Sobol's Sequence Based Method for Fitting Nonlinear Mixed Effects Model: A Comparative View

M. Rashedul Hoque[1] and A.H.M. Mahbub Latif[2]
[1, 2] Institute of Statistical Research and Training (ISRT), University of Dhaka, Dhaka-1000, Bangladesh
[1]Corresponding author: M. Rashedul Hoque, e-mail: rhoque@isrt.ac.bd

Abstract

Nonlinear mixed effects models are mixed effects models in which some of the fixed and random effects parameters enter nonlinearly to the model function. Nonlinear models are parsimonious so that we can capture the nonlinear variation with a minimum number of parameters. Due to their great importance, fitting of these models are also of crucial matters. A number of methods for fitting nonlinear mixed effects model are available in literature, most of the methods require approximating wither the model function or the likelihood function. A new method is proposed which numerically evaluate the integrations involved in the likelihood function with Monte Carlo integration using Sobol's sequence. The methods are compared using simulation studies and the method based on Laplace approximation is found to fit the nonlinear mixed effects model the best. The proposed Sobol's sequence based method performs better than some of the existing methods, especially in some cases; it produces good result in estimating random effects parameter. Thus, the Sobol's sequence based proposed method is very much compatible with the existing ones as well as the approximation based methods are quite handy.

Keywords: Laplace approximation, Lindstrom and Bates approximation, intractable integrations, quasi Monte Carlo integration.

1. Introduction

In nonlinear models responses are expressed as a nonlinear function of explanatory variables and are preferred over linear models due to parsimony by capturing nonlinear behavior of the system. There are some fixed effects parameters as well as some random effects parameters in mixed effects model. Such models can easily handle unbalanced repeated-measures data, can allow for flexible variance-covariance structures of the response vector, and are intuitively appealing. Thus nonlinear mixed effects model have properties of both nonlinear model and mixed effects model. There are two types of parameters in such models. These are the regression parameters associated with the fixed effects and the variance component parameters associated with the random effects. Regression parameters have subject-specific interpretation while variance component parameters have population average interpretation.

Through maximum likelihood and restricted maximum likelihood several estimation methods have been proposed for mixed effects models where the later is generally adopted for linear mixed effects models (Harville 1977). In recent years several different nonlinear mixed-effects models have been proposed (Sheiner and Beal 1980; Lindstrom and Bates 1990; Davidian and Gallant 1992; Vonesh and Carter 1992). Most commonly used one is proposed by Lindstrom and Bates (1990). There are different estimation methods for the parameters in the nonlinear mixed effects model (Davidian and Giltinan 1993) and there is an ongoing debate in the literature about the most adequate methods. One of the reasons for this variety of estimation methods is related to the numerical complexity involved in the derivation of REML estimates in the nonlinear mixed effects model. This complexity is due to the fact that the likelihood function in the nonlinear mixed effects model does not usually have a closed form expression. Different approximations have been proposed to try to circumvent this problem (Lindstrom and Bates 1990; Davidian and Gallant 1993). There are alternative approximations to the log-likelihood based on the Laplacian approximation (Tierney and Kadane 1986), importance sampling (Geweke 1989), and Gaussian quadrature (Davis and Rabinowitz 1984).

Sometimes in statistical problems the integrand may be a probability density function not easily expressible in a form suitable for computation, but at the same time it may be easy to sample from the distribution. In such cases Monte Carlo methods of Integration is a plausible candidate. Quasi Monte Carlo methods converge much faster than normal Monte Carlo is nothing but the use of quasi-random, rather than random, numbers in Monte Carlo methods. These methods are now widely used in scientific computation, especially in estimating integrals over multidimensional domains and in many different financial computations (Dimov 2008). The Sobol's sequence is one of the standard quasi-random sequences and is widely used in quasi Monte Carlo applications. During likelihood construction and fitting of nonlinear mixed effects model, we have to solve intractable integrations, which can be multidimensional. For solving such intractable integrations this study attempts to use Monte Carlo method of integration based on Sobol's sequence which is the newly proposed one along with some other existing methods of fitting such models. We compare them based on their computational and statistical properties, using simulation results. Section 2 contains a description of the Sobol's sequence based method for fitting the nonlinear mixed-effects model. Section 3 presents a comparison of the different approximations along with the proposed method based on simulated data and finally Section 4 gives some concluding remarks.

