ON THE EXISTENCE OF THE LEAST SQUARES ESTIMATE IN NONLINEAR GROWTH CURVE MODELS OF EXPONENTIAL TYPE

Eugene Demidenko

Dartmouth Medical School,
Section of Biostatistics and Epidemiology, HB 7927
Hanover, New Hampshire 03755-3861

Published in
Communications in Statistics.
Theory and Methods, 25, 159-182 (1996)

Key words and Phrases: Existence; Growth curve, Nonlinear regression; Least squares; Exponential model; Logistic model; Gompertz curve.

ABSTRACT

A general criterion for the existence of a global minimizer of a continuous function on a noncompact set is developed. Criteria for the existence of the least squares estimate in some popular nonlinear growth curve models of exponential type are derived: the quasilinear regression model, two- and three-parameter exponential model, modi¯ed exponential model, Gompertz curve, and logistic model. The concept of the "existence level", as the minimum of the sum of squares on the boundary of the parameter set, is introduced. Simple procedures for checking

whether a nonlinear least squares estimate exists, and suitable initial starting values for particular growth curve models are presented. These concepts and derived criteria are illustrated using the logistic model on a real life biomedical example of mouse tumor growth.

1. INTRODUCTION

Almost all procedures of point estimation lead to minimizing (or maximizing) some function. The most commonly used optimization procedures in statistics are maximum likelihood (ML) and the least squares (LS) methods of estimation. Even if the minimization function is continuous, when the parameter set is not compact (for instance, coinciding with the whole space), the problem of existence of the estimate is not trivial. Makelainen & Schmidt and Styan (1981) addressed this problem for the particular case of the maximum likelihood estimate. A number of other authors have considered the issue of the existence of the ML-estimate for speci¯c statistical models. The existence and uniqueness of the ML-estimate for generalized linear models was considered by Wedderburn (1967). Albert & Anderson (1984) established necessary and su±cient conditions for the existence of the ML-estimate for the logistic regression model. Recently, Lesa®re & Kaufmann (1993) derived a similar criterion for the multivariate probit model. A general criterion for the existence of a minimizer of a continuous function on a noncompact set was formulated by Demidenko (1981, 1989). Nakamura (1984) provides conditions for the existence of the ML-estimate for censored data. Nakamura & Lee (1993) formulate criteria for the existence of the ML-estimate for a binary model based on a general criterion for the existence similar to formulated by Demidenko (1989).

The most important examples of the functions to minimize in statistics are: the minus log-likelihood function and the sum of squares. We adapt that general approach to derive criteria for the existence of the least squares estimate for some popular nonlinear growth curve models of exponential type (e.g. Seber & Wild, 1989). In the present paper the parameter set coincides with the whole space which simpli¯es formulas and development of the appropriate criteria.

The availability of those criteria allows us:

² quickly to verify whether the a least squares problem has a solution,

- to generate a suitable initial starting value for an iterative minimization algorithm.

The aim of the present paper is to develop a general criterion for the existence of a global minimizer and to apply it to the sum of squares for some popular nonlinear growth curve models. The structure of the paper is as follows. In section 2, the concept of "existence level" is introduced, and a general existence criterion is developed. In section 3 the existence criterion is derived for the quasilinear regression model, a generic type of growth curve models considered later. Section 4 and 5 are devoted to the nonlinear growth curve models of exponential type: modified exponential model, Gompertz curve, and logistic model. These models are very popular in economics, engineering and biology.

2. EXISTENCE LEVEL

A growth curve model can be formulated in terms of a nonlinear regression model

$$y_t = f(x_t; \alpha) + \varepsilon_t; \qquad t = 1; \ldots; n \tag{2.1}$$

where y_t is the dependent variable (in the setting of a growth curve model it can be interpreted as the observation of the growth variable); f is a continuous response or growth function, x_t is the fixed vector of independent variables; t is the time index; ε_t is the unobservable error term with $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_j) = 0$ for $t \neq j$ and $E(\varepsilon_t^2) = \mathrm{var}(\varepsilon_t) = \sigma^2$; n is the sample size; α is an $m \times 1$ vector of unknown parameters, $\alpha \in R^m$: Since x_t is fixed, the regression equation (2.1) may be written in a more convenient form,

$$y_t = f_t(\alpha) + \varepsilon_t; \quad t = 1; \ldots; n: \tag{2.2}$$

The least squares estimate (LSE) of α; by definition, returns the global minimum of the sum of squares (SS):

$$S(\alpha) = \sum_{t=1} (y_t - f_t(\alpha))^2; \quad \alpha \in R^m: \tag{2.3}$$

The fundamental question is whether given data $y_1; \ldots; y_n$ the LSE exists, or in other words, whether the infimum of the function $S(\alpha)$ is attainable on R^m. The following simple yet general criterion for the existence will be adapted.

General criterion for the existence of a global minimizer of a continuous function on a noncompact set (Demidenko, 1981). Let $F(u)$ be a general continuous function of $u \in R^m$. The "existence level" (EL) for F is defined as

$$F_E = \lim_{r \to \infty} \inf_{\|u\|>r} F(u). \tag{2.4}$$

If there exists $u_0 \in R^m$ such that $F(u_0) < F_E$, then the global minimizer of F exists, and the level set

$$S_0 = \{u \in R^m : F(u) \leq F(u_0)\}$$

is compact.

In other words, the EL is the infimum of F at infinity. Makelainen & Schmidt and Styan (1981) proved this criterion under the assumption that $F(u)$ is constant at infinity, $\lim_{u \to \infty} F(u) = \text{const}$.

For practical purposes, an equivalent and more convenient formula in terms of disconvergent sequences of the argument (we say that u_k are disconvergent if $\|u_k\| \to \infty$) will be used:

$$F_E = \inf_{\|u_k\| \to \infty} \{\underline{\lim}_k F(u_k)\} \tag{2.5}$$

where by definition $\underline{\lim}_k F(u_k) = \lim_{p \to \infty} \min_{k>p} F(u_k)$. In other words, the EL is equal to the lower bound of F at infinity. It is easy to see that the limit in (2.5) always exists (it is permissible for F_E to equal $-\infty$). If for all $u \in R^m$ we have $F(u) > F_E$ then the infimum of F is not attainable, or in statistical context the estimate does not exist.

A vector u_0 that satisfies the inequality $F(u_0) < F_E$ is called a "suitable starting value". In particular, starting from u_0 any iterative algorithm for F will produce a sequence which has at least one limit point if $F(u_k) \leq F(u_0)$, $k = 1, 2, \ldots$.

We will adapt the general criterion of existence to the sum of squares (2.3) as a function of the unknown vector parameter α. Formally $S = F$ and $\alpha = u$. In this case the global minimizer is called the LSE. The typical form of the sum of squares with asymptote as the EL is shown in Figure 1. Here, if the initial parameter value is large enough and lies in the region of the unsuitable parameter value any minimization algorithm will produce a disconvergent sequence of α's.

To prove the existence of the LSE for a specific growth curve model, we first have to calculate the existence level S_E. Second, one has to find a suitable starting value which satisfies the inequality $S(a_0) < S_E$: Then the LSE exists and the level set

$$S_0 = \{\alpha \in R^m : S(\alpha) \le S(a_0)\} \tag{2.6}$$

is compact. The issue of uniqueness of the LSE is beyond the scope of the present paper.

