \inf\{\liminf G(x_k) \,:\, (x_k) \in N(\Lambda)\} \tag{4}$$

which we call *the infimum of the function G at the end.*

Since $G^\star \leq G(\lambda)$ for all $\lambda \in \Lambda$, from the definition (**??**) one can easily prove the following inequality:

Proposition 1. *$G^\star \leq \overline{G}_E$.*

Example 1. *Let $G(\lambda) = \lambda^2$.*

a) If $\Lambda = R$, then $G^\star = 0$, $\overline{G}_E = \infty$,

b) If $\Lambda = [0, \infty)$, then $G^\star = 0$, $\overline{G}_E = \infty$,

c) If $\Lambda = (0, \infty)$, then $G^\star = 0$, $\overline{G}_E = 0$,

d) If $\Lambda = (-\infty, -1] \cup (1, \infty)$, then $G^\star = 1$, $\overline{G}_E = 1$.

Note that the problem (P1) has no solution only in the case c).

Proposition 2. *If $G^\star < \overline{G}_E$, then the problem (P1) has a solution.*

Proof. Let (λ_k) be a sequence in Λ such that $G^\star = \lim_{k\to\infty} G(\lambda_k)$. By our assumption $G^\star < \overline{G}_E$, and therefore neither the sequence (λ_k), nor any of its subsequence, belong to the set $N(\Lambda)$. This means that the sequence $||\lambda_k||$ does not tend to ∞, and hence the sequence (λ_k) has a bounded subsequence. By the Bolzano-Weierstrass theorem, the sequence (λ_k) has a convergent subsequence (λ_{k_i}). Let $\lambda_{k_i} \to \lambda^\star$. Then $\lambda^\star \in \Lambda$, since otherwise, the sequence (λ_{k_i}) would belong to the set $N(\Lambda)$. By the continuity of the function G we obtain $G^\star = \lim_{i\to\infty} G(\lambda_{k_i}) = G(\lim_{i\to\infty} \lambda_{k_i}) = G(\lambda^\star)$. □

Theorem 1. *The problem (P1) has a solution if and only if there exists a point $\lambda^\star \in \Lambda$ such that $G(\lambda^\star) \leq \overline{G}_E$.*

Proof. a) Suppose the problem (P1) has a solution. Then there exists a point $\lambda^\star \in \Lambda$ such that $G(\lambda^\star) = G^\star$. By *Proposition* **??** we obtain $G(\lambda^\star) \le \overline{G}_E$.

b) Suppose the problem (P1) has no solution. Then $G(\lambda) > G^\star$ for all $\lambda^\star \in \Lambda$. Furthermore, by *Proposition* **??**, $G^\star = \overline{G}_E$. Hence, if the problem (P1) has no solution, then $G(\lambda) > \overline{G}_E$ for all $\lambda \in \Lambda$. □

Remark 1. *In [**?**], the infimum of the function G at the end, is defined as the number $\overline{G}^\star_E := \inf\{\liminf G(x_k) \,:\, (x_k) \in N^\star(\Lambda)\}$, where $N^\star(\Lambda)$ is the set of all sequences (x_k) in Λ, such that either $||x_k|| \to \infty$ or $x_k \to x^\star \in \mathrm{Cl}\,\Lambda \setminus \mathrm{Int}\,\Lambda$.*

Since $\mathrm{Cl}\,\Lambda \setminus \Lambda \subseteq \mathrm{Cl}\,\Lambda \setminus \mathrm{Int}\,\Lambda$, one has $\overline{G}^\star_E \le \overline{G}_E$. Furthermore, it is easy to show that Theorem **??** *holds also if $\overline{G}_E$ is replaced by $\overline{G}^\star_E$.*

Example 2. *We illustrate the application of* Theorem **??** *on the following simple nonlinear ordinary least squares problem. Let the data (p_i, t_i, f_i), $i = 1, \ldots, m$, $m \ge 3$, be such that $t_1 < t_2 < \ldots < t_m$, and $f_i > 0$, $i = 1, \ldots, m$. For the model-function we take the exponential function $f(t; \lambda) = e^{\lambda t}$, $\lambda \in R$. We consider the existence problem for the global minimum of the functional $S : R \to R$ given by*

$$S(\lambda) = \sum_{i=1}^{m} p_i [f_i - e^{\lambda t_i}]^2.$$

Let $I = \{1, \ldots, m\}$, $I_0 := \{i : t_i = 0\}$. Note that the set I_0 is either empty or contains only one point. Let us show that in the set R there exists a point of the global minimum for the functional S.

a) If some of the numbers $t_1, \ldots, t_m$ are strictly positive and some are strictly negative, then obviously $\overline{S}_E = \infty$. By Theorem **??** *there exists a point of the global minimum for the functional S on R.*

