Evaluation of the expected Fisher information matrix without linearization, in nonlinear mixed effect models for discrete and continuous outcomes

Marie-Karelle Riviere, Sebastian Ueckert and France Mentré
INSERM, IAME, UMR 1137, F-75018 Paris, France;
Univ Paris Diderot, Sorbonne Paris Cité, F-75018 Paris, France

Introduction

Nonlinear mixed effect models (NLMEM) are widely used for the analysis of longitudinal data. Longitudinal data include continuous measures but also discrete measures such as binary, ordinal or count data. To evaluate and compare NLME designs, the Fisher information matrix (FIM) can be used as its inverse is the lower bound of the variance covariance matrix of any unbiased parameters estimator according to the Cramer-Rao inequality. However, in NLMEM, the Fisher information matrix has no analytical form and its calculation, which require multiple integrations, can be challenging. Therefore, an expression based on first-order linearization (FO) around the expectation of the random effects was proposed by Mentre et al. In recent years, estimation algorithms for NLMEMs have transitioned from linearization-based approaches towards more exact higher-order methods. Optimal design, on the other hand, has mainly relied on FO to calculate the expected FIM. Although very efficient in general, FO precludes the application of optimal design with complex non-linear models and in studies with discrete endpoints. The objective of this work was to apply integration algorithms, which have proven to be efficient for estimation, to evaluate the asymptotically exact FIM in NLMEM for both discrete and continuous outcomes.

Methods

Nonlinear mixed effect models

▷ For continuous data: $y_i = f(g(\mu, b_i), \xi_i) + \varepsilon_i$

▷ For discrete data: $P(y_i|b_i) = \prod_{j=1}^{n_i} h(y_{ij}, g(\mu, b_i), \xi_i)$

with

y_i response for individual i
f, h structural model
ξ_i elementary design for subject i
g individual parameters vector, function of μ and b_i
μ vector of fixed effects
b_i vector of random effects for individual i, $b_i \sim \mathcal{N}(0, \Omega)$
Ω variance-covariance matrix
ε_i vector of residual errors, $\varepsilon_i \sim \mathcal{N}(0, \Sigma)$
Σ is a diagonal matrix

Fisher Information Matrix

$$M(\psi, \xi) = E\left(\frac{\partial \log(L(y,\psi))}{\partial \psi}\frac{\partial \log(L(y,\psi))}{\partial \psi}^T\right)$$

Markov Chains Monte Carlo - MCMC

FIM: After calculation, (k, l) term of the FIM $\Longleftrightarrow$

$$\int_y \int_{b_1} \frac{\partial\left(\log(p(y|b_1,\psi)p(b_1))\right)}{\partial \psi_k} \underbrace{\frac{p(y|b_1,\psi)p(b_1)}{\int p(y|b,\psi)p(b)db}}_{\substack{\text{conditional density}\\ \text{of } b \text{ given } y}} db_1 \cdot \int_{b_2} \frac{\partial\left(\log(p(y|b_2,\psi)p(b_2))\right)}{\partial \psi_l} \underbrace{\frac{p(y|b_2,\psi)p(b_2)}{\int p(y|b,\psi)p(b)db}}_{\substack{\text{conditional density}\\ \text{of } b \text{ given } y}} db_2 \cdot \underbrace{L(y,\psi)}_{\substack{\text{marginal}\\ \text{density of } y}} dy$$

Monte Carlo - MC Markov Chains Monte Carlo - MCMC

Partial derivatives calculation:

- By hand. For example, for continuous data:
$$-2.\frac{\partial\left(\log(p(y|b,\psi)p(b))\right)}{\partial \psi_k} = Tr\left(V_b^{-1}\frac{\partial V_b}{\partial \psi_k}\right) - 2(y-E_b)^T V_b^{-1}\frac{\partial E_b}{\partial \psi_k} - (y-E_b)^T V_b^{-1}\frac{\partial V_b}{\partial \psi_k}V_b^{-1}(y-E_b) + Tr\left(\Omega^{-1}\frac{\partial \Omega}{\partial \psi_k}\right) - b^T\Omega^{-1}\frac{\partial \Omega}{\partial \psi_k}\Omega^{-1}b$$
with $E_b = f(g(\mu, b), \xi)$ and $V_b = \Sigma$
- Using **STAN**: MCMC sampler, calculate directly the partial derivatives (for all types of outcome)