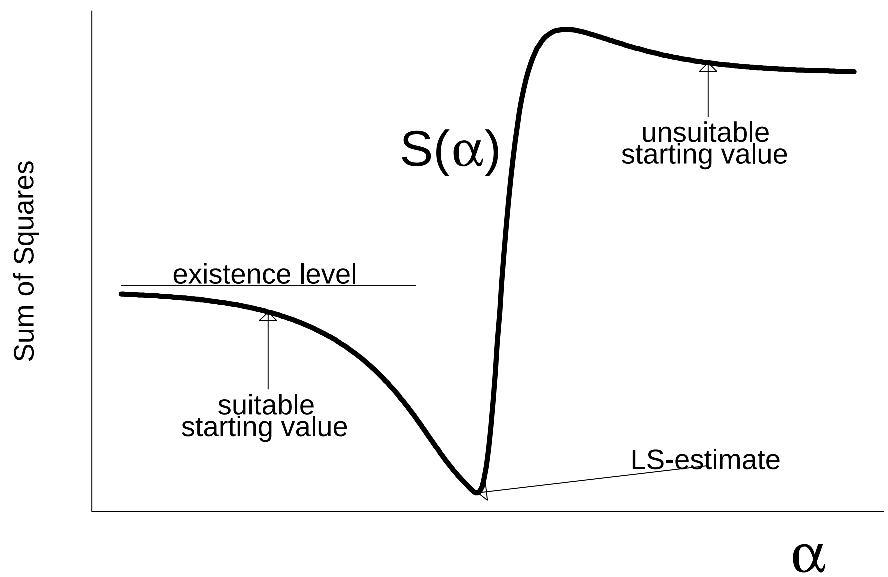

FIG 1. A typical form of the sum of squares and the existence level.

The concept of the limit regression is introduced for finite limit of f_t with disconvergent sequences of α:

$$\lim_{\|\alpha\| \to \infty} f_t(\alpha) = \bar{f}_t(\beta); \tag{2.7}$$

where β is a parameter vector indexing theses α's: If we denote

$$\bar{S} = \min_{\beta} \sum_{i=1} \left(y_i - \bar{f}_t(\beta) \right)^2 \tag{2.8}$$

then by the definition S_E is equal to the minimum of $\bar{S}$ over all limit regression functions $\bar{f}_t$. This concept will be illustrated later.

To illustrate the concept of the existence level the following simple example is considered (see Fig. 1).

Example. Let $y_t = e^{\alpha x_t} + \varepsilon_t$, $t = 1, \ldots, n$, where without loss of generality one can assume that the x_t are ordered: $x_1 \le x_2 \le \ldots \le x_n$. If $x_1 < 0$ and $x_n > 0$ then $S_E = \infty$. Indeed, if $\alpha \to +\infty$ then $f_1(\alpha) = \exp(\alpha x_n) \to \infty$ and $S(\alpha) \to \infty$. If $\alpha \to -\infty$ then $f_1(\alpha) = \exp(\alpha x_1) \to \infty$ and again $S(\alpha) \to \infty$. Now we consider the case when $x_1 > 0$. Then $\lim_{\alpha \to -\infty} f_t(\alpha) = 0 = \bar{f}_t$, the limit regression function with $\bar{S} = \sum_1^n y_t^2$. Thus, in this case $S_E = \sum_1^n y_t^2$. $\Box$

In the following sections we will illustrate the calculation of the existence levels of several popular growth curve models and suggest suitable initial starting values. We start with a generic type of growth curve model which is a prototype for the growth curves considered later.

3. QUASILINEAR REGRESSION

One of the simplest type of a nonlinear regression model was introduced by Kubicek, Marek and Eckert (1971) and was called "quasilinear regression". In this model the response function is $f_t(\alpha) = g(\alpha' x_t)$, where $\alpha, x_1, \ldots, x_n \in R^m$, and $g(s)$ is a strictly monotonic function of $s \in (-\infty, \infty)$. The technique developed in this section will be applied to other regression models containing linear combinations of parameters through $\alpha' x_t$, as will be shown in the following sections.

It is easy to show that, without loss of generality, we can assume $g(\cdot)$ to be strictly increasing and $x_t \neq 0$. We also assume that the $n \times m$ matrix X with x_t' as the tth row is of full rank $m < n$. The following two cases are considered:

Case A. The function g is unbounded: $g(s) \to +\infty$ when $s \to +\infty$ and $g(s) \to -\infty$ when $s \to -\infty$.

Case B. The function g is bounded from below, and unbounded from above: $\lim_{s \to -\infty} g(s) = g_*$ and $\lim_{s \to +\infty} g(s) \to +\infty$.

It is easy to verify that in Case A we have $S_E = +\infty$. Indeed, when $\|\alpha\|$ is disconvergent

$$\sum (\alpha' x_t)^2 = \alpha' (\sum x_t x_t') \alpha = \alpha' X' X \alpha \ge \|\alpha\|^2 \lambda_{\min}(X'X) \to \infty \qquad (3.1)$$

where $\lambda_{min}(X'X)$, the minimal eigenvalue of matrix $X'X$, is positive by the previous assumption. Further, as follows from (3.1), $|\alpha' x_j| \to \infty$ which implies $f_j(\alpha) = g(\alpha' x_j) \to \infty$ for at least one index $j \in \{1, ..., n\}$. Thus, $S(\alpha) = \|y - f(\alpha)\|^2 \geq \|f(\alpha)\| - \|y\| \to \infty$. As a consequence, for all $a_0 \in R^m$ the level set (2.6) is compact.

The calculation of the EL for Case B requires more analysis. The theory of polyhedral cones and some concepts of convex analysis will be exploited for this purpose (e.g. Kamthan & Gupta, 1985 or Cottle & Pang and Stone, 1992).

Definition. The system of vectors $x_1, \ldots, x_n$ is called unidirectional, if there is a vector $\gamma \in R^m$ such that $\gamma' x_t > 0$ for all $t = 1, \ldots n$. The system of vectors is called multidirectional if it is not unidirectional. It is called strictly multidirectional if for all $\gamma \in R^m, \gamma \neq 0$ there is $j \in \{1, ..., n\}$ such that $\gamma' x_j < 0$.

First we will show that for a quasilinear regression with strictly multidirectional vectors $S_E = +\infty$. For this purpose we will use an analog of formula (2.5) for the sum of squares:

$$S_E = \inf \underline{\lim}_{\|\alpha_k\| \to \infty} S(\alpha_k). \tag{3.2}$$

Let $\{\alpha_k\}$ be a disconvergent sequence of parameters, $\|\alpha_k\| \to \infty$. We consider the limit points γ^* of the sequence of normalized vectors: $\{\gamma_k = \alpha_k / \|\alpha_k\|, k = 1, 2, \ldots\}$. Let $\{\alpha_p\}$ be a subsequence such that $\alpha_p / \|\alpha_p\| \to \gamma^*$ when $p \to \infty$ (since γ_k belong to a compact set any sequence has at least one limit point). Because vectors $\{x_t\}$ are strictly multidirectional there is an index j such that $x_j' \gamma^* > 0$. Then there is a $\delta > 0$ and a p_0 such that for all $p > p_0$ we have $x_j' \alpha_p / \|\alpha_p\| \geq \delta$. Thus, $x_j' \alpha_p \geq \delta \|\alpha_p\| \to \infty$ when $p \to \infty$ and we obtain $g(x_j' \alpha_p) \to \infty$. This implies $S_E = +\infty$, which completes the proof.

