Maximum likelihood estimation in nonlinear mixed effect models: adaptive Gaussian quadrature by sparse grid sampling

PAGE 2009
St. Petersburg 22-26 June
Wan Hui O.Clausen¹, Birgitte B.Rønn¹, Ib M.Skovgaard²

¹:Biometrics, Genmab a/s. ²:Department of Basic Sciences and Environment, University of Copenhagen
Pinheiro and Bates (1995):

Comparison of estimation algorithms, NLME, Laplacian, AGQ and importance sampling

Conclusion: AGQ precise, but less efficient
Compartment model for Theophyline data

\[ c_t = \frac{D K_e K_a}{C_l (K_a - K_e)} (e^{-K_e t} - e^{-K_a t}) \]

- **D**: Dose
- **Ka**: Absorption rate
- **Ke**: Elimination rate
- **Cl**: Clearance
Theophylline serum concentrations
Boeckmann et al (1994)
Maximum likelihood estimation
• Non-linear mixed effect model:

\[ Y = h(\beta, b) + \varepsilon \]

where \( Y \) is the observation vector, \( h \) is the mean function, possible nonlinear in the fixed parameter vector, \( \beta \), and the random effect vector, \( b \).

The random effects are assumed to follow a multivariate normal distribution, \( b \sim N(0, \Omega) \), independent of the residual error, \( \varepsilon \), also assumed to follow a multivariate normal distribution, \( \varepsilon \sim N(0, \Sigma) \).
Maximum likelihood estimation in non-linear mixed effects models

- Likelihood function:

\[ L(Y, \beta, \Omega, \Sigma) = \int p(Y, \beta, b, \Omega, \Sigma) db \]

\[ = \int p(Y, \beta, \Sigma | b) p(b, \Omega) db \]

where \( p \) is the density function for the relevant normal distributions

- MLE: parameter values of \( \beta, \Omega, \Sigma \) that maximize the likelihood function

The final integral cannot be solved explicitly, hence approximations is needed, e.g. first order approximation, Laplace or AGQ.
Laplace approximation to $\int h(b)p(b)db$

Best Gaussian approximation with “mean” $\arg_b \max h(b)p(b)$ and “variance” $d^2(h(b)p(b)) db^{2-1}$
Laplace approximation to \( \int h(b) p(b) db \)

- The integral is approximated by the exact integral of the approximation.
- The approximation is exact when the random effects occur linearly in the mean function.
- The approximation works well if \( h(b)p(b) \) is approximately quadratic in \( b \).
Adaptive Gaussian Quadrature of $\int h(b) p(b) \, db$

Best approximation through the three abscissas

$h(b)p(b)$

$\hat{b} = \arg_b \max h(b)p(b)$

$\hat{b} - x$  $\hat{b}$  $\hat{b} + x$
Adaptive Gaussian Quadrature of $\int h(b) p(b) \, db$

- The integral is approximated by a weighted sum:
  \[ \int h(b) p(b) \, db = \sum_{i=1}^{n} h(x_i) p(x_i) w_i \]
- Where the $x_i$, $i=1, \ldots, n$ are the abscissas and $w_i$ the weight functions. Adaptive Gauss Hermite Quadrature have abscissas equal to roots of Hermite polynomials.
- The approximation is exact when $h(b)p(b)$ is a polynomial of degree $2n-1$ (or less)
Adaptive Gaussian Quadrature, with 2-dim $b$ and 5 abscissas, of $\int h(b) p(b) db$ (product rule)
AGQ for multivariate random effect

- The number of function evaluations grow exponentially with dimension $d$ of $b$:

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9</td>
<td>27</td>
<td>81</td>
<td>243</td>
<td>59049</td>
</tr>
<tr>
<td>7</td>
<td>49</td>
<td>343</td>
<td>2401</td>
<td>16807</td>
<td>0.2*10^9</td>
</tr>
<tr>
<td>13</td>
<td>169</td>
<td>2197</td>
<td>28561</td>
<td>371293</td>
<td>138*10^9</td>
</tr>
</tbody>
</table>

- AGQ becomes heavy computationally
- However, not all evaluation points are important for precision
Smolyak’s rule

- Old news: Smolyak (1963)
- We follow the tensor product construction by Gerstner and Griebel (2003)
- Example:

Smolyak's level 2 rule for d-dimensions:

\[
\int h(b)e^{-\frac{1}{2}b^2} \, db \approx \sum_{i=1}^{d} \left\{ h(-\sqrt{3}e_i) + h(\sqrt{3}e_i) \right\} + (1 - \frac{d}{3})h(0)
\]

- Weights equals 1/6 for the 2d axial points and 1-1/3 for center
- Exact for polynomials of degree 3 or less
- Number of function evaluations reduced significantly:

<table>
<thead>
<tr>
<th></th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9 5</td>
<td>27 7</td>
<td>81 9</td>
<td>243 11</td>
<td>59049 21</td>
</tr>
<tr>
<td>7</td>
<td>49 7</td>
<td>343 31</td>
<td>2401 49</td>
<td>16807 71</td>
<td>0.2*10^9 241</td>
</tr>
<tr>
<td>13</td>
<td>169 45</td>
<td>2197 105</td>
<td>28561 201</td>
<td>371293 341</td>
<td>138*10^9 1981</td>
</tr>
</tbody>
</table>
Examples
Theophylline serum concentrations

- Model:
  \[ c_t = \frac{D K_e K_a}{C l (K_a - K_e)} (e^{-K_e t} - e^{-K_a t}) \]

- Absorption rate, elimination rate and clearance are random

- Smolyak algorithm converged with non-singular covariance matrix of parameter estimates

- Smaller (slightly) residual variance compared to Laplace
Theophyline data - observed and predicted profiles
Indomethacin data - observed and predicted profiles