Adaptive Gaussian Quadrature - AGQ

FIM: $M(\psi, \xi) = E_y\left(\frac{\partial L(y,\psi)}{\partial \psi}\frac{\partial L(y,\psi)}{\partial \psi}^T L(y,\psi)^{-2}\right) = \int_y \frac{\partial L(y,\psi)}{\partial \psi}\frac{\partial L(y,\psi)}{\partial \psi}^T L(y,\psi)^{-2}.L(y,\psi)\ dy$

with

$$\frac{\partial L(y,\psi)}{\partial \psi} = \int_b \frac{\partial(p(y|b,\psi)p(b))}{\partial \psi}\ db = \int_\eta \frac{\partial p(y|\Omega^{\frac{1}{2}}\eta,\psi)}{\partial \psi}\phi(\eta)\ d\eta + \int_\eta -\frac{1}{2}\left[Tr\left(\Omega^{-1}\frac{\partial \Omega}{\partial \psi_k}\right) - \eta^T\Omega^{-\frac{1}{2}}\frac{\partial \Omega}{\partial \psi_k}\Omega^{-\frac{1}{2}}\eta\right] p(y|\Omega^{\frac{1}{2}}\eta,\psi)\phi(\eta)\ d\eta$$

$$L(y,\psi) = \int_b p(y|b,\psi)p(b)\ db = \int_\eta p(y|\Omega^{\frac{1}{2}}\eta,\psi)\phi(\eta)\ d\eta$$

Monte Carlo - MC Adaptive Gaussian Quadrature - AGQ

Partial derivatives calculation: Finite differencing

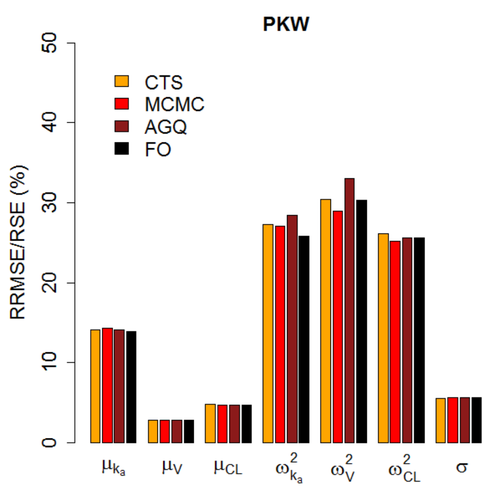
Simulation and Results

We compared 3 approaches to evaluate the FIM: MCMC-based approach, AGQ-based approach and Linearization (FO), with clinical trial simulations (CTS). We used several criteria: the RSE / RRMSE, and the normalized determinant of the FIM / normalized determinant of the inverse of the observed variance-covariance matrix, and also investiagted computation time. We used two continuous examples: a pharmacokinetics Warfarin model (PKW), and an Emax dose-response model (SC1); as well as two discrete examples: a logistic model for longitudinal binary outcome (LLB), and an exponential model with constant hazard for repeated time-to-event (RTTE).

PKW: One compartment model with first order absorption and elimination:

$$f(\phi = (k_a, V, CL), t) = \frac{70}{V}\frac{k_a}{k_a - \frac{CL}{V}}\left(e^{-\frac{CL}{V}t} - e^{-k_a t}\right)$$

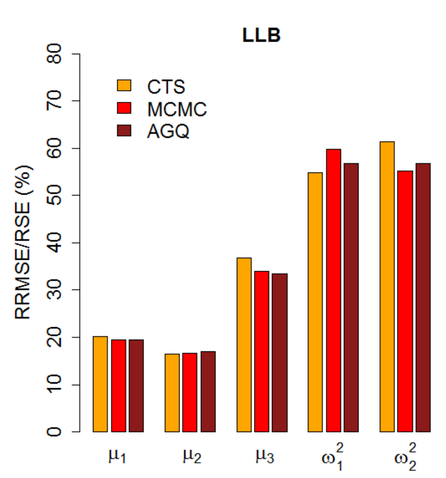
- Fixed effects: $(\mu_{k_a}, \mu_V, \mu_{CL}) = (1.00, 8.00, 0.15)$
- Exponential random effects with variances: $(\omega^2_{k_a}, \omega^2_V, \omega^2_{CL}) = (0.60, 0.02, 0.07)$
- Proportional residual error: $\sigma_{\text{slope}} = 0.1$
- 8 times: $t = (0.5, 1, 2, 6, 24, 36, 72, 120)$
- $N = 32$ patients



LLB: Probability of "one-response" at time t for group δ:

$$P(y = 1|b) = \frac{\exp(\beta_1 + \beta_2(1 - \mu_3\delta)t)}{1 + \exp(\beta_1 + \beta_2(1 - \mu_3\delta)t)}$$