The case is richer when the vectors $\{x_t\}$ are unidirectional. The reader is referred to Figure 2 for a geometrical illustration in the case of a two-parameter quasilinear regression. In the next section this regression is considered in details for $x_t = (1, x_t)'$. It is easy to show that $\{x_t\}$ are unidirectional if the intercept term is present.

As follows from formula (3.2), to obtain finite values for $\underline{\lim} S(\alpha_k)$ we have to consider those sequences of $\{\alpha_k\}$ which either produce finite limits of $x_t' \alpha_k$ or $x_t' \alpha_k \to -\infty$ for all $t = 1, \ldots, n$, because if $x_j' \alpha_k \to +\infty$ for some $j \in \{1, ..., n\}$, then $S_k = S(\alpha_k) \to +\infty$. For further analysis we will introduce two cones associated with n vectors $\{x_t\}$. The first one is generated by $\{x_t\}$:

$$K = \{\alpha \in R^m : \alpha = \sum_t \lambda_t x_t, \quad \lambda_t \geq 0\}. \tag{3.3}$$

It is known that the cone K is polyhedral and spanned by a certain number of vectors from $\{x_t\}$. We denote these $\{x_{t_1}, \dots, x_{t_p}\}$ where $t_i \in \{1, \dots, n\}$, $i = 1, \dots, p$. Vectors x_{t_i} are called "edges". To simplify, we assume that any edge cannot be expressed through other edges as a linear combination with positive coefficients. Since the system $\{x_t, t = 1, \dots, n\}$ is of full rank the system of vectors $\{x_{t_1}, \dots, x_{t_p}\}$ is of full rank too.

The second cone will be denoted as K^- and called the "negative conjugate cone". It consists of all vectors that have a right or dull angle with vectors from $\{x_t\}$:

$$K^- = \{\alpha \in R^m : x_t'\alpha \le 0, \ \forall t = 1, \dots, n\}. \tag{3.4}$$

The number of edges of K^- is equal to p and each edge of K^- is orthogonal to $m-1$ edges of K. Let us denote by $G_1, \dots, G_P$ those sets of $m-1$ edges of K which define the edges of K^-. Then, as is proved in the Appendix, the EL for quasilinear regression is defined by the following formula:

$$S_E = \sum_{t=1} (y_t - g_*)^2 - \max_{k \in [1;p]} \sum_{j \in G_k} (y_j - g_*)^2. \tag{3.5}$$

From (3.5) a simple and useful formula can be derived. Without loss of generality, we can reindex our observations so that $y_{(1)} \le y_{(2)} \le \dots \le y_{(n)}$. Then since $p \ge m$ we obtain

$$S_E \ge \sum_{t=1}^{n-m+1} (y_{(t)} - g_*)^2. \tag{3.6}$$

Therefore, applying the General criterion for existence from section 2 we conclude that if $S(a_0) < S_E$ for some a_0, or $S(a_0)$ is less than the right hand side of (3.6), then the LSE for a quasilinear regression exists and the level set (2.6) is compact. Based on formula (3.5) one can prove the following simple criterion.

Criterion for the LSE existence for quasilinear regression. *If $y_t > g_*$ for all $t = 1, \dots, n$ then the LSE for the quasilinear regression exists.*

Proof. See Appendix.

Now we shall consider the problem of the calculation of a suitable initial guess. Recall that if the $\{x_t\}$ are multidirectional then $S_E = \infty$ and any parameter vector can be considered as a suitable initial guess. Let us assume that the $\{x_t\}$ are unidirectional, for instance, an intercept term is present, and $y_t > g_*$, $t = 1, \dots, n$. Without loss of generality one can assume that the maximum in the second sum of (3.5) is achieved for the first $m-1$ set of vectors G_1, $k = 1$.

Let $P = \{(x;u) \in R^{m+1} : u = x'\beta\}$ be a plane in the space $(x; g^{-1}(y))$ passing through points $\{(x_t; g^{-1}(y_t)); t \in G_1\}$. It is easy to see that this plane can go through an additional observation point $(x_k; \ln(y_k)); k \notin G_1$; and that there are no other points under this plane (see Fig. 3). Since the $m - 1$ vectors from G_1 are linearly independent, the plane P; which is specified by the vector β; exists: For this plane

$$S(\beta) = \sum_{t \notin G_1} [y_t - g(\beta' x_t)]^2 - [y_k - g(\beta' x_k)]^2 < S_E$$

which proves the existence of such a vector $a_0 = \beta$ that $S(a_0) < S_E$:

In the next section, two examples of quasilinear regression will be considered.

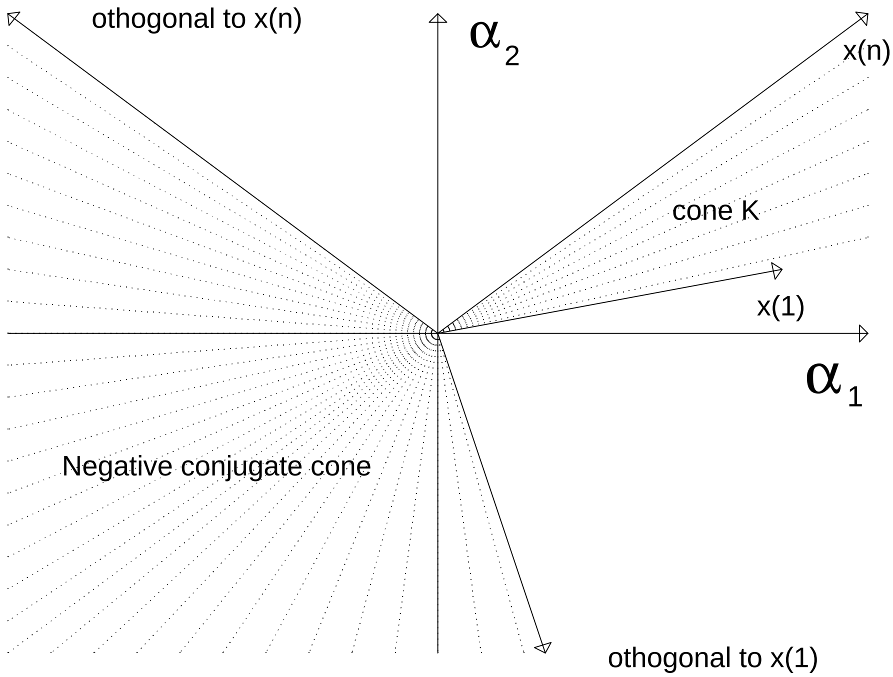

FIG 2. The pair of cones for a two-parameter quasilinear regression model.

