Analysis of Laplace Approximation for Pharmaceutical Nonlinear Mixed Effects Models

Mohamed Tarek[1,2]; Christopher Rackauckas[1,3,4]; Patrick K. Mognesen[1,5], Vijay Ivaturi[1,4], and Andreas Noack[1,5]
[1]Pumas AI Inc., [2]University of New South Wales at Canberra, [3]Massachusetts Institute of Technology, [4]University of Maryland, [5]Julia Computing Inc.

Abstract

Laplace approximation and its various approximations are some of the most commonly used techniques in pharmaceutical nonlinear mixed effects modelling to marginalise the the random effects. Classically, only Gaussian random effects have been used with transformations applied after the fact. However recently Pumas.jl introduced the ability to use non-Gaussian random effects in a nonlinear mixed effects model. In this work, the Laplace approximation framework is mathematically analysed showing that the approximation converges to the true integral as the number of observations per subject goes to infinity given some sufficient conditions. Those sufficient conditions were shown to be satisfiable by models with generic classes of non-Gaussian random effects. Additionally, a class of pharmacokinetic models was shown to violate the sufficient conditions for convergence when the infinitely many observations are collected by prolonging the study period rather than by collecting more observations per unit time in a finite time study. Numerous examples of pharmacokinetic dynamics models belonging to this class were then presented.

Introduction

Nonlinear mixed effects (NLME) modelling is commonly used in pharmaceutical modelling and simulation tools such as Pumas.jl [Rackauckas et. al., 2020]. Nonlinear mixed effects models are models with deterministic parameters θ , known as fixed effects, and random parameters η, known as random effects. NLME models used in pharmacometrics are typically discriminative models predicting the probability distribution of a response y given some covariates x and the fixed effects θ marginalising out the random effects η:

$$p(y|z,\theta) = \int_{D_\eta} p(y|z,\theta,\eta)p(\eta,\theta)d\eta$$

where D_η is the support of the random variables η. Further, these models are typically two-level hierarchical models where the fixed effects θ are shared among a population of subjects but the random effects, covariates and response are individualised. Let z_i, y_i and η_i be the i^{th} subject's covariates, response and random effects respectively and let $z = [z_1, z_2, \dots, z_N]$, $y = [y_1, y_2, \dots, y_N]$ and $\eta = [\eta_1, \eta_2, \dots, \eta_N]$ where N is the number of subjects. The probability of the response y can therefore be written as:

$$p(y|z,\theta) = \prod_{i=1}^{N} p(y_i|z_i,\theta) = \prod_{i=1}^{N} \int_{D_{\eta_i}} p(y_i|z_i,\theta,\eta_i)p(\eta_i|\theta)d\eta_i$$

The probability $p(y_i\,|z_i,\theta,\eta_i)$ describes the response probability distribution typically centered around the nonlinear solution of an ordinary differential equation (ODE). The ODE solution can be a function of z_i, η_i and θ. A class of such ODE-based models in pharmacometrics are pharmacokinetic and pharmacodynamic (PKPD) models. In PKPD models, y_i is usually a time-series response of longitudinal data such as the drug concentration in a subject's blood sample at different points in time. Let $y_{i,j}$ be response at the j^{th} point in time for the i^{th} subject and let $y_i = [y_{i,1}, y_{i,2}, \dots, y_{i,M_i}]$ where M_i is the number of observations for subject i. The response probability can therefore be written as:

$$p(y_i|z_i,\theta,\eta_i) = \prod_{j=1}^{M_i} p(y_{i,j}|y_{i,1:j-1},z_i,\theta,\eta_i)$$

where $y_{i,1:j-1}$ are the 1^{st} to $(j-1)^{st}$ observations associated with subject i. The nonlinearity in an NLME model makes the above integral analytically intractable in general thus forcing us to rely on approximate integration techniques to evaluate $p(y_i\,|z_i,\theta)$. Many of the common algorithms used in pharmacometrics and Pumas.jl are different variants and simplifications of the Laplace approximation method.

