Statistics and Probability Letters 120 (2017) 87–94

Contents lists available at ScienceDirect

Statistics and Probability Letters

journal homepage: www.elsevier.com/locate/stapro

Constructing initial estimators in one-step estimation procedures of nonlinear regression[☆]

Yu.Yu. Linke, I.S. Borisov[*]

Sobolev Institute of Mathematics, Novosibirsk State University, Novosibirsk, 630090, Russia

ARTICLE INFO

Article history:
Received 5 January 2016
Received in revised form 21 September 2016
Accepted 22 September 2016
Available online 3 October 2016

Keywords:
Nonlinear regression
One-step M-estimator
Initial estimator
α_n-consistency
Asymptotic normality

ABSTRACT

We discuss an approach to construct explicitly calculable consistent estimators for parameters of some nonlinear regression models. The estimators of such a kind can be used as initial estimators in one-step estimation procedures for unknown parameters of these models.

1. Introduction

We consider the classical model of nonlinear regression when one needs to estimate an unknown parameter θ (for simplicity, we assume here that $\theta \in \mathbb{R}$) based on observations $X_1, \ldots, X_n$ of the form

$$X_i = f_i(\theta) + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

where $\{f_i(\cdot)\}$ are known functions and the errors $\{\varepsilon_i\}$ form a sequence of independent random variables with zero mean. To estimate the parameter of such a model there are some important methods like the least-squares procedure and its modifications, the quasi-likelihood and maximal likelihood methods, and some others, which are connected with the so-called M-estimation procedures. By an M-estimator we mean a statistic which is a solution to the equation (with respect to t)

$$\sum_{i=1}^{n} M_i(t, X_i) = 0, \tag{2}$$

where $\{M_i(t,x)\}$ are some functions, with $\mathbf{E}M_i(\theta, X_i) = 0$ for all i. Usually the choice of the functions $\{M_i(t,x)\}$ guarantees that the corresponding M-estimators possess required properties. For example, if $\mathbf{E}\varepsilon_i^2 = \sigma^2$ (the parameter σ^2 may be unknown) then, under some regularity conditions, the least-squares estimator is defined by Eq. (2), with $M_i(t, X_i) = f_i'(t)(X_i - f_i(t))$. If $\mathbf{E}\varepsilon_i^2 = \sigma^2/w_i(\theta)$, where the functions $\{w_i(\cdot)\}$ are known, then the quasi-likelihood estimator is defined by

☆ Supported by the Russian Foundation for Basic Research grant 14–01–00220.
* Corresponding author.
E-mail addresses: linke@math.nsc.ru (Yu.Yu. Linke), sibam@math.nsc.ru (I.S. Borisov).

http://dx.doi.org/10.1016/j.spl.2016.09.022
0167-7152/

Eq. (2), with $M_i(t, X_i) = w_i(t)f_i'(t)(X_i - f_i(t))$. Asymptotic properties of such estimators are well known and they are often optimal in some sense (for example, see Heyde, 1997, Seber and Wild, 1989, Draper and Smith, 1998). The main problem of the above-mentioned methods is how to calculate a consistent estimator defined by Eq. (2). To solve this problem the most effective approaches are Newton-type iterative methods.

The problem of finding such implicit estimators by the use of iterative methods is connected with the possible existence of a large number of the roots of Eq. (2). Hence, if an initial point of an iterative process is too far from the parameter, then this procedure approximates the root of (2) closest to the initial point chosen.

The situation may be much simpler if we have a consistent estimator θ_n^* approximating the parameter with a suitable proximity, which can be used as an initial point of the iterative process under consideration. As a rule, we need only one step in a Newton-type iterative process to construct an explicit estimator possessing the same asymptotic properties as the corresponding M-estimator. A so-called *one-step M-estimator* θ_n^{**} is defined by the relation

$$\theta_n^{**} = \theta_n^* - \sum_{i=1}^n M_i(\theta_n^*, X_i) \Big/ \sum_{i=1}^n M_i'(\theta_n^*, X_i). \tag{3}$$