4. SIMPLE EXPONENTIAL MODELS

4.1. Two-parameter exponential model

The two-parameter exponential model is a particular case of quasilinear regression with $g = \exp$ and $x_t = (1; x_t)'$:

$$f_t(\alpha_1; \alpha_2) = \exp(\alpha_1 + \alpha_2 x_t); \qquad t = 1; \ldots; n \tag{4.1}$$

where x_t can be an increasing function of time t. We assume that $\sum (x_t - \bar{x})^2 > 0$ and some y_t are allowed to be negative. It should be noticed that for the model (4.1) the multiplicative error term is usually assumed: $y_t = \exp(\alpha_1 + \alpha_2 x_t)\exp(\varepsilon_t)$; that leads to a linear model $\ln(y_t) = \alpha_1 + \alpha_2 x_t + \varepsilon_t$. However, sometimes additive errors are more appropriate. In particular, (4.1) could be derived from the model $y_t^* = c + \exp(\alpha_1 + \alpha_2 x_t) + \varepsilon_t$ with a known constant c. Then putting $y_t = y_t^* - c$, it is quite possible to encounter negative values for y_t. (The latter model with unknown c will be considered in the following section.)

In terms of quasilinear regression for model (4.1), $g_* = 0$ and vectors $x_t = \{(1; x_t)'\}$ are unidirectional, so that we have Case B. Without loss of generality we can assume that observations are indexed in such a way that $x_1 \le \ldots \le x_n$. For the sake of simplicity we also assume that $x_1 < x_2$ and $x_{n-1} < x_n$. In terms of the previous section these assumptions imply that the cone K is spanned by $\{(1; x_t)'; t = 1; \ldots; n\}$ and has two edges: $(1; x_n)'$ and $(1; x_1)'$. The negative conjugate cone is spanned by the pair of vectors $(x_1^\perp; x_n^\perp)$ (see Figure 2). If $\alpha_k = (\alpha_{1k}; \alpha_{2k})'$ lies in the interior of K^- then $f_t \to 0$ and $S_k \to \sum_1^n y_t^2$ when $\|\alpha_k\|^2 \to \infty$. If $y_1 > 0$ then we can reduce the limit of the SS to $\sum_2^n y_t^2$. If the direction coincides with the second edge and $y_n > 0$ then the limit of the SS can be reduced to $\sum_1^{n-1} y_t^2$. Thus, for the two-parameter exponential model the limit regression functions are:

$$\text{a) } \bar{f}_t(\beta) = \begin{cases} \beta; & \text{if } t = 1 \\ 0 & \text{if } t > 1 \end{cases}, \quad \text{b) } \bar{f}_t(\beta) = \begin{cases} 0 & \text{if } t < n \\ \beta & \text{if } t = n \end{cases}. \tag{4.2}$$

The two SS are thus equal to

$$S_1 = \begin{cases} \sum_1^n y_t^2; & \text{if } y_1 < 0 \\ \sum_2^n y_t^2 & \text{if } y_1 \ge 0 \end{cases} \quad \text{and} \quad S_2 = \begin{cases} \sum_1^n y_t^2; & \text{if } y_n < 0 \\ \sum_1^{n-1} y_t^2 & \text{if } y_n \ge 0 \end{cases}.$$

Thus, the EL for the two-parameter exponential models is equal to

$$S_E = \min\{S_1; S_2\}. \tag{4.3}$$

Now let us consider the problem of finding a suitable initial guess. First, let us consider the case when all the observations y_t are positive. Let $y_1 \le y_n$. Then for the following pair

$$a_2 = \max_{t<n}[\ln(y_n) - \ln(y_t)]/(x_n - x_t); \quad a_1 = \ln(y_n) - a_2 x_n \tag{4.4}$$

it is easy to see that $S(a_1, a_2) < S_E$. The solution (4.4) has a geometrical interpretation: it corresponds to the straight line which goes through the point $(x_n, \ln(y_n))$ and another point $(x_k, \ln(y_k)), k < n$ in such a way that all other observation points lie either on the line or above it (see Figure 3). As a consequence, if a minimization procedure starts with initial values (4.4), the sequence (a_{1k}, a_{2k}) has at least one limit point. It is interesting to note that $\mathbf{a}_{OLS}$, as the pair of the OLS-coefficients in regressing $\ln(y_t)$ on x_t, does not necessarily satisfy the inequality $S(\mathbf{a}_{OLS}) < S_E$ and may generate a disconvergent sequence of parameters.

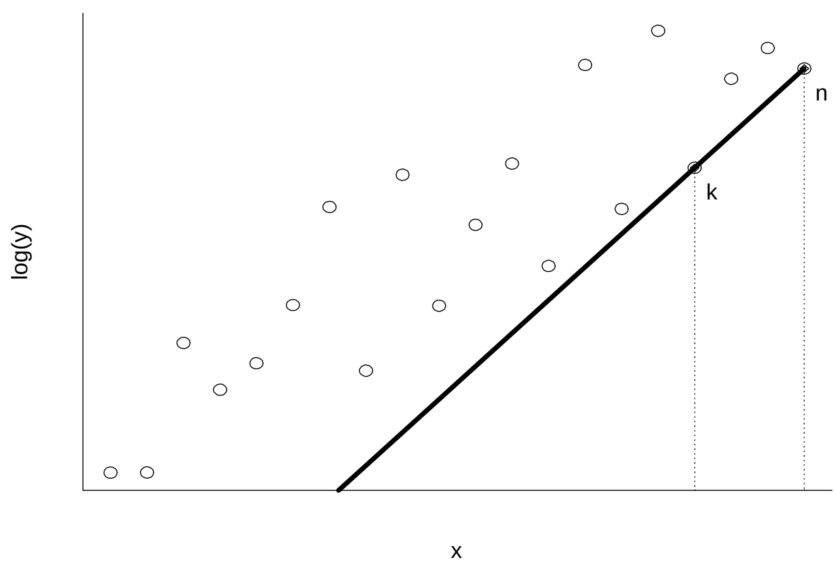

FIG 3. The existence of a suitable starting value for the simple exponential model. The line goes through two points n and k, and there are no observation points under the line.

Clearly, if all observations y_t are positive the LSE exists. In the presence of negative observations the following criterion can be formulated (for proof, see the Appendix).

Criterion for the LSE existence for the exponential model. If one of the two following conditions are fulfilled

$$\text{a) } y_n > y_1;\ y_n > 0;\ y_{n-1} > 0 \quad \text{or} \quad \text{b) } y_1 > y_n;\ \ y_1 > 0;\ y_2 > 0;$$

then the LSE for the exponential growth curve model (4.1) exists.