Laplace Approximation

Laplace approximation [Laplace, 1996] is a well known approximate integration technique that can be used under certain conditions. Let the integrand be $F(\xi)$, the integration variable be ξ, and the integration domain be $\mathbb{R}^K$, where K is the number of variables in ξ. Let F be continuous and twice differentiable and let $\log F(\xi)$ have a unique global maximum. The second-order Taylor series expansion of $\log F(\xi)$ around ξ_0 is:

$$logF(\xi) \approx \log F(\xi_0) + G^T(\xi-\xi_0) + \frac{1}{2}(\xi-\xi_0)^T H(\xi-\xi_0)$$

where G is the gradient vector whose i^{th} component is $G_i = \frac{d(\log F)}{d\xi_i}(\xi_0)$ and H is the Hessian matrix whose $(i,j)^{th}$ component is $H_{i,j} = \frac{d^2(\log F)}{d\xi_i\, d\xi_j}(\xi_0)$. When ξ_0 is taken to be the global maximiser of $\log F(\xi)$, $G = 0$ and H is negative definite. The integral $\int F(\xi)\, d\xi$ can therefore be approximated by:

$$\int F(\xi)d\xi \approx F(\xi_0)\sqrt{(2\pi)^K/|-H|}$$

In the remaining of this poster, only models with a single random effect will be considered. Multiple random effects and multivariate random effects will be looked at in a future work.

Sufficient Conditions for Convergence

Assume there is a single integration variable ξ, i.e. $K = 1$, and let $F(\xi;\, M) = f(\xi)\, e^{\sum_{i=1}^{M}\phi_i(\xi)}$. There are a number of proofs for the Laplace method with different conditions on the integral, e.g. see Bleistein and Handelsman [1975] and Bender and Orszag [1978] where the Laplace approximation of integrals of the form:

$$\int e^{-\lambda\Phi(\xi)} f(\xi) d\xi$$

were shown to converge to the true integral as λ goes to ∞, subject to a number of assumptions. In this work, a different proof strategy was used which:

1. Exploits the fact that $f(\xi)$ is positive in NLME models thus significantly simplifying the mathematics, and
2. Uses NLME model-specific notation and assumptions.

The following lemma describes some sufficient conditions for the convergence of the Laplace approximation of a specific form of integrals commonly used in NLME. The full proof is not shown here due to space limitation but can be found in the draft manuscript (available upon request).

Lemma 1: If the following conditions are satisfied:

1. $f(\xi) > 0$
2. $\phi(\xi;M) = \frac{1}{M}logF(\xi;M) = \frac{1}{M}logf(\xi) + \frac{1}{M}\sum_{i=1}^{M}\phi_i(\xi)$ is continuous and twice differentiable
3. For every $M > M_0$, there exists a unique finite global maximiser $a < \xi_0(M) < b$ such that:
   a) $\phi'(\xi_0(M);M) = 0$
   b) $\phi''(\xi_0(M);M) < 0$
4. $\lim_{M\to\infty} \phi''(\xi_0(M);M) < 0$
5. $\int_a^b e^{\phi(\xi;M)} < \infty$ for all $M > M_0$

then:

$$\lim_{M\to\infty} \frac{\int_a^b F(\xi;M)d\xi}{F(\xi_0(M);M)\sqrt{\frac{2\pi}{-\frac{d^2}{d\xi^2}logF(\xi_0(M);M)}}} = 1$$

Sufficient Conditions for Convergence

The following definitions formulate the NLME marginal in the form compatible with Lemma 1 for all random effects.

$$\phi_j(\xi) = \log p(y_j|y_{1:j-1},z,\theta,\eta(\xi))$$
$$f(\xi) = p(\eta(\xi)|\theta)$$

If the integration bounds are the bounds of the support of the random effect, then $\eta(\xi) = \xi$ can be used. Alternatively, a domain transformation can be used while still guaranteeing the convergence of the Laplace method under mild conditions. Refer to the paper for more details.

Violating the Sufficient Conditions

When the sufficient conditions are violated by a function, the convergence guarantees are lost. Recall that convergence is only guaranteed in the limit as $M \to \infty$ which mean collecting more observations/samples from a subject. This can be done by:

1. Collecting more data by prolonging the study and taking the end time T of the study to ∞, or
2. Collecting more data in a fixed study time window $[0, T]$ by taking samples more frequently.

