Bayesian analysis for heavy-tailed nonlinear mixed effects models

Cibele M. Russo[1], Danilo A. Silva[1]

[1] Departamento de Matemática Aplicada e Estatística, Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, SP, Brazil

E-mail for correspondence: cibele@icmc.usp.br

Abstract: Nonlinear models have many applications in different areas such as pharmacokinetics and pharmacodynamics, and random effects are often included to take into account the correlation between observations taken within the same subject. In this context, we propose a bayesian analysis for heavy-tailed nonlinear mixed effects models, which may produce more robust estimates for the parameters in the model.

Keywords: Nonlinear model, mixed model, bayesian inference

1 Introduction

Nonlinear mixed effects models are suitable for different applications such as longitudinal data or growth curves, specially in pharmacokinetics. In this work we discuss a bayesian approach for fitting nonlinear mixed effects models with scale mixture of normal distributions for the random effects and errors by using a stochastic formulation. The assumption of heavy-tailed distributions for the random effects and errors enables the model to produce more robust estimates against outlying or influential observations (see, for instance, Meza et al., 2011 and Russo et al., 2009). Results are applied to the theophylline data set, frequently used to exemplify the absorption and elimination of a substance in the body.

2 The model

It is usual to consider the normal distribution for the random effects and errors. One alternative to the normality would be, for instance, the scale mixture of normal (SMN) distributions. Let $\mathbf{Y}$ be an $m-$dimensional random vector following a distribution in its stochastic form

$$\mathbf{Y} = \boldsymbol{\mu} + \kappa(U)^{1/2}\mathbf{Z},$$

where $\boldsymbol{\mu}$ is the location vector, U is a positive random variable with cumulative distribution function (cdf) $\mathrm{H}(u, \boldsymbol{\nu})$ and probability density function (pdf) $h(u, \boldsymbol{\nu})$. Here, $\boldsymbol{\nu}$ is a scalar or vector parameter indexing the distribution of U, $\kappa(U)$ is the weight function, $\mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with $\mathbf{Z}$ and U independent. Given $U = u$, $\mathbf{Y}$ follows a multivariate normal distribution with mean $\boldsymbol{\mu}$ and variance-covariance $\kappa(u)\boldsymbol{\Sigma}$. In other words, the SMN distribution is a scale mixture of normal distributions, where the distribution of the scale factor U is the mixing distribution. The marginal pdf of $\mathbf{Y}$ may be written as

$$f(\mathbf{y}) \quad = \quad \int_0^\infty \phi_m(\mathbf{y}|\boldsymbol{\mu}, \kappa(u)\boldsymbol{\Sigma}) d\mathrm{H}(u, \boldsymbol{\nu}), \tag{1}$$

where $\phi_m(.|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stands for the probability density function of the m–variate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We will use the notation $\mathbf{Y} \sim \mathrm{SMN}_m(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \mathrm{H})$. Here, we consider specifically the cases where $U = 1$ (normal model), $U \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right)$ (Student-t model) and $U \sim Beta(\nu, 1)$ (slash model).

2.1 Nonlinear mixed-effects models with scale mixture of normal distributions

Suppose that $\mathbf{y} = (\mathbf{y}_1^\top, \ldots, \mathbf{y}_n^\top)^\top$ is a vector of observed continuous multivariate responses with $\mathbf{y}_i$ a $(n_i \times 1)$ vector containing the observations for the experimental unit i, $i = 1, \ldots, n$, such that

$$\begin{array}{l} \mathbf{y}_i = \mathbf{g}(\boldsymbol{\phi}_i, \mathbf{X}_i) + \boldsymbol{\epsilon}_i, \ i = 1, \ldots, n, \\ \boldsymbol{\phi}_i = \mathbf{A}_i\boldsymbol{\beta} + \mathbf{b}_i, \end{array} \tag{2}$$