2. Fitting Nonlinear Mixed Effects Models using Sobol's Sequences

Nonlinear mixed effects model is specified in two stages, at the first stage the response corresponding to the j-th observation on the i-th subject is modeled as

$$y_{ij} = f\left(\boldsymbol{\beta}_i, \boldsymbol{x}_{ij}\right) + \varepsilon_{ij}, \quad i = 1,2,\ldots,m,\ j = 1,2,\ldots,n_i, \tag{1}$$

where f is a nonlinear function of a subject-specific parameter vector $\boldsymbol{\beta}_i$, $\boldsymbol{x}_{ij}$ is the predictor vector, ε_{ij} is a normally distributed noise term, i.e. $\varepsilon_{ij} \sim N(0, \sigma^2)$. In the second stage the subject-specific parameter vector is modeled as $\boldsymbol{\beta}_i = \boldsymbol{A}_i\boldsymbol{\beta} + \boldsymbol{B}_{\boldsymbol{i}}\boldsymbol{b}_{\boldsymbol{i}}$, where $\boldsymbol{\beta}$ is a p- dimensional vector of fixed population parameters and $\boldsymbol{b}_{\boldsymbol{i}}$ is a q-dimensional vector of random effects, and $\boldsymbol{A}$ and $\boldsymbol{B}$ are design matrices corresponding to the fixed and random effects respectively. The random effects term $\boldsymbol{b}_{\boldsymbol{i}}$ is assumed to follow a normal distribution, e.g. $\boldsymbol{b}_{\boldsymbol{i}} \sim N_q(\boldsymbol{0},\ \boldsymbol{D})$. It is further assumed that given the random effects $\boldsymbol{b}_{\boldsymbol{i}}$, response made on different subjects are independent and that the ε_{ij}s are independent of the $\boldsymbol{b}_{\boldsymbol{i}}$**.** For the nonlinear mixed effects model in equation (1) the marginal density of the response $\boldsymbol{y}_i$ is

$$p(\boldsymbol{y}_i; \boldsymbol{\beta}, \boldsymbol{D}, \sigma^2) = \int p_1(\boldsymbol{y}_i |\ \boldsymbol{b}_i, \boldsymbol{\beta}, \boldsymbol{D}, \sigma^2)\, p_2(\ \boldsymbol{b}_i; \boldsymbol{D}) d\boldsymbol{b}_i\,, \tag{2}$$

where p_1 is the conditional density of $\boldsymbol{y}_i$ given $\boldsymbol{b}_{\boldsymbol{i}}$ and p_2 is the marginal density of $\boldsymbol{b}_{\boldsymbol{i}}$**.** In general the integral needed to evaluate the marginal distribution of the response and hence to evaluate the likelihood function does not have a closed-form expression when the model function is nonlinear in random effects $\boldsymbol{b}_{\boldsymbol{i}}$.

Sobol's sequences, the quasi-random sequences are also called low-discrepancy sequences where discrepancy is a measure of deviation from uniformity of a sequence of points in U = ([0, 1]^s). The low-discrepancy sequences cover the unit cube as 'uniformly' as possible by reducing gaps and clustering of points (Paskov 1997). Here the outputs are constrained by a low-discrepancy requirement that has a net effect of points being generated in a highly correlated manner (i.e., the next point 'knows' where the previous points are). In computational problems such a sequence is extremely useful where numbers are computed on a grid, but it is not known in advance how fine the grid must be to obtain accurate results. Using a quasi random sequence allows stopping at any point where convergence is observed, whereas the usual approach of halving the interval between subsequent computations requires a huge number of computations between stopping points. For the purpose of likelihood construction in nonlinear mixed effects model we can solve the intractable integration through Monte Carlo integration based on Sobol's sequences and then proceed. We assume that,

$$y_{ij}\,|\boldsymbol{b}_i \;=f_{ij}\;(\boldsymbol{\beta},\boldsymbol{b}_i)\;+\;\epsilon_{ij}\,,i\;=\;1,\cdots,m,\qquad j\;=\;1\cdots,n,$$