Let the ODE describing the dynamics of a pharmacokinetic model be:

$$u' = f(u;\eta)$$
$$u(t_0) = u_0$$

If the Jacobian $\frac{df(u;\eta)}{du}$ is negative definite and M is taken to ∞ by taking T to ∞, then the sufficient conditions of the Laplace method are violated. A proof of this can be found in the draft manuscript (available upon request). Some common examples satisfying this condition include:

$$[Central'] = \left[-\frac{CL}{V}\right][Central] \qquad \begin{bmatrix} Depot' \\ Central' \end{bmatrix} = \begin{bmatrix} -Ka & 0 \\ Ka & -\frac{CL}{Vc} \end{bmatrix} \begin{bmatrix} Depot \\ Central \end{bmatrix}$$

$$\begin{bmatrix} Central' \\ Peripheral' \end{bmatrix} = \begin{bmatrix} -\frac{CL+Q}{Vc} & \frac{Q}{Vp} \\ \frac{Q}{Vc} & -\frac{Q}{Vp} \end{bmatrix} \begin{bmatrix} Central \\ Peripheral \end{bmatrix} \qquad \begin{bmatrix} Central' \\ CPeripheral' \\ Metabolite' \end{bmatrix} = \begin{bmatrix} -\frac{CL+Q+CLfm}{Vc} & \frac{Q}{Vp} & 0 \\ \frac{Q}{Vc} & -\frac{Q}{Vp} & 0 \\ \frac{CLfm}{Vc} & 0 & -\frac{CLm+Qm}{Vm} \end{bmatrix} \begin{bmatrix} Central \\ CPeripheral \\ Metabolite \end{bmatrix}$$

$$\begin{bmatrix} Depot' \\ Central' \\ Peripheral' \end{bmatrix} = \begin{bmatrix} -Ka & 0 & 0 \\ Ka & -\frac{CL+Q}{Vc} & \frac{Q}{Vp} \\ 0 & \frac{Q}{Vc} & -\frac{Q}{Vp} \end{bmatrix} \begin{bmatrix} Depot \\ Central \\ Peripheral \end{bmatrix} \qquad \begin{bmatrix} Central' \\ CPeripheral' \\ Metabolite' \\ MPeripheral' \end{bmatrix} = \begin{bmatrix} -\frac{CL+Q+CLfm}{Vc} & \frac{Q}{Vp} & 0 & 0 \\ \frac{Q}{Vc} & -\frac{Q}{Vp} & 0 & 0 \\ \frac{CLfm}{Vc} & 0 & -\frac{CLm+Qm}{Vm} & \frac{Qm}{Vmp} \\ 0 & 0 & \frac{Qm}{Vmp} & -\frac{Qm}{Vmp} \end{bmatrix} \begin{bmatrix} Central \\ CPeripheral \\ Metabolite \\ MPeripheral \end{bmatrix}$$

Conclusion and Future Work

In this research, a mathematical justification for the use of Laplace approximation with non-Gaussian random effects in NLME models was presented. It was proven that the choice of the random effect and its support does not affect the convergence guarantee of the Laplace method under some mild assumptions. Additionally, a class of pharmacokinetic models was presented for which the Laplace method's convergence sufficient conditions are not satisfied when more observations are collected by prolonging the study rather than by increasing the number of observations per unit time. In the future, we hope to theoretically and experimentally explore the connection between the Laplace method's convergence and an NLME model's identifiability as well as analyse the uniformity of convergence for specific models.

References

Carl M. Bender and Steven A. Orszag. Advanced Mathematical Methods for Scientists and Engineers. McGraw-Hill, Inc., 1978.
Norman Bleistein and Richard A. Handelsman. Asymptotic Expansion of Integrals. Holt. Rinehart and Winston, 1975.
Pierre Simon Laplace. Memoir on the probability of the causes of events. Statistical Science, 1(3):364–378, 1986.ISSN 08834237.
Chris Rackauckas, Yingbo Ma, Andreas Noack, Vaibhav Dixit, Patrick Kofod Mogensen, Simon Byrne, Shubham Maddhashiya, Jos´e Bayo´an Santiago Calder´on, Joakim Nyberg, Jogarao V.S. Gobburu, and Vijay Ivaturi. Accelerated predictive healthcare analytics with pumas, a high performance pharmaceutical modeling and simulation platform. bioRxiv, 2020.doi: 10.1101/2020.11.28.402297.