Based on this criterion it is possible to provide some approximation for the probability that the LSE exists when $n \to \infty$. Let us assume that the error term is normally distributed, in other words, $y_t \sim N(e^{\alpha_1+\alpha_2 x_t}; \sigma^2)$: Let us also assume that $0 < x_1 < \dots < x_n$ and $\alpha_2 > 0$: Then, as follows from the criterion for the existence

$$\begin{aligned} P(\text{LSE exists}) &\geq P\{(y_1 > 0;\ y_2 > 0) \cup (y_{n-1} > 0;\ y_n > 0)\} \\ &\geq P(y_{n-1} > 0;\ y_n > 0) = P(y_{n-1} > 0)P(y_n > 0) \end{aligned} \tag{4.5}$$

because the $\{y_t\}$ are independent. Let $\Phi(\cdot)$ denote the normal distribution function. Then

$$P(y_n > 0) = 1 - \Phi(-\sigma^{-1}e^{\alpha_1+\alpha_2 x_n});\ \ P(y_{n-1} > 0) = 1 - \Phi(-\sigma^{-1}e^{\alpha_1+\alpha_2 x_{n-1}}):$$

If $x_{n-1} \to \infty$ and $x_n \to \infty$ when $n \to \infty$; then as follows from (4.5) the probability that the LSE exists approaches 1 very rapidly, as implied by the fact $\Phi(-s) \simeq \phi(s)/s$ for large s (e.g. Feller, 1957). It is interesting to note that if $x_n = n$ and parameter α_2 is negative the LSE is inconsistent:

Furthermore, for the two-parameter exponential model the following simple criterion can be formulated.

Criterion for the nonexistence of the LSE for the exponential model. If $y_1 \leq 0; \dots; y_{n-1} \leq 0$ or $y_2 \leq 0; \dots; y_n \leq 0$ then the LSE does not exist.

The proof follows from the following inequality:

$$S(\alpha) = \sum_1 [y_t - \exp(\alpha_1 + \alpha_2 x_t)]^2 \geq \sum_1^{n-1} [\exp(\alpha_1 + \alpha_2 x_t) - y_t]^2 > \sum_1^{n-1} y_t^2 = S_E:$$

4.2. Three-parameter exponential model

Another particular case of (4.1) is the quadratic exponential curve with three parameters

$$f_t(\alpha_1; \alpha_2; \alpha_3) = \exp(\alpha_1 + \alpha_2 x_t + \alpha_3 x_t^2)\ ; \quad t = 1; \dots; n > 3:$$

For the sake of simplicity we assume that observations on x_t are different and ordered: $x_1 < x_2 < \dots < x_{n-1} < x_n$; and matrix X with tth row $x_t = (1; x_t; x_t^2)'$ has full rank $3 < n$ (some observations y_t may be negative). The cone K defined in the previous section belongs to the three dimensional space and has n planes spanned by the pair of neighboring vectors $(x_t; x_{t+1})$ where by convenience we let $x_{n+1} = x_1$ and $y_{n+1} = y_1$. Also, the observations y_t corresponding to the edges of K are assumed nonnegative. Based on the results of section 3 we can state that the EL for this model is

$$S_E = \sum_{1} y_t^2 - \max_j (y_j^2 + y_{j+1}^2). \tag{4.6}$$

To find the suitable initial vector of parameters with the property $S(a_0) < S_E$ we again draw n points $(x_t; \ln(y_t))$. One can easily prove that for this set of points there exists a parabola that goes through the neighboring points k and $k+1$, some other point p; and all other points lie either on or above the parabola (see Figure 4). With (4.6) in mind it becomes clear that this choice satisfies the desired inequality $S(a_0) < S_E$.

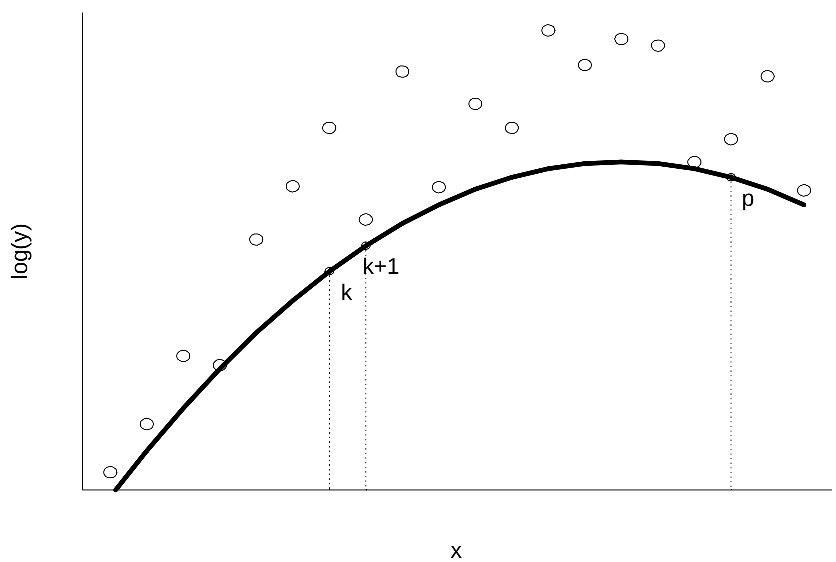

FIG 4. The existence of a suitable starting value for a quadratic exponential model. The parabola goes through three points $n; k$ and $k+1$ and there are no observation points under the parabola.

5. MODIFIED EXPONENTIAL MODEL AND GOMPERTZ CURVE

In this section we shall first study the problem of the existence of a LSE for the so-called "modified exponential model", which is a generalization of the exponential model considered in the previous section. Then based on the derived formula for the modified exponential model, it is easy to provide the EL for the Gompertz curve.

The regression function of the modified exponential model has the form

$$f_t(\alpha_1;\alpha_2;\alpha_3) = \alpha_1 - \exp(\alpha_2 + \alpha_3 x_t); \qquad t = 1;\ldots;n \tag{5.1}$$

where $\alpha = (\alpha_1;\alpha_2;\alpha_3) \in R^3$, and x_t are fixed numbers. Sometimes this curve is called the "monomolecular growth model" (Seber & Wild, 1989). By reindexing the observations, without loss of generality, we can assume $x_1 \le x_2 \ldots \le x_n$. In addition it is assumed that $x_1 < x_2$ and $x_{n-1} < x_n$.

Our first aim is to calculate the existence level by formula (3.2). So let $\alpha_k = (\alpha_{1k};\alpha_{2k};\alpha_{3k})'$ and $\| \alpha_k \| \to \infty$. We are looking for minimal limit values of the sum of squares over all disconvergent sequences $\| \alpha_k \| \to \infty$: The index k will be suppressed. There are three possibilities:

1. $\alpha_1 \to \pm\infty; \alpha_2^2 + \alpha_3^2$ is bounded,
2. $\alpha_1 \to \pm\infty; \alpha_2^2 + \alpha_3^2 \to +\infty$,
3. α_1 is bounded, $\alpha_2^2 + \alpha_3^2 \to +\infty$.