where $y_{ij}\,|\boldsymbol{b}_i$, $j = 1\cdots,n'$ s are independent of one another with $y_{ij}\,|\boldsymbol{b}_i{\sim}N(f_{ij}\,(\boldsymbol{\beta},\boldsymbol{b}_i),\sigma^2)$. So, $p(y_i|\boldsymbol{b}_i;\boldsymbol{\beta},\sigma^2){\sim}\;N_n(\boldsymbol{f_i};\;\boldsymbol{I}\sigma^2)$, where $\boldsymbol{f_i}$=$(f_{i1},\cdots,f_{in})'$ and $\boldsymbol{I}$ is an $n\times n$ identity matrix. Also, $\boldsymbol{b_i}\sim N(\boldsymbol{0},\boldsymbol{D})$. Hence,

$$p(y_i|\boldsymbol{b}_i;\boldsymbol{\beta},\boldsymbol{D},\sigma^2)=\frac{1}{(2\pi\sigma^2)^{n/2}}\exp\left[-\frac{1}{2\sigma^2}\textstyle\sum_{j=1}^{n}(y_{ij}-f_{ij})^2\right],$$

and from equation (2), the marginal density of the response $\boldsymbol{y_i}$ is

$$p(y_i;\boldsymbol{\beta},\sigma^2,\boldsymbol{D})=\int\frac{1}{(2\pi\sigma^2)^{n/2}(2\pi|\boldsymbol{D}|)^{1/2}}\exp\left[-\frac{1}{2\sigma^2}\sum_{j=1}^{n}\left(y_{ij}-f_{ij}\right)^2-\frac{b_i'\boldsymbol{D}^{-1}b_i}{2}\right]db_i. \quad (3)$$

The log-likelihood function is defined as,

$$l_S=\sum_{i=1}^{m}\log p(y_i;\boldsymbol{\beta},\sigma^2,\boldsymbol{D}). \quad (4)$$

In equation (3) the integration is evaluated using the Monte Carlo integration rule with the help of Sobol's sequence. The values of $\boldsymbol{b}_i$ are generated from normal Sobol's sequence (with appropriate dimensions), with specific mean vector $\boldsymbol{0}$ and variance-covariance matrix $\boldsymbol{D}$. This is to be optimized for estimating all fixed effects and random effects parameters considered here as this is a function of these parameters.
The other three methods considered for the comparison are Taylor series approximation, Lindstrom and Bates approximation and Laplace approximation. The first one approximates the nonlinear (in random effects) model function into a linear function (in random effects) and then constructs likelihood function using this approximated function (Sheiner and Beal 1980). Lindstrom and Bates approximation is very close to Taylor series approximation and both of these works in two steps for constructing likelihood. Main difference of this with the previous one is that it takes first-order Taylor expansion of the nonlinear model function around the conditional modes of the random effects whereas the previous one takes around the expected value of the random effects implying some modification of the first one (Lindstrom and Bates 1990). Laplace approximation approximates the intractable integration and then constructs the likelihood (Tierney and Kadane 1986; Wolfinger 1993).

3. Simulation Study and Results

This section contains a comparison among some methods for fitting nonlinear mixed effects models. As an example a simulation study is performed for a nonlinear mixed effect Michaelis-Menten model for the comparison.

3.1. The Michaelis-Menten Model

Michaelis-Menten equation is often used in enzyme kinetics for examining the relationship between rate of reaction and substrate concentration. The model takes the following form for relating reaction rate (v) to substrate concentration (x) as

$$v = \frac{V_{max}\, x}{K_m + x}\,, \tag{5}$$

where V_{max} represents the maximum rate achieved by the system, at maximum (saturating) substrate concentrations and the Michaelis constant K_m is the substrate concentration at which the reaction rate is half of V_{max}.

Consider a hypothetical experiment which involves liver sample from one of the m selected subjects and a level of concentration of the substrate under investigation. In this case the response is the measured rate of reaction from each experiment, e.g. y_{ij} is the response corresponding to the experiment involving the j-th level of substrate concentration for the i-th subject. For each subject a number of experiments are performed with different substrate concentrations so that a Michaelis-Menten equation can be fitted. Assume that only the subject specific V_{max} parameter can describe the subject- specific variation in the Michaelis-Menten equations, so the following model is assumed for the response corresponding to the $(i,j)th$ treatment

$$y_{ij} = \frac{V_{max\,i}\, x_{ij}}{K_m + x_{ij}} + \epsilon_{ij} = \frac{e^{\theta_i} x_{ij}}{e^{\beta} + x_{ij}} + \epsilon_{ij}\,, \quad i = 1, \cdots, m, \qquad j = 1 \cdots, n, \tag{6}$$