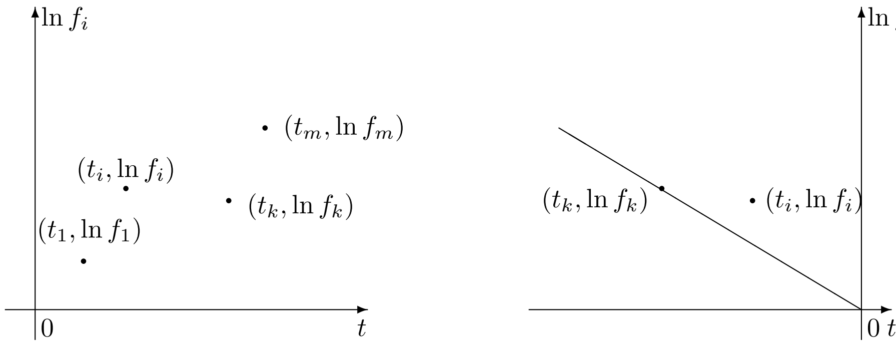
b) If $0 \le t_1$, then $\overline{S}_E = \sum_{i \in I \setminus I_0} p_i f_i^2 + \sum_{i \in I_0} p_i (f_i - 1)^2$. Let (see Figure **??**.a*)*

$$\lambda^\star = \min_{i \in I \setminus I_0} \frac{\ln f_i}{t_i} = \frac{\ln f_k}{t_k}.$$

Geometrically, (see Figure **??**.a*), $\lambda^\star$ is the slope of the line through the origin and the point $(t_k, \ln f_k)$, and none of the other points $(t_i, \ln f_i)$, $i \in I \setminus I_0$, lies below this line. Since $f_k = e^{\lambda^\star t_k}$ and $f_i \ge e^{\lambda^\star t_k}$, $i \ne k$, we have*

$$\begin{aligned} S(\lambda^\star) &= \tfrac{1}{2} \sum_{i=1}^{m} p_i [f_i - e^{\lambda^\star t_i}]^2 = \tfrac{1}{2} \sum_{\substack{i=1 \\ i \ne k}}^{m} p_i [f_i - e^{\lambda^\star t_i}]^2 \\ &\le \tfrac{1}{2} \sum_{\substack{i \in I \setminus I_0 \\ i \ne k}} p_i f_i^2 + \tfrac{1}{2} \sum_{i \in I_0} p_i (f_i - 1)^2 < \overline{S}_E. \end{aligned}$$

By Theorem **??** *in this case there aslo exists a point of the global minimum for the functional S on R.*



a) *b)*

Figure 2:

c) *If* $t_m \le 0$, *then* $\overline{S}_E = \frac{1}{2}\sum_{i\in I\setminus I_0} p_i f_i^2 + \frac{1}{2}\sum_{i\in I_0} p_i(f_i-1)^2$. *Let (see* Figure **??**.b*)*

$$\lambda^\star = \max_{i\in I\setminus I_0} \frac{\ln f_i}{t_i} = \frac{\ln f_k}{t_k}.$$

Proceeding similarly as in b), one can show that there exists a point $\lambda^\star \in R$ *such that* $S(\lambda^\star) < \overline{S}_E$.

Therefore, our least square problem always has a solution.

References

[1] A. BJÖRCK, *Least Squares Methods*, in Handbook of Numerical Analysis, Vol. I, P. G. Ciarlet and J. L. Lions, eds., Elsevier Science Publishers B.V., North - Holland, Amsterdam, 1990, 467–653.

[2] P. T. BOGGS, R. H BYRD, R. B. SCHNABEL, *A stable and efficient algorithm for nonlinear orthogonal distance regression*, SIAM J. Sci. Statist. Comput. **8**(1987), 1052–1078.

[3] E. Z. DEMIDENKO, *Optimizacia i regresija*, Nauka, Moskva, 1989.

[4] P. E. GILL, W. MURRAY, M. H. WRIGHT, *Practical Optimization*, Academic Press, London, 1981.

[5] D. JUKIĆ, R. SCITOVSKI, *The existence of optimal parameters of the generalized logistic function*, Appl. Math. Comput., to appear.

[6] H. Schwetlick, V. Tiller, *Numerical methods for estimating parameters in nonlinear models with errors in the variables*, Technometrics **27**(1985), 17–24.

[7] R. Scitovski, *A special nonlinear least squares problem*, J. Comput. Appl. Math. **53**(1994), 323–331.

[8] R. Scitovski, D. Jukić, *A method for solving the parameter identification problem for ordinary differential equations of the second order*, Appl. Math. Comput. **74**(1996), 273–291.

[9] R. Scitovski, D. Jukić, *Total least squares problem for exponential function*, Inverse Problems, to appear.