in which $\mathbf{X}_i = (\mathbf{X}_{i1}, \ldots, \mathbf{X}_{in_i})^\top$ is a matrix of explanatory variables for the i-th unit, $\mathbf{b}_i$ is a $(q \times 1)$ vector of random effects, $\boldsymbol{\epsilon}_i$ is an $(n_i \times 1)$ vector of random errors values for $i = 1, \ldots, n$, $\boldsymbol{\beta}$ is a $(p \times 1)$ location vector and $\mathbf{A}_i$ is a full rank $(p \times p)$ matrix of known constants. We assume that

$$\begin{pmatrix} \boldsymbol{\epsilon}_i \\ \mathbf{b}_i \end{pmatrix} \overset{ind.}{\sim} \mathrm{SMN}_{n_i+q}\left(\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{D} \end{pmatrix}; \mathrm{H}\right), \tag{3}$$

where $\mathbf{D}$ and $\boldsymbol{\Sigma}_i$ are positive-definite dispersion matrices. We assume that $\mathbf{D} = \mathrm{diag}(\boldsymbol{\tau})$ is a unstructured matrix and denote its elements by $\boldsymbol{\tau} = \mathrm{diag}(\tau_1, \tau_2, \ldots, \tau_q)^\top$. The matrix $\boldsymbol{\Sigma}_i$ with dimension $(n_i \times n_i)$ is typically dependent upon i through its dimension, and it will be considered, for example, $\boldsymbol{\Sigma}_i = \sigma^2\mathbf{I}_{n_i}$ for $i = 1, \ldots, n$ and $\sigma > 0$ a scalar. Since $\mathbf{A}_i$ and $\mathbf{X}_i$ are known matrices, we will simplify the notation by writing $\mathbf{g}(\boldsymbol{\beta}, \mathbf{b}_i)$ to represent $\mathbf{g}(\boldsymbol{\phi}_i, \mathbf{X}_i) = \mathbf{g}(\mathbf{A}_i\boldsymbol{\beta} + \mathbf{b}_i, \mathbf{X}_i)$. Finally, $\mathrm{H} = \mathrm{H}(\cdot, \boldsymbol{\nu})$ is the cdf generator that determines the specific SMN model that was assumed. Under a bayesian framework, the unobserved quantities are considered as random variables. The prior distributions considered were $\boldsymbol{\beta}_i \sim N(0, v)$, $\tau_i \sim \text{Gamma}(a_i, b_i)$, $\sigma^2 \sim \text{Gamma}(c, d)$.

3 Numerical illustration

Pinheiro & Bates (2000) and Meza et al (2001) analysed the kinetic study of the agent theophylline. In that experiment, serum concentration (in mg/L) of theophylline was measured in eleven times (in h) after the administration of d dose (in mg/kg) in each of twelve patients.
This type of problem involves the absorption and elimination of a substance on the organism, and it is usual to model the mean theophylline concentration Y by using the nonlinear function of the time T and dose d as follows

$$\mathrm{E}(Y) = d\exp(lK_a + lK_e - lC_l)\frac{[\exp(-e^{lK_e}T) - \exp(-e^{lK_a}T)]}{e^{lK_a} - e^{lK_e}}.$$

This model incorporate the following interpretations for the parameters: lK_a represents the logarithm of the substance absorption rate, lK_e is the logarithm of the substance elimination rate and lC_l represents the logarithm of plasma clearance.
Thus, a nonlinear mixed effects model for the vector of observations $\mathbf{y}_i$ would be as in (2) with $\mathbf{g}(\boldsymbol{\phi}_i, \mathbf{X}_i) = \mathbf{g}(\boldsymbol{\phi}_i, \mathbf{T}_i) = (g(\boldsymbol{\phi}_i, T_{1i}), \ldots, g(\boldsymbol{\phi}_i, T_{m_i i}))^\top$ with

$$\boldsymbol{\phi}_i = (lK_e + b_{1i}, lK_a + b_{2i}, lCl + b_{3i})^\top$$

and

$$g(\boldsymbol{\phi}_i, T_{m_i i}) = d\exp(\boldsymbol{\phi}_{2i} + \boldsymbol{\phi}_{1i} - \boldsymbol{\phi}_{3i})\frac{[\exp(-e^{\boldsymbol{\phi}_{1i}}T) - \exp(-e^{\boldsymbol{\phi}_{2i}}T)]}{e^{\boldsymbol{\phi}_{2i}} - e^{\boldsymbol{\phi}_{1i}}}.$$

The Monte Carlo estimates using OpenBUGS were obtained generating chains of size 50000 spaced by 50. The posterior means and the corresponding standard deviations of the parameters are presented in Table 1 and fitted profiles show a suitable fit.