where $\theta_i = \log(V_{max\,i}) = \theta + b_i$, $\beta = \log(K_m)$, $\epsilon_{ij} \sim N(0, \sigma^2)$ and b_i is random effects for subject i with $b_i \sim N\left(0, \sigma_b^2\right)$, and b_i and ϵ_{ij} are assumed to be independent. Both the fixed effects parameters (θ, β) and random effects parameters $(\sigma_b^2,\ \sigma^2)$ are of interest, and $\rho = \sigma_b^2/(\sigma^2 + \sigma_b^2)$ are known as the intra-class correlation coefficient, which indicates the strength of association between responses within a specific subject.

3.2. Simulation Study

The model (6) is used to generate responses for the hypothetical experiment described above where $\theta = \log(5)$ and $\beta = \log(2)$ are used as the true value of the fixed effects parameters, and ${\sigma_b}^2 = 0.50^2$, $\sigma^2 = 0.10^2$ are used as the true values of variance components which corresponds to the intra-class correlation coefficient $\rho = 0.96$. A design with 20 equally distant levels of substrate concentration ranging from 0.5 to 10 and the largest substrate concentration 50 i.e. $x = \{0.5, 1.0, 1.5, \ldots, 9.5, 10, 50\}$ is considered in the experiment for each of the subjects considered for the study. Different numbers of subjects are considered in the simulation study to examine the effect of it in estimation methods, which are A (with $m = 10$), B (with $m = 5$) and C (with $m = 20$). The same true values of the fixed and random effects are considered for the three cases *A*, *B*, and *C*. For comparing estimation procedure for nonlinear mixed effects models based on Sobol's sequence method with Lindstrom-Bates, Taylor series and Laplace approximation methods bias, estimated standard error (se), and simulated standard error (sd) based on 1000 simulations are reported in Table **1**. It is noted that the simulation programs are written in R.

Table **1** shows that estimators of all the fixed effects parameters have negligible bias when 10 subjects are considered (i.e. Case *A*), which is true for all the methods of estimation. Among them Laplace approximation based method has least bias for V_{max} and K_m than the others. The estimates are treated as unbiased in the sense that the bias is more than 10 times smaller than the simulated standard deviation of the corresponding estimates. The standard errors of the estimators of the fixed effect parameters are correctly computed by all the methods except the Sobol's sequence based method for which only the standard error of estimator $\widehat{K}_m$ is correctly computed. The estimates of random effects parameter $\hat{\sigma}_b$ are unbiased for Sobol's sequence based method and Laplace approximation and other two methods produced some biased estimates. Both the estimators corresponding to the fixed effects parameters are found to be unbiased for the

Case B where only 5 subjects are considered for the simulation. In this case, the estimator of the random effects parameter is found to be more biased compare to the Case A and only Laplace approximation based method produces unbiased estimate.

Table 1. Simulation results of parameter estimation in four different estimation methods for Cases A, B, and C

Cases	Parameters		Estimation methods			
			Taylor	Lindstorm-Bates	Laplace	Sobol's sequence
Case A	V_{max}	Bias	0.0006	0.0014	0.0002	0.0019
		se	0.1482	0.1482	0.0215	0.0468
		sd	0.1564	0.1564	0.0208	0.0823
	K_m	Bias	-0.0015	-0.0002	-0.0001	-0.0002
		se	0.0320	0.0319	0.0314	0.0320
		sd	0.0335	0.0335	0.0338	0.0335
	σ_b	Bias	0.0369	0.0370	0.0032	0.0056
		sd	0.1068	0.1068	0.1142	0.1355
Case B	V_{max}	Bias	-0.0085	-0.0077	0.0006	-0.0128
		se	0.1895	0.1894	0.0304	0.0631
		sd	0.2225	0.2225	0.0287	0.1478
	K_m	Bias	-0.0001	0.0011	-0.0001	0.0011
		se	0.0453	0.0449	0.0442	0.0452
		sd	0.0452	0.0452	0.0464	0.0452
	σ_b	Bias	0.0831	0.0832	0.0160	0.0319
		sd	0.1534	0.1534	0.1615	0.1843
Case C	V_{max}	Bias	-0.0019	-0.0012	0.0006	-0.0023
		se	0.1096	0.1096	0.0152	0.0645
		sd	0.1122	0.1122	0.0142	0.1011
	K_m	Bias	-0.0008	0.0004	0.0007	0.0005
		se	0.0226	0.0226	0.0222	0.0426
		sd	0.0225	0.0225	0.0229	0.0452
	σ_b	Bias	0.0150	0.0151	-0.0034	-0.0091
		sd	0.0758	0.0758	0.0806	0.1615