We will examine each case separately. It is obvious that in the first case $S(\alpha) \to +\infty$. To make the limit SS finite in the second case we have to put $\alpha_3 = 0$ with the minimal value

$$S_1 = \sum_{1} (y_t - \bar{y})^2: \tag{5.2}$$

The third case needs more analysis. For the simple exponential model we consider the cone K spanned by $(1; x_t)'$. The minimal limit value of the SS is achieved when the sequence of $(\alpha_2;\alpha_3)$ approaches one of the edges of the negative conjugate cone

K^- spanned by the pair (x_1^*, x_n^*) (see Figure 2). The two limit regression functions are:

$$\text{a) } \bar f_t(\alpha_1, c) = \alpha_1 - \begin{cases} c, & \text{if } t = 1 \\ 0 & \text{if } t > 1 \end{cases} \qquad \text{b) } \bar f_t(\alpha_1, c) = \alpha_1 - \begin{cases} 0 & \text{if } t < n \\ c & \text{if } t = n \end{cases}. \tag{5.3}$$

where $c \geq 0$ is a parameter. The minimal limit of the SS in the first case is

$$\min_{\alpha_1, c \geq 0} \left\{ (y_1 - \alpha_1 + c)^2 + \sum_2^n (y_t - \alpha_1)^2 \right\}. \tag{5.4}$$

Let us denote $\bar y_n = \sum_2^n y_t/(n-1)$. It is easy to verify that if $y_1 \leq \bar y_n$ then (5.4) is equal to $\sum_2^n (y_t - \bar y_n)^2$; otherwise it is equal to (5.2). So, we denote

$$S_2 = \begin{cases} \sum_2^n (y_t - \bar y_n)^2 & \text{if } y_1 \leq \bar y_n \\ S_1 & \text{otherwise} \end{cases}. \tag{5.5}$$

A similar result is obtained for the second limit regression function,

$$S_3 = \begin{cases} \sum_1^{n-1} (y_t - \bar y_1)^2 & \text{if } y_n \leq \bar y_1 \\ S_1 & \text{otherwise} \end{cases} \tag{5.6}$$

where $\bar y_1 = \sum_1^{n-1} y_t/(n-1)$. Finally, the EL for the modified exponential model is $S_E = \min\{S_2, S_3\}$. Based on this formula it is possible to establish conditions for the LSE existence.

Criterion for the LSE existence for modified exponential model. If $y_1 < y_n, y_1 < \bar y_n, y_2 < \bar y_n$ or $y_n < y_1, y_n < \bar y_1, y_{n-1} < \bar y_1$ then the LSE exists.

Proof. See the Appendix.

It is easy to generalize the modified exponential model to a multiparameter version: $f_t(\alpha_0, \alpha) = \alpha_0 + \exp(\alpha' x_t)$, where $\alpha, x_t \in R^m$. Repeating the technique and using arguments from sections 3 and considering the same three cases of the disconvergence it is easy to derive the existence level and formulate sufficient conditions for the LSE existence.

Finally, let us consider the Gompertz curve which can be written as

$$f_t(\alpha_1, \alpha_2, \alpha_3) = \exp\left(\alpha_1 - \exp(\alpha_2 + \alpha_3 x_t)\right), \qquad t = 1, \ldots, n.$$

Because it is the exponent of the modified exponential curve, we can directly apply all results derived for the model (5.1). In particular, assuming that all observations are positive the limit regression functions for the Gompertz curve are exponents of (5.3) with the same EL and criterion for existence.

6. LOGISTIC MODEL

In this section we consider another popular type of nonlinear growth curve, the logistic model:

$$f_t(\alpha_1;\alpha_2;\alpha_3) = \frac{\alpha_1}{1+\exp(\alpha_2 - \alpha_3 x_t)}. \tag{6.1}$$

It is assumed that $x_1 < x_2 \le \cdots \le x_{n-1} < x_n$, and all observations on the growth are positive: $y_t > 0$, $t = 1,\ldots,n$. As in the case of the modified exponential model there are three possibilities for disconvergence:

1. $\alpha_1 \to \pm\infty$, $\alpha_2^2+\alpha_3^2$ is bounded,
2. $\alpha_1 \to \pm\infty$, $\alpha_2^2+\alpha_3^2 \to +\infty$,
3. α_1 is bounded, $\alpha_2^2+\alpha_3^2 \to +\infty$.

It is clear that in the first case $S(\alpha) \to \infty$ because $f_t \to \pm\infty$. In the second case if $\alpha_1 \to -\infty$, then $f_t \le 0$ and $\underline{\lim} S(\alpha) \ge \sum_1^n y_t^2$. Let us consider the case $\alpha_1 \to +\infty$. If we put $\alpha_1 = \exp(\alpha_2)\exp(\beta_1)$ and $\alpha_2 \to +\infty$, where β_1 is an arbitrary number, the limit regression is

$$\lim_{\alpha_2\to\infty} f_t = \lim_{\alpha_2\to\infty} \frac{\exp(\beta_1)}{\exp(-\alpha_2)+\exp(-\alpha_3 x_t)} = \exp(\beta_1+\alpha_3 x_t) \equiv \bar f_t(\beta_1;\alpha_3) \tag{6.2}$$

with the minimal limit value of the SS

$$S_4 = \min_{\beta_1;\alpha_3} \sum_1^n (y_t - \exp(\beta_1+\alpha_3 x_t))^2. \tag{6.3}$$

Thus, at infinity, the logistic model includes the exponential model as a special case. The third case is standard: if asymptotic directions of $(\alpha_2;\alpha_3)$ are interior to K^- then the minimal limit regression functions are:

a) $\bar f_t(\alpha_1;c) = \begin{cases} \alpha_1/(1+c), & \text{if } t = 1 \\ \alpha_1 & \text{if } t > 1 \end{cases}$ b) $\bar f_t(\alpha_1;c) = \begin{cases} \alpha_1 & \text{if } t < n \\ \alpha_1/(1+c) & \text{if } t = n \end{cases}$

where $c \ge 0$. Therefore, $S_E = \min\{S_2;S_3;S_4\}$. Thus, if for some a_0 one has $S(a_0) < S_E$ then the level set $S_0 = \{\alpha : S(\alpha) \le S(a_0)\}$ is compact and the LSE exists.

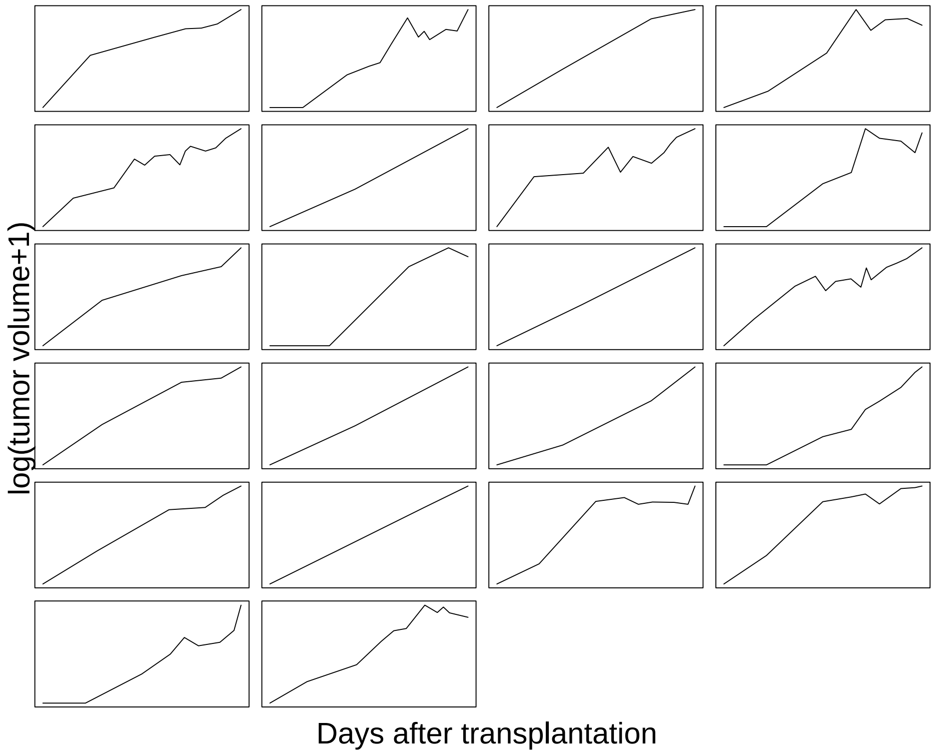

FIG 5. The data set for tumor growth of 22 rats.

