ACTA UNIVERSITATIS APULENSIS No 11/2006

Proceedings of the *International Conference on Theory and Application of Mathematics and Informatics ICTAMI 2005* - Alba Iulia, Romania

A SYNTHESIS OF THE ESTIMATORS PROPERTIES FOR THE NONLINEAR MODELS WITH RANDOM PARAMETERS

Alina Constantinescu

Abstract. The models with random parameters are used for the mathematical design of repeated measures data. For this models, the parameters estimation is a uneasy problem. We deal with a classical parameter approach for the nonlinear model. We also consider that the parameters are estimated by optimization of the likelihood function.
In this paper, after a presentation of the various existent estimation methods, we state a properties synthesis of the obtained estimators. This synthesis cover a emptiness from the specialized literature in which, generally, the main accent is put on the numerical aspects of the methods.

2000 *Mathematics Subject Classification*: 62F12.

1. Introduction

There is a great interest, in the specialized literature, concerning the models with random parameters. Many authors had focussed their attention especially on nonlinear models because in this case the difficulty degree of parameters estimation is higher. The major difficulty is caused by the complexity of likelihood function. In this paper, we explore the development of research on the nonlinear models with random parameters in a classical parametric approach. For this case, the model is:

$$\begin{cases} Y_i = f(t, \beta_i) + \varepsilon_i \\ \beta_i = \beta + B_i \\ B_i \sim N_p(0, \Gamma) \quad i.i.d. \\ \varepsilon_i \sim N_p(0, \sigma_\varepsilon^2 Id_j) \quad i.i.d. \end{cases} \tag{1}$$

where I is the individuals set and J is the observations set ($i \in I$, $j \in J$).

The parameter vector which we must estimate, is:

$$\theta = \left(\beta, \Gamma, \sigma_\varepsilon^2 \right).$$

If we chose the maximum likelihood method for the parameter estimation, then the likelihood function has the following expression:

$$V(Y \mid \beta, \Gamma, \sigma_\varepsilon^2) = (2\pi)^{\frac{IJ+Ip}{2}} \sigma_\varepsilon^{-IJ} \mid \Gamma \mid^{-\frac{I}{2}}$$
$$\int exp\Big[\sum_{i=1}^{I}\Big(-\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{J}\Big[y_{ij} - f(t_j, \beta + B_i)\Big]^2 - \frac{1}{2}B_i^T \Gamma^{-1} B_i\Big)\Big] dB_i \quad (2)$$

and the log-likelihood function is:

$$l(Y \mid \beta, \Gamma, \sigma_\varepsilon^2) = \frac{IJ+Ip}{2} log(2\pi) - \frac{IJ}{2} log\sigma_\varepsilon^2 - \frac{I}{2} log \mid \Gamma \mid +$$
$$\sum_{i=1}^{I}\Big\{ log\Big[\int exp\Big(-\frac{1}{2\sigma_\varepsilon^2}\sum_{j=1}^{J}\Big[y_{ij} - f(t_j, \beta + B_i)\Big]^2 - \frac{1}{2}B_i^T \Gamma^{-1} B_i\Big) dB_i\Big]\Big\} \quad (3)$$

The optimum finding for the previous function is, in generally, a difficult job because there is an multiple integral. That is why various technics appeared for find a solution in this problem.
In the following section we present the existent estimation procedures based on the optimization of the likelihood function. In section 3 we analyze the properties of the estimators which are produced with this methods.

2. Estimation procedures which use the maximum likelihood method

The existent estimation methods which use the maximum likelihood criterion are: the Expectation Maximization algorithm (EM), Laplace's approximation, Monte Carlo, Gaussian Quadrature, Pseudo-likelihood and Linearisation.
The EM algorithm is the only one method which work out the maximum likelihood of the initial model. The general principle of this algorithm consists in calculation of the maximum of the posteori density (Dempster et al. [6]). The algorithm use an iterative scheme and yields an estimator sequence which converge to the maximum likelihood solution. The Laplace's approximation, Monte Carlo and Gaussian Quadrature are methods which maximizes an approximation of the initial model likelihood. In this methods are applied the classical numerical integration technics for the integral approximation (Dahlquist G., Bjorck A. et Andersen [3], Davis P.J. et Rabinowitz P.