The standard errors are correctly computed for the estimator $\widehat{K}_m$ for all the methods. For the Case C where 20 subjects are considered in the simulation, both the fixed effect estimators $\widehat{V}_{max}$ and $\widehat{K}_m$ are found to be unbiased for all the methods and the estimator $\hat{\sigma}_b$ is found to be slightly biased for two methods whereas the Sobol's sequence based method and Laplace approximation based method produced unbiased estimate once again. The standard errors of $\widehat{K}_m$ are found to be correct for all the methods, but the standard error of $\widehat{V}_{max}$ is slightly incorrect for the Sobol's sequence based method. However, all the results improve for the Case C compared to the Case A and Case B.

4. Conclusions

Fitting of the nonlinear mixed effects model is an important issue because of the parameters involved with it and the way of their estimation procedure. Here the score functions are nonlinear in parameters so that no close form solution is available. Instead the iterative procedure is applied for parameter estimation in such model. A number of estimation procedures are available for fitting nonlinear mixed effects models, of which three methods, namely Lindstrom and Bates approximation, Taylor series approximation and Laplace approximation are considered in this paper. These methods consider different approaches for estimation; e.g. both the Lindstrom-Bates and Taylor series approximation based methods approximately linearize the nonlinear model function and hence construct likelihood function, where as the Laplace approximation based method approximates the likelihood of the nonlinear mixed effects model. A new method of estimation for nonlinear mixed effects model is proposed in this paper, which is exact in the sense that it does not require any approximation as it uses Sobol's sequence, a quasi-random sequence, to evaluate the integrations involved in the resulting likelihood function.

The simulation results show that there is no difference between these four methods in estimating the fixed effects parameters. The Laplace approximation based method is found to be the best and the proposed Sobol's sequence based method gives better result than Lindstrom-Bates approximation and Taylor series approximation methods in fitting random effects parameter, which depends on the number of subjects used in the simulation. As expected the estimates of random effects parameter tends to more accurate as the number of subjects increases. The study shows that the Sobol's sequence based method is quite compatible with the existing methods as well as approximation based methods are quite handy in fitting of nonlinear mixed effects models.

References

Davidian, M. and Gallant, A. (1992) "Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine," *Journal of Pharmacokinetics and Pharmacodynamics*, 20(5), 529-556.

Davidian, M. and Gallant, A. (1993) "The nonlinear mixed effects model with a smooth random effects density," *Biometrika*, 80(4), 475-488.

Davidian, M. and Giltinan, D. (1993) "Some simple methods for estimating intra-individual variability in nonlinear mixed effects models," *Biometrics*, 59-73.

Davis, P. and Rabinowitz, P. (1984) *Methods of numerical integration*, Academic Press.

Dimov, I. T. (2008) *Monte Carlo methods for applied scientists*, World Scientific Pub Co Inc.

Geweke, J. (1989) "Bayesian inference in econometric models using Monte Carlo integration," *Econometrica: Journal of the Econometric Society*, 57(6), 1317-1339.

Harville, D. (1977) "Maximum likelihood approaches to variance component estimation and to related problems," *Journal of the American Statistical Association*, 72, 320-338.

Lindstrom, M. and Bates, D. M. (1990) "Nonlinear mixed effects models for repeated measures data," *Biometrics*, 46, 673-687.

Paskov, S. (1997) "New methodologies for valuing derivatives," *Mathematics of Derivative Securities*, 545-582.

Sheiner, L. and Beal, S. (1980) "Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model: routine clinical pharmacokinetic data," *Journal of Pharmacokinetics and Pharmacodynamics*, 8(6), 553-571.

Tierney, L. and Kadane, J. (1986) "Accurate approximations for posterior moments and marginal densities," *Journal of the American Statistical Association*, 81, 82-86.

Vonesh, E. and Carter, R. (1992) "Mixed effects nonlinear regression for unbalanced repeated measures," *Biometrics*, 48, 1-17.

Wolfinger, R. (1993) "Laplace's approximation for nonlinear mixed models," *Biometrika*, 80(4), 791-795.