7. EXAMPLE

We adapt an example described in Stukel & Demidenko (1995) on modelling the tumor growth of pancreatic carcinoma in 22 laboratory rats. The data set consists of 22 time series where time is measured in days and the growth is measured as $y_t = \log(v_t + 1)$ where v_t is the tumor's volume at time x_t (see Figure 5). Since at least theoretically, the level of growth has an asymptote, the logistic model (6.1) was chosen to model the individual tumor growth curves. Due to the short length of time series we face the lack of convergence when minimizing the SS for approximately one third of all rats. The analysis in the previous section can help us to understand why. Let us consider one of the mice with data $x = (28,32,34,36,39,41,42)$ and $y = (1.143426, 1.913239, 2.577790, 2.240843, 2.388763, 2.848913, 3.846417)$. Many attempts to estimate the parameters of the logistic model for this data fail due to lack of convergence. In particular, parameter α_1 tends to a very large value. After about 100 iterations we obtained $a = (207.6; 6.79; 0.0642)$ with $S(a) = 0.861.$ To understand the reasons for non-convergence, the profile SS was computed. Since the parameter α_1 is linear in (6.1) it can be removed by substituting its OLS-estimate, which leads to the profile SS

$$\tilde{S}(\alpha_2; \alpha_3) = \sum (y_t - \hat{\alpha}_1 z_t)^2 \tag{7.1}$$

where

$$z_t = \frac{1}{1 + \exp(\alpha_2 - \alpha_3 x_t)}; \quad \hat{\alpha}_1 = \frac{\sum y_t z_t}{\sum z_t^2}.$$

The contour plot for $\tilde{S}$ is shown on the left graph of Figure 6.

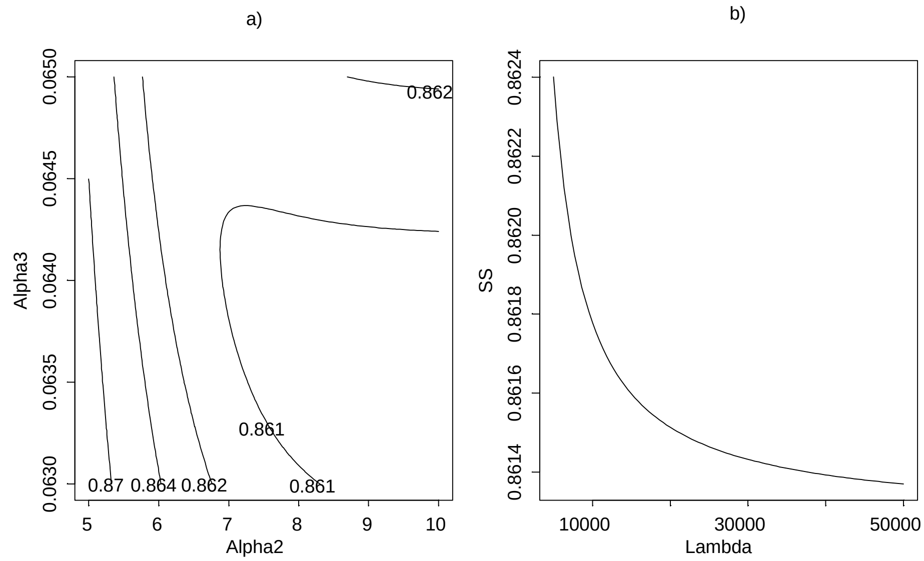

FIG 6. The LSE for the logistic model does not exist. a) a typical contours plot of the profile sum of squares (7.2), b) the SS (7.2) vs λ; the convergence takes place from above.

As one can see, the contour corresponding to 0.8610 goes to the right and possibly to infinity, which might be the case if the LSE did not exist. To check this an exponential model (6.2) was fitted with parameters $(.2361; .0634)$ and $S_4 = 0.8606 < 0.8610$ which is the first sign that the LSE for the logistic model does not exist. To confirm this another sum of squares was calculated

$$S(\lambda) = \sum_{1} \left(y_t - \frac{0.2361\lambda}{1 + \lambda \exp(-.0634 x_t)} \right)^2 \tag{7.2}$$

shown on the right side of Figure 6. As we expect $S(\lambda) \to 0.8606$. It is important that this convergence is from above, not below. Therefore, we conclude that the LSE for this logistic model does not exist.

8. APPENDIX. PROOFS

Formula (3.5). We consider successive directions of $\{\alpha_k, \|\alpha_k\| \to \infty\}$ which produce finite limits of the SS. First we observe that because $\{x_t\}$ are unidirectional, $K^- \neq \{0\}$ and $K^- \neq R^m$. The reader can refer to Fig. 2 for the geometrical interpretation. Obviously, if $\{\alpha_k\}$ belong to K^- then the asymptotic limit of $S_k = S(\alpha_k)$ is finite, otherwise not. Our task is to find directions in this cone which produce the minimal asymptotic limit of S_k. If an asymptotic direction belongs to the interior of K^-, or more precisely if $\cos(x_t, \alpha_k) \le -\delta < 0$ for all $k > k_0$, then

$$x_t'\alpha_k = \|x_t\|\|\alpha_k\| \cdot \cos(x_t, \alpha_k) \le -\delta\|x_t\|\|\alpha_k\| \to -\infty, \quad t = 1, \ldots, n$$

and $S_k \to \sum_1^n (y_t - g_*)^2$. A specific case is when the asymptotic direction coincides with one of the edges of cone K^-. Then we could lower the limit value of S_k in the following way. Let us denote this direction as γ; then, as was mentioned before, γ is orthogonal to some $m-1$ edges of K. We denote this as the set G_1. Formally it means

$$x_t'\gamma = 0, \; t \in G_1 \qquad \text{and} \qquad x_t'\gamma < 0, \; t \notin G_1. \tag{8.1}$$

As follows from (8.1) $\lim g(x_t'\alpha_k) = g_*$ for all $t \in G_1$ when $\|\alpha_k\| \to \infty$, and therefore

$$\underline{\lim}_k \sum_1 (y_t - g(x_t'\alpha_k))^2 \ge \sum_{t \notin G_1} (y_t - g_*)^2. \tag{8.2}$$

Now we shall prove that for a specific choice of $\{\alpha_k\}$ the inequality (8.2) becomes an equality. For simplicity we assume that for all observations y_t which correspond to the edges of cone K we have $y_t > g_*$. Let us consider the ray $\alpha_k = \lambda_k\gamma + r$, where $\lambda_k \to +\infty$ and $r \in R^m$. For all $t \in G_1$ because of (8.1) we have $\alpha_k' x = \lambda_k\gamma' x_t + r'x_t = r'x_t$. Furthermore, we can choose the beginning of the ray, r so that

$$g(r'x_t) = y_t, \quad t \in G_1. \tag{8.3}$$

By the previous assumption $y_t > g_*$ and vectors $\{x_t; t \in G_1\}$ are linearly independent. Therefore the system (8.3) with $m-1$ linear equations has a solution. This means that the inequality (8.2) becomes an equality. Furthermore, we can consider other groups $G_2, \ldots, G_P$ and select the one that leads to the minimal asymptotic value of the SS.