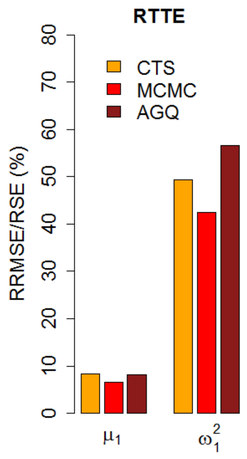
- Fixed effects: $(\mu_1, \mu_2, \mu_3) = (-1.0, 4.0, 0.4)$
- Additive random effects with variances: $(\omega^2_1, \omega^2_2) = (0.5, 4.0)$
- 2 groups: $\delta = 0$ and $\delta = 1$
- 13 time points equally spaced between 0 and 1 time units for each patient
- $N = 25$ patients per group



RRTE: Exponential distribution for repeated time-to-event with constant hazard:

$$P(y|b) = \lambda_1 \exp(-\lambda_1 t)$$

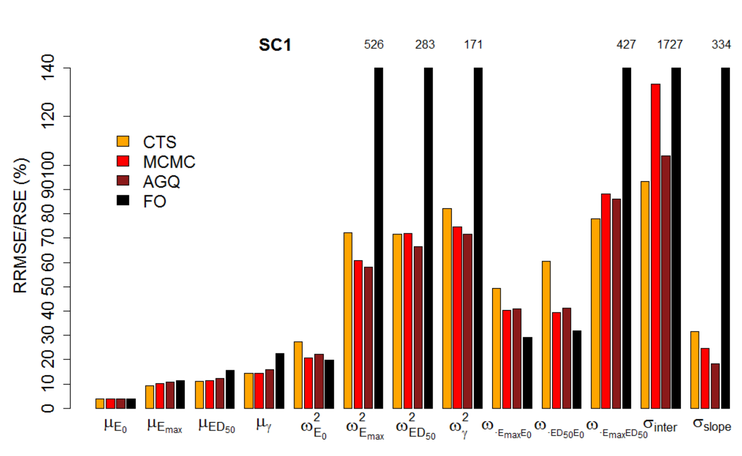
- Fixed effects: $\mu_1 = 1.0$
- Exponential random effects with variances: $\omega^2_1 = 0.1$
- 10 repeated measures per patient
- $N = 50$ patients



SC1: Sigmod E_{max} model:

$$f(\phi = (E_{max}, ED_{50}), d) = E_0 + \frac{E_{max}d^\gamma}{ED_{50}^\gamma + d^\gamma}$$

- Fixed effects: $(\mu_{E_0}, \mu_{E_{max}}, \mu_{ED_{50}}, \mu_\gamma) = (5, 30, 500, 3)$
- Exponential random effects with variance-covariance:
$$\Omega = \begin{pmatrix} 0.09 & 0.06 & 0.06 & 0 \\ 0.06 & 0.09 & 0.06 & 0 \\ 0.06 & 0.06 & 0.09 & 0 \\ 0 & 0 & 0 & 0.09 \end{pmatrix}$$
- Combined residual error: $(\sigma_{\text{inter}}, \sigma_{\text{slope}}) = (0.2, 0.2)$
- 4 doses: $d = (0, 100, 300, 1000)$
- $N = 100$ patients



	$\det(\text{FIM})^{\frac{1}{p}}$				Time			
	PKW	LLB	RTTE	SC1	PKW	LLB	RTTE	SC1
CTS	1116	7	526	177	>5h	>5h	>5h	>5h
MCMC	1122	8	370	146	≈ 6min	≈ 10min	≈ 2min	≈ 10min
AGQ	1088	8	464	149	≈ 2min	≈ 2min	≈ 10s	≈ 13min
FO	1133	-	-	104	<5s	-	-	<5s

Results: Both AGQ and MCMC-based approaches showed good performance on scenarios for continuous outcomes with RSEs close to the RRMSEs obtained by simulations. In general, RSE predicted by linearization gave close results for rich designs, but showed larger deviations for sparse designs and very non-linear models. On the contrary to FO, AGQ and MCMC can be applied to discrete data and showed good performance. Computation of the FIM with AGQ took only seconds for models with few random effects, but time computation increases exponentially with the number of random effects and models with more than 4 random effects became infeasible. The MCMC approach on the other hand was slower than AGQ for simple models, but can be applied to complex ones with similar time calculation.

Conclusion

Two complementing methods for calculating the exact FIM were proposed and evaluated. Both approaches showed similar performance in terms of RSE for model with continuous and discrete data. AGQ is a fast algorithm that can be used for simple models. MCMC suited even for large complex models where FO fails to correctly evaluate the FIM. We developped two R packages that will be soon available.