[4]). The yielded estimators are those which maximizes an approximated initial model likelihood. The laplace's approximation consists in a Taylor series development and keeping the first and second terms. The Monte Carlo method consists in application of the Monte Carlo technic for the integral calculation and for the Gaussian Quadrature method is applied the quadrature formula for the integral. For these three previous methods the likelihood functions have the following expressions:

$$l_{lap}(\theta \mid Y) = -\frac{IJ}{2} log(2\pi\sigma_\varepsilon^2) - \frac{I}{2} log \mid Y \mid + \sum_{i=1}^{I} F(\beta, \theta, Y_i, \hat{B}_i) - \frac{1}{2}\sum_{i=1}^{I} log \mid F^{"}_{**}(\hat{B}_i) \mid$$

$$l_{MC}(\theta \mid Y) = -\frac{IJ}{2} log(2\pi\sigma_\varepsilon^2) - \frac{1}{2}\sum_{i=1}^{I} log \mid F^{"}_{**}(\hat{B}_i) \mid +$$
$$+ \sum_{i=1}^{I} log\Big\{\frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} exp\Big[F(\theta, Y_i, B_i^*) + \frac{1}{2} z_k^{*T} z_k^*\Big]\Big\}$$

$$l_{QG}(\theta \mid Y) = -\frac{IJ}{2} log(2\pi\sigma_\varepsilon^2) +$$
$$+ \sum_{i=1}^{I} \Big\{ \sum_{k_1=1}^{N_{QG}} \cdots \sum_{k_p=1}^{N_{QG}} exp\Big(-\frac{1}{2\sigma_\varepsilon^2} \sum_{j=1}^{J} \Big[Y_{ij} - f\Big(t_j, \beta + \Gamma^{\frac{1}{2}} z_{k_1,\cdots k_p}\Big)\Big]^2\Big) \prod_{l=1}^{p} \omega_{jl} \Big\}.$$

The Pseudo-likelihood and linearisation are the methods for which the initial model was modified into a model for which the likelihood function is not difficult. The Pseudo-likelihood was proposed by Concordet and Nuñez [2]. The linearisation method principle consists in approximation of the nonlinear model with one linear and thus the linear model technics are applicable (Searle et al. [12], Gaybill [7]). The linearisation consists in a Taylor series development and keeping the first term.

3. A synthesis of the estimators properties

We shall specify the possibility in which the asymptotic behavior can be attacked before we relate in detail the estimators properties.

So, in the repeated measures data framework we dispose of two directions for asymptoticity. One direction is I the number of individuals and the other direction is the number J of observations. Whit that circumstances, it is possible to taking into account the following convergence cases:

1. $I \to \infty$ and $J \to \infty$ when we have a disposable large number of individuals and observations;
2. $I \to \infty$ and $J < \infty$ when the number of observations is limited;
3. $I < \infty$ and $J \to \infty$ when the number of individuals is limited.

In the following exposition we shall reject the situation numbered whit three because, in generally, we can count on a large number of individuals (i.e. $I \to \infty$).

Few papers are dedicated to study the estimators properties in nonlinear models with random parameters. The references [1,5,11,13,15] are almost the only papers which treat that subject.

We propose an analyze of that few papers from the point of view of the following estimators properties:

- convergence;
- asymptotical normality;
- efficiency.

The EM algorithm implemented for nonlinear models with random parameters delivers the maximum likelihood estimator [15]. So, this properties are obvious. Among the methods which maximizes an numerical approximation of the likelihood function, the Laplace's method is the only one for which the statistical properties of estimators was studied. In [13] Vonesh studies the estimator from the Laplace's method. We think that Vonesh's proof for convergence is not explicitly and they rather calculate the convergence speed of the estimator.