Criterion for the LSE existence for the quasilinear regression. Let the EL (3.5) be attained on the ray $\alpha = \lambda\beta + r$ where $r'x_t = g^{-1}(y_t)$ and $x_t'\beta = 0$ for $t \in G_1$ (see the previous proof). Then for this ray $S(\lambda) \equiv S(\lambda\beta + r) \to S_E = \sum_{t\notin G_1}(y_t - g_*)^2$ when $\lambda \to +\infty$. Our current aim is to show that $S(\lambda)$ approaches S_E from below. We consider the difference

$$\begin{aligned} S(\lambda) - S_E &= \sum_1^n [y_t - g(\lambda\beta' x_t + r' x_t)]^2 - \sum_{t\notin G_1}(y_t - g_*)^2 \\ &= -2\sum_{t\notin G_1}(y_t - g_*)[g(\lambda\tau_t + \theta_t) - g_*] + \sum_{t\notin G_1}[g(\lambda\tau_t + \theta_t) - g_*]^2 \end{aligned} \tag{8.4}$$

where it is denoted $\tau_t = \beta' x_t$ and $\theta_t = r' x_t$; $t = 1, \ldots, n$. The difference (8.4) converges to zero if $\lambda \to \infty$ because $\tau_t < 0$ for $t \notin G_1$ and $g(s) \to g_*$ when $s \to -\infty$. We shall prove that there is a λ_0 such that $S(\lambda_0) < S_E$. Let $\tau_k = \max(\tau_t; t \notin G_1)$. Then (8.4) can be rewritten as

$$S(\lambda) - S_E = [g(\lambda\tau_k + \theta_k) - g_*]\frac{S(\lambda) - S_E}{g(\lambda\tau_k + \theta_k) - g_*} = [g(\lambda\tau_k + \theta_k) - g_*]P(\lambda);$$

where

$$P(\lambda) = -2\sum_{t\notin G_1}(y_t - g_*)\frac{g(\lambda\tau_t + \theta_t) - g_*}{g(\lambda\tau_k + \theta_k) - g_*} + \sum_{t\notin G_1}\frac{[g(\lambda\tau_t + \theta_t) - g_*]^2}{[g(\lambda\tau_k + \theta_k) - g_*]^2}.$$

It is easy to see that the each term in the second sum of P tends to zero with a higher order convergence than the corresponding term of the first sum. Since the first term is negative we conclude that $P(\lambda) < 0$ for sufficient large λ: that means $S(\lambda)$ converges to S_E from below.

Criterion for the LSE existence for the exponential model. We prove the case a); the proof of b) is quite similar. As follows from the previous considerations, the ray which corresponds to the existence level (4.3) is: $\alpha_1(\lambda) = -\lambda x_n + \ln(y_n)$; $\alpha_2(\lambda) = \lambda \to +\infty$. The sum of squares along this ray is written as

$$S(\lambda) = \sum_{t=1}^{n-1}[y_t - y_n \exp(-\lambda(x_n - x_t))]^2.$$

For the existence of the LSE it suffices to prove that there exist λ_0 such that for all $\lambda > \lambda_0$

$$dS/d\lambda = 2y_n \sum_1^{n-1} [y_t - y_n \exp(-\lambda(x_n - x_t))] \cdot (x_n - x_t) \exp(-\lambda(x_n - x_t)) > 0.$$

It is obvious that the sign of this derivative is defined by the term $t = n - 1$ when $\lambda \to \infty.$

Criterion for the LSE existence for the modified exponential model. We prove the LSE existence in the case $y_1 < y_n; y_1 < \bar{y}_n; y_2 < \bar{y}_n$. We fix $a_{01} = \bar{y}_n$ and consider the SS for the exponential curve (5.1) with $y_t^* = \bar{y}_n - y_t;\ t = 1; ...; n$. It is easy to see that $y_1^* > y_n^*; y_1^* > 0$ and $y_2^* > 0$ by this assumption. Thus, as follows from the criterion for the LSE existence for the simple exponential model, there exists a pair $a_0 = (a_{02}; a_{03})$ that

$$S(a_0) = \sum [\exp(a_{02} + a_{03}x_t) - y_t^*]^2 = \sum [\exp(a_{02} + a_{03}x_t) + y_t - \bar{y}_n]^2$$
$$< \sum_2^n (y_t^*)^2 = \sum_2^n (y_t - \bar{y}_1)^2 = S_E.$$

Therefore, from the General criterion for the global minimizer existence it follows the LSE for the modified exponential model exists. The proof is similar for the second case.

ACKNOWLEDGMENT

This work was supported by National Cancer Institute grants CA52192 and CA47327.

References

[1] Albert, A.and Anderson, J.A. (1984). "On existence of maximum likelihood estimates in logistic regression model", Biometrika, 71, 1-10.

[2] Cattle, R.W. and Pang, J.-S. and Stone R.E. (1992). The Linear Complementarity Problem, Academic Press, New York.

[3] Demidenko, E. (1981). Linear and Nonlinear Regression (in Russian), Finansy and Statistika, Moscow.

[4] Demidenko, E. (1989). Optimization and Regression (in Russian), Nauka, Moscow.

[5] Edwards, C.H. (1973). Advanced Calculus of Several Variables, Academic Press: New York.

[6] Feller, W. (1957). An Introduction to Probability Theory and its Application. v.1, Wiley: New York.

[7] Kamthan, P.G. and Gupta, M. (1985). Theory of Basis and Cones, Pitman Advanced Publications, Boston.

[8] Kubicek, M. and Marek, M. and Eckert, E. (1971). "Quasilinear regression", Technometrics, 13.

[9] Lesa®re, E. and Kaufmann, H.K. (1992). "Existence and uniqueness of the maximum likelihood estimator for a multivariate probit model", JASA, 87, 805-811.

[10] Makelainen, T. and Schmidt, K. and Styan, G. (1981). "On the existence and uniqueness of the maximum likelihood estimate of a vector-valued parameter in ¯xed-size samples", The Annals of Statistics, 9, 758-767.

[11] Nakamura, T. (1984). "Existence theorems of a maximum likelihood estimate from a generalized censored data sample", Ann. Inst. Statist. Math, 36, 375-393.

[12] Nakamura, T. and Lee, C (1993). "On the existence of minimum contrast estimates in binary response model", Ann. Inst. Statist. Math, 45, 741{758.

[13] Stukel, T. and Demidenko, E. (1995). "E±cient estimation for general linear growth curves models", Submitted.

[14] Seber, G.A.F. and Wild, C.J. (1989). Nonlinear Regression. John Wiley: New York.

[15] Wedderburn, R.W.M. (1976). "On existence and uniqueness of the maximum likelihood estimates for certain generalized linear models". Biometrika, 63, 27-32.