	Normal		**Student-t$_4$**		**Slash$_4$**	
	mean	(sd)	mean	(sd)	mean	(sd)
lK_e	-2.465	(0.102)	-2.452	(0.099)	-2.450	(0.099)
lK_a	0.462	(0.142)	0.460	(0.146)	0.459	(0.147)
lC_l	-3.232	(0.095)	-3.221	(0.097)	-3.220	(0.094)
$(\sigma^2)^{-1}$	0.636	(0.104)	0.552	(0.145)	0.123	(0.035)
$(\tau_1)^{-1}$	5.580	(1.691)	5.053	(1.492)	5.052	(1.505)
$(\tau_2)^{-1}$	13.060	(3.458)	13.150	(3.451)	13.210	(3.451)
$(\tau_3)^{-1}$	11.830	(3.087)	11.740	(3.082)	11.760	(3.085)

4 Discussion and remarks

Nonlinear mixed effects models plays an important role in nonlinear problems with correlated data. Considering a stochastic formulation in a bayesian

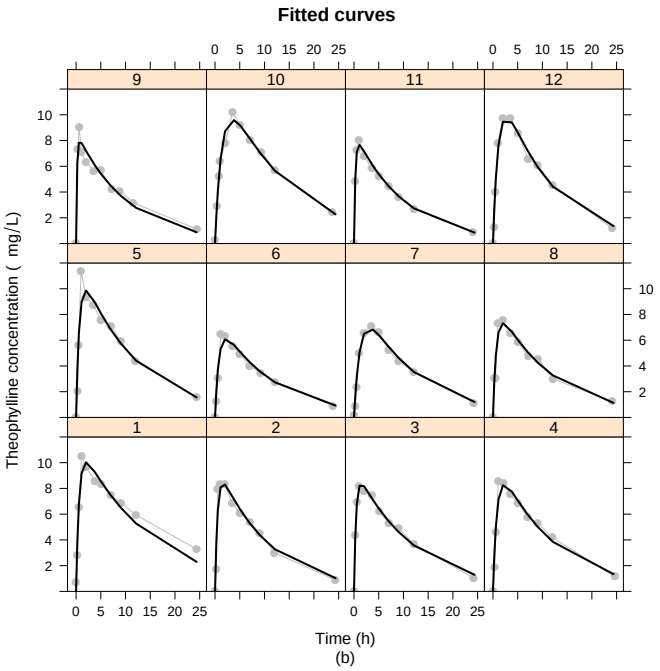

FIGURE 1. Fitted curves for theophylline problem under slash$_4$ model.

approach, we propose the use of heavy-tailed distributions in a bayesian context to provide alternatives to the gaussian model. As a result, the heavy-tailed models provide more robust estimates for the parameters of the model.

Acknowledgments: The authors thank to Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP, Brazil, for supporting this research.

References

Meza, C., Osorio, F. and De la Cruz, R. (2011) Estimation in nonlinear mixed-effects models using heavy-tailed distributions. Statistics and Computing 51, 4354–4368, doi 10.1007/s11222-010-9212-1.

Pinheiro, J. C. and Bates, D. M. (2000) Mixed-Effects Models in S and S-Plus, Springer, New York.

Russo, C. M., and Paula, G. A. and Aoki, R. (2009). Influence diagnostics in nonlinear mixed-effects elliptical models. *Computational Statistics & Data Analysis* 53, 4143–4156.