The estimator θ_n^{**} is a one-step approximation to a solution of (2) obtained with the help of Newton's method, where $t = \theta_n^*$ is the initial point. Under certain assumptions, this estimator is asymptotically normal with the same asymptotic variance as the M-estimator $\widetilde{\theta}_n$. The idea of such one-step estimation originates to R. Fisher who used a similar approach to construct explicit estimators equivalent to the corresponding implicit maximal likelihood estimators. Later the estimators of such a kind (like (3)) were called *one-step estimators*. The first one-step estimators for non-identically distributed observations go back to Bickel (1975), where special estimators for linear regression models were considered. Since this time one-step estimation procedures are successfully used in studying various special statistical problems (see Simpson et al., 1992, Müller, 1994, Fan and Chen, 1999, Welsh and Ronchetti, 2002, Zou and Li, 2008, Jurečková, 2012, Fan et al., 2014).

Thus, one of the key points of the above-mentioned one-step estimation methodology is presence of sufficiently good initial estimators. At present, only for a few particular types of nonlinear regression models, there are known consistent explicit estimators, for example, for the intrinsically linear models and for the linear-fractional regression models (see Draper and Smith, 1998, Linke and Sakhanenko, 2000), and these constructions essentially depend on the form of regression functions. We did not find in the literature any methods allowing to construct explicitly calculable statistical estimators for nonlinear regression parameters for sufficiently broad class of regression models.

The goal of the paper is to propose an approach to construct explicitly calculable consistent estimators for a broad class of nonlinear regression models, which can be considered as initial estimators for the above-mentioned one-step estimation procedures in such regression models.

2. Univariate parameter and nonrandom design

Consider the following generic regression model.

(**R**) *The response variables $\{X_i\}$ have the structure $X_i = f(\theta, z_i) + \varepsilon_i$, $i = 1, \dots, n$, where the numerical sequence $\{z_i\}$ (a design) and the function $f(\theta, z)$ are known, the mean-zero random errors $\{\varepsilon_i\}$ are independent, and $\theta \in \Theta = (a, b)$, where the interval boundaries a and b (one or both) may be infinite.*

In the paper we discuss an approach to construct consistent estimators of θ in Model (R). A similar approach was suggested by Müller (1988) for nonparametric regression models. This method is based on constructing special weighted sums of the original response variables $\{X_i\}$, which have the structure of Riemann sums for some integrals. To this end it is convenient to rewrite the values $z_1, \dots, z_n$ in increasing order, i.e., $z_{n:1} \le \dots \le z_{n:n}$. The corresponding renumbered values of the response variables and the errors will be denoted by $\{X_{ni}\}$ and $\{\varepsilon_{ni}\}$, respectively. In this case the response variable from (R) can be rewritten as follows: $X_{ni} = f(\theta, z_{n:i}) + \varepsilon_{ni}$, $i = 1, \dots, n$, where the mean-zero random errors $\{\varepsilon_{ni}\}$ are independent.

We first explain the main idea of the above-mentioned construction. In what follows we assume that the design $\{z_i\}$ is chosen from an interval $[c, d]$, where for simplicity only the value d may be infinite. Put $\Delta z_{ni} := z_{n:i} - z_{n:i-1}$, $z_{n:0} := c$, $z_{n:n+1} := d$ if $d < \infty$, and $z_{n:n+1} := z_{n:n}$ otherwise. So we have

$$\sum_{i=1}^n \Delta z_{ni} X_{ni} = \sum_{i=1}^n \Delta z_{ni} f(\theta, z_{n:i}) + \sum_{i=1}^n \Delta z_{ni}\varepsilon_{ni}. \tag{4}$$

We need the following important assumption.

(**C**) *The design $\{z_i\}$ forms a partition of* $[c, d]$ *such that* $\lim_{n\to\infty} \max_{i\le n+1} \Delta z_{ni} = 0$.

Notice that, in the case $d = \infty$, assumption (C) allows that $z_{n:n} \to \infty$. Under condition (C) and some additional minor restrictions,

$$\sum_{i=1}^n \Delta z_{ni} f(\theta, z_{n:i}) \to \int_c^d f(\theta, z)dz \equiv T(\theta) \quad \text{and} \quad \sum_{i=1}^n \Delta z_{ni}\varepsilon_{ni} \xrightarrow{p} 0;$$