The estimators obtained from the linearisation and pseudo-likelihood methods have properties which was proved in the specialized literature. Dispute on their large popularity, for the linearisation methods very few studies concerning the estimators properties are available. One paper by Denidenko [5] and one by Ramos et Patula [11] are the only two which study the properties of the estimators obtained from the linearisation method gave by Lindstrom et Bates [9]. Ramos et Patula [11] and Demidenko [5] use the same technic for non-convergence proof when J is finite and constant. For the second asymptotic situation ($I \to \infty$ and $J \to \infty$), Demidenko [5] proves that the estimator from the Lindstrom and Bates algorithm are equivalent with the maximum

likelihood estimator and then it has the convergence, normality and efficiency properties.
For pseudo-likelihood method, Concordet et *Nuñez* [2] deliver an estimator which has the convergence and asymptotic normality properties when $I \to \infty$. Much more, Nuñez gives the convergence speed for this estimator depending on the number of the individuals I and on the value of the Monte Carlo sample size N_{MC}. The following relation:

$$\frac{\sqrt{I}}{N_{MC}}\left(\theta_I^{N_{MC}} - \theta\right) \to^{n.c.}_{I\to\infty, N_{MC}\to\infty} 0$$

shows that the convergence is achieved when J is fixed. In the other words, is not necessary to dispose a large number of observations on individual.
The following table makes a sum of the estimators properties.

Table 1: The properties of estimators which are obtained from various methods: 1 and 2 marks the asymptotic situations for convergence, • marks a proved property, ○ marks the absence of the property demonstration and ? marks the absence of the result for the property.

Method of estimation	Convergence	Normality	Efficiency	References
EM	1	•	•	[15]
Pseudo-verosimility	1	•	○	[2]
Linearisation	2	•	•	[5]
Laplace	2?	•	•	[13]
Monte Carlo	?	?	?	
Gaussian Quadrature	?	?	?	

4. Conclusions

The presented synthesis may be used in a comparative studies when the parameter estimation is involved. Also, it shows which is the research level in that field.

References

[1] D. Concordet, *Modèle non linéaire à effects mixtes: quelques questions ouvertes*, in XXXIII Journées de Statistiques, (2001), pp. 43-48.

[2] D. Concordet et O. Nuñez, *A simulated pseudo-maximum likelihood estimator for nonlinear mixed models*, Computational Statistics and Data Analysis, (2002).

[3] G. Dahlquist, A. Bjorck et N. Andersen, *Numerical Methods*, Prentice-Hall, 1974, Prentice-Hall series in automatic computation.

[4] P.J. Davis et P. Rabinowitz, *Methods of Numerical Integration*, Academic Press, 1984, Computer Science and Applied Mathematics.

[5] E. Demidenko, *Asymptotic properties of nonlinear mixed-effects models*, Lecture notes in statistics, vol. 122, 1997, pp. 49-62.

[6] A.P. Dempster, N.M. Laird et D.B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society B, vol 39, (1977), pp. 1-38.

[7] F.A. Gabril, *Theory and application of the linear model*, Duxbury press, 1976.

[8] B. Lapeyre, E. Pardoux et R. Sentis, *Méhodes de Monte-Carlo pour les équations de transport et de diffusion*, Springer, 1998, Mathématiques & Applications.

[9] M.J. Lindstrom et D.M. Bates, *Nonlinear mixed effects models for repeated measures data*, Biometrics, vol. 46, (1990), pp. 673-687.

[10] J.C. Pinheiro et D.M. Bates, *Approximation to the log-likelihood function in the nonlinear mixed-effects models*, Journal of Computational and Graphical Statistics, vol.4, (1995), pp.12-35.

[11] R.Q. Ramos et S.G. Pantula, *Estimation of nonlinear random coefficient models*, Statistics & Probability Letters, vol. 24, (1995), pp. 49-56.

[12] S. Searle, G. Casella et McCulloch, *Variance components*, Wiley, New York, 1992.

[13] E.F. Vonesh, *A note of the use of laplace's approximation for nonlinear mixed-effects models*, Biometrica, vol. 83, (1996), pp. 447-452.

[14] E.F. Vonesh et R.L. Carter, *Mixed-effects nonlinear regresion for unbalanced repeated measures*, Biometrics, vol. 48, (1992), pp. 1-17.

[15] S. Walker, *An EM algorithm for nonlinear random effects models*, Biometrics, vol. 52, (1996), pp. 934-944.

Alina Constantinescu
Department of Mathematics,Valahia University of Târgovişte
Str. G. Enescu, No 1, Romania
email:*alinaconsta@yahoo.com*