Available online at www.sciencedirect.com
SCIENCE DIRECT
Computational Statistics & Data Analysis 49 (2005) 1020 – 1038
COMPUTATIONAL STATISTICS & DATA ANALYSIS
www.elsevier.com/locate/csda

Maximum likelihood estimation in nonlinear mixed effects models

E. Kuhn, M. Lavielle*

Université Paris Sud, Bât. 425, 91405 Orsay, France

Received 10 September 2003; received in revised form 15 July 2004; accepted 16 July 2004
Available online 12 August 2004

Abstract

A stochastic approximation version of EM for maximum likelihood estimation of a wide class of nonlinear mixed effects models is proposed. The main advantage of this algorithm is its ability to provide an estimator close to the MLE in very few iterations. The likelihood of the observations as well as the Fisher Information matrix can also be estimated by stochastic approximations. Numerical experiments allow to highlight the very good performances of the proposed method.

Keywords: Mixed effects model; Nonlinear model; Maximum likelihood estimation; EM algorithm; SAEM algorithm

1. Introduction

The mixed effects models were introduced mainly for modeling responses of individuals that have the same global behavior with individual variations (see the book of Pinheiro and Bates (2000) and the many references therein, for example). In fact, we consider that all the responses follow a common known functional form that depends on unknown effects. Some of them are fixed (i.e. the same for all the individuals), the others are random, so they depend on the individuals (or on sub-groups of the population). Then, the model has two types of parameters: global parameters that correspond to the fixed effects and parameters which vary among the population that correspond to the random effects. This kind of observations are usually the result of repeated measurements: some individuals are

* Corresponding author. Tel.: +31-1-69-15-57-43; fax: +31-1-69-15-72-34.
E-mail addresses: estelle.kuhn@math.u-psud.fr (E. Kuhn), marc.lavielle@math.u-psud.fr (M. Lavielle).

0167-9473/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.csda.2004.07.002

repeatedly observed under different experimental conditions. This approach seems to be adapted to many situations, particularly in the fields of pharmacokinetic, biological growth, epidemiology or econometry.

Let us consider here the following general nonlinear mixed effects model:

$$y_{ij} = g(\boldsymbol{\phi_i}, \boldsymbol{\beta}, x_{ij}) + h(\boldsymbol{\phi_i}, \boldsymbol{\beta}, x_{ij})\varepsilon_{ij}, \quad 1 \leqslant i \leqslant n, \quad 1 \leqslant j \leqslant m_i, \tag{1}$$

where y_{ij} is the jth observation of the ith individual, at some known instant x_{ij}. Here, n is the number of individuals and m_i is the number of observations of individual i. The within-group errors (ε_{ij}) are supposed to be i.i.d. Gaussian random variables with mean zero and unknown variance σ^2. The model is nonlinear means that g or h are nonlinear functions of $\boldsymbol{\phi_i}$. The random vector $\boldsymbol{\phi_i}$ is modelized by

$$\boldsymbol{\phi_i} = \boldsymbol{A_i}\boldsymbol{\mu} + \boldsymbol{\eta_i} \quad \text{with} \quad \boldsymbol{\eta_i} \sim_{\text{i.i.d.}} N(0, \boldsymbol{\Gamma}),$$

where $\boldsymbol{\mu}$ is an unknown vector of population parameters. The individual matrix $\boldsymbol{A_i}$ is supposed to be known. The vector $\boldsymbol{\beta}$ denotes also unknown population parameters, which do not appear in the random effect $\boldsymbol{\phi_i}$. We suppose that the ε_{ij} and the $\boldsymbol{\eta_i}$ are mutually independent.

Our purpose is to propose a method for computing the maximum likelihood estimate of the unknown parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\mu}, \boldsymbol{\Gamma}, \sigma^2)$ and to compare this method with other existing methods, particularly those based on the maximum likelihood approach.

In the case of a linear model, the estimation of the unknown parameters can be treated with the usual EM algorithm (Dempster et al., 1977) or with a Newton–Raphson algorithm (Pinheiro and Bates, 2000). A nonlinear function is often more suitable for modeling the physical problem, but requires a specific approach for estimating the parameters. Different methods, based generally on linearization of the log-likelihood, were suggested for dealing with nonlinear models. A Laplace approximation was proposed by Edward F. Vonesh in Vonesh (1996), a Bayesian approach was proposed by Racine-Poon (1985), Wakefield et al. (1994), Wakefield (1996). Walker (1996) uses a Monte-Carlo EM algorithm, whereas a simulated pseudo maximum likelihood estimator for these specific models is developed by Concordet and Nunez (2002).

In this paper, we show that the SAEM algorithm (stochastic approximation version of EM) is very efficient for computing the maximum likelihood estimate of $\boldsymbol{\theta}$. This iterative procedure consists at each iteration, in successively simulating the random effects with the conditional distribution, and updating the unknown parameters of the model. This algorithm was shown to converge under very general conditions by Delyon et al. (1999). When this algorithm is coupled with a MCMC procedure for the simulation step, Kuhn and Lavielle (2004) also established the convergence of the algorithm toward the MLE. Furthermore, the observed likelihood and the Fisher Information matrix can also be estimated by using also a stochastic approximation procedure. This method has the very nice advantage to converge very quickly to a neighborhood of the Maximum Likelihood Estimate. Then, only a few seconds are required for computing a MLE confidence interval, in any of the usual models used in the practice. The SAEM can be used for estimating homoscedastic models, but also heteroscedastic models. For the latter, the parameters related to the fixed effects are estimated in a Bayesian context in term of their expectations. By the way, the SAEM could also be used in an empirical Bayesian context for estimating the prior distribution of