- Model:

\[ c_t = A_1 \exp(-r_{c1} \cdot \text{time}) + B_1 \exp(-r_{c2} \cdot \text{time}) \]

- A1, log(rc1) and B1 are random

- Smolyak algorithm converged with nonsingular covariance matrix
Indomethacin data

![Graph showing indomethacin concentration over time](image)
Simulations
Simulation: First-order open compartment model (Theophyline data)

- 500 simulations
- 12 subjects
- 10 concentrations at $t=0, 0.25, 0.5, 1, 2, 3.5, 7, 9, 12, 24h$
- Dose=4.5mg
- $\text{lKa} = 0.5$, $\text{lCl} = -3$, $\text{lKe} = -2.5$
- $\text{lKa}$ and $\text{lCl}$ are allowed to vary randomly, $b_i \sim \text{N}(0, \psi)$, where $\psi$ is diagonal, 0.36 and 0.04 respectively
### Simulation results (Theophyline data)

#### 500 simulations based on Theophylline data

<table>
<thead>
<tr>
<th></th>
<th>lKe</th>
<th>lKa</th>
<th>lCl</th>
<th>psi.lKa</th>
<th>psi.lCl</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True Value</strong></td>
<td>-2.5</td>
<td>0.5</td>
<td>-3</td>
<td>0.6</td>
<td>0.2</td>
<td>0.7</td>
</tr>
<tr>
<td><strong>NLME (starting values)</strong></td>
<td>-2.4940</td>
<td>0.4792</td>
<td>-3.0040</td>
<td>0.5799</td>
<td>0.1928</td>
<td>0.6993</td>
</tr>
<tr>
<td>sd</td>
<td>0.0403</td>
<td>0.1174</td>
<td>0.0405</td>
<td>0.0853</td>
<td>0.0300</td>
<td>0.0296</td>
</tr>
<tr>
<td><strong>Smolyak</strong></td>
<td>-2.5010</td>
<td>0.5057</td>
<td>-3.0010</td>
<td>0.5716</td>
<td>0.1935</td>
<td>0.6929</td>
</tr>
<tr>
<td>sd</td>
<td>0.0407</td>
<td>0.1210</td>
<td>0.0410</td>
<td>0.1087</td>
<td>0.0313</td>
<td>0.0361</td>
</tr>
<tr>
<td><strong>AGQ</strong></td>
<td>-2.5010</td>
<td>0.5004</td>
<td>-3.0010</td>
<td>0.5893</td>
<td>0.1952</td>
<td>0.6984</td>
</tr>
<tr>
<td>sd</td>
<td>0.0404</td>
<td>0.1196</td>
<td>0.0406</td>
<td>0.0889</td>
<td>0.0304</td>
<td>0.0296</td>
</tr>
</tbody>
</table>
Serum Concentration of Indomethacin
1000 simulations
6 subjects
11 concentrations at
t=0.25, 0.5, 0.75, 1, 1.25, 2, 3, 4, 5, 6, 8h
\( \alpha = 2.8 \), \( \text{lKe}_1 = 0.7 \), \( \beta = 0.4 \), \( \text{lKe}_2 = -1.5 \)
\( \alpha \) and \( \text{lKe}_1 \) are allowed to vary randomly, \( b_i \sim N(0, \psi) \), where \( \psi \) is diagonal, 0.36 and 0.04 respectively
## Simulation results (Indomethecin data)

### 1000 simulation based on Indomethecin data

<table>
<thead>
<tr>
<th></th>
<th>A1</th>
<th>lrc1</th>
<th>A2</th>
<th>lrc2</th>
<th>psiA1</th>
<th>psilrc1</th>
<th>sigma</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>True value</strong></td>
<td>2.8</td>
<td>0.7</td>
<td>0.4</td>
<td>-1.5</td>
<td>0.6</td>
<td>0.2</td>
<td>0.09</td>
</tr>
<tr>
<td><strong>Starting values (NLME)</strong></td>
<td>2.806</td>
<td>0.6644</td>
<td>0.3637</td>
<td>-1.6967</td>
<td>0.4975</td>
<td>0.1636</td>
<td>0.08772</td>
</tr>
<tr>
<td>sd</td>
<td>0.2701</td>
<td>0.1242</td>
<td>0.0953</td>
<td>0.3856</td>
<td>0.1859</td>
<td>0.0700</td>
<td>0.0089</td>
</tr>
<tr>
<td><strong>Smolyak</strong></td>
<td>2.814</td>
<td>0.7122</td>
<td>0.4169</td>
<td>-1.5432</td>
<td>0.5151</td>
<td>0.1920</td>
<td>0.08661</td>
</tr>
<tr>
<td>sd</td>
<td>0.2690</td>
<td>0.1335</td>
<td>0.11547</td>
<td>0.35370</td>
<td>0.1712</td>
<td>0.06829</td>
<td>0.008654</td>
</tr>
<tr>
<td><strong>AGQ</strong></td>
<td>2.798</td>
<td>0.7039</td>
<td>0.4111</td>
<td>-1.5572</td>
<td>0.4738</td>
<td>0.1817</td>
<td>0.08758</td>
</tr>
<tr>
<td>sd</td>
<td>0.2701</td>
<td>0.1311</td>
<td>0.11434</td>
<td>0.35904</td>
<td>0.1828</td>
<td>0.06894</td>
<td>0.008803</td>
</tr>
</tbody>
</table>
Conclusion
Conclusion

- AGQ precise method for MLE in nonlinear mixed effects models
- For multivariate parameters AGQ becomes difficult
- Smolyaks rule reduces the number of function evaluations significant
- The method works well on the examples aswell as in the simulation study
• References: