Approximation of the Fisher information and design in nonlinear mixed effects models

Tobias Mielke

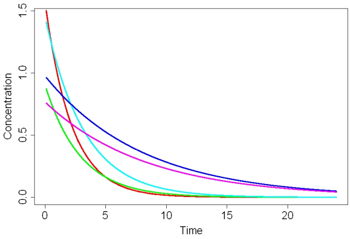
Optimum Design for Mixed Effects Non-Linear and Generalized Linear Models

PODE - 2011

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG

Outline

Mixed Effects Models



- Similar functions for different individuals
- Every individual has its own individual parameters
- Vectors of individual parameters are realizations of random vectors

$\rightarrow$ Mixed Effects Models

Mixed Effects Models

Two-stage-model:

- 1. stage (intra-individual variation):

$$Y_{ij} = \eta(\beta_i, x_{ij}) + \epsilon_{ij},\ j = 1, \dots, m_i,\ \epsilon_{ij} \sim N(0, \sigma^2)$$

- 2. stage (inter-individual variation):

$$\beta_i = \beta + b_i,\ i = 1, \dots, N,\ b_i \sim N_p(0, \sigma^2 D)$$

- b_i and ϵ_{ij} are assumed to be independent.
- Variance parameters σ^2 and D assumed to be known.
- Aim: Estimation of β.

Review

Individual observation vector:

$$Y_i = \eta(\beta_i, \xi_i) + \epsilon_i,$$

- m_i - number of observations,
- $\xi_i = (x_{i1}, ..., x_{im_i})$ - experimental settings,
- $\eta(\beta_i, \xi_i) := (\eta(\beta_i, x_{i1}), ..., \eta(\beta_i, x_{im_i}))^T$.

Design matrix:

$$F_\beta := (\frac{\partial \eta(\beta_i, \xi_i)}{\partial \beta_i^T}|_{\beta_i = \beta}).$$

Review

- Optimal designs for estimating $\beta \to cov(\hat{\beta})$?
- For Y_i with density $f_{Y_i}(y_i)$ and

$$l(\beta; y_i) := \log(f_{Y_i}(y_i)) :$$

Maximization of the Fisher information

$$\mathfrak{M}_\beta = E(\frac{\partial l(\beta; Y_i)}{\partial \beta} \frac{\partial l(\beta; Y_i)}{\partial \beta^T}).$$

- Often applied: Linearization.

Review

For η nonlinear in β_i:

- Linearization of the model (1):

$$\begin{aligned} Y_i &= \eta(\beta_i, \xi_i) + \epsilon_i \\ &\approx \eta(\beta_0, \xi_i) + F_{\beta_0}(\beta - \beta_0) + F_{\beta_0}(\beta_i - \beta) + \epsilon_i, \end{aligned}$$

  with a guess β_0 of β (or β_i).

- Distribution assumptions yield:

$$Y_i \sim N(\eta(\beta_0, \xi_i) + F_{\beta_0}(\beta - \beta_0), \sigma^2 V_{\beta_0})$$

  with $V_{\beta_0} := I_{m_i} + F_{\beta_0} D F_{\beta_0}^T$

  $\rightarrow$ Linear mixed effects model.

Review

For η nonlinear in β_i:

- Linearization of the model (2):

$$\begin{aligned} Y_i &= \eta(\beta_i, \xi_i) + \epsilon_i \\ &\approx \eta(\beta, \xi_i) + F_\beta(\beta_i - \beta) + \epsilon_i \end{aligned}$$

  with the true β.

- Distribution assumptions yield:

$$Y_i \sim N(\eta(\beta, \xi_i), \sigma^2 V_\beta)$$

  with $V_\beta := I_{m_i} + F_\beta D F_\beta^T$

  $\rightarrow$ Heteroscedastic nonlinear normal model.

Review

Calculation of the Fisher information:

Assumption: new model is true.

- Linear Mixed Information [Retout et. al (2001)]:

$$\Rightarrow \mathfrak{M}_{1,\beta}(\xi) = \frac{1}{\sigma^2} F_{\beta_0}^T V_{\beta_0}^{-1} F_{\beta_0}$$

- Nonlinear heteroscedastic Information

  [Retout et. al (2003)]:

$$\Rightarrow \mathfrak{M}_{2,\beta}(\xi) = \frac{1}{\sigma^2} F_{\beta}^T V_{\beta}^{-1} F_{\beta} + \frac{1}{2} S,$$

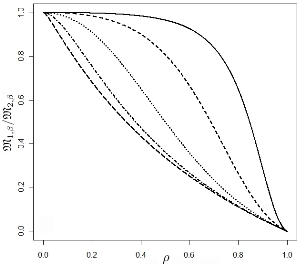
where $S \geq 0$, with

$$S_{j,k} = Tr\,[V_\beta^{-1} \frac{\partial V_\beta}{\partial \beta_j} V_\beta^{-1} \frac{\partial V_\beta}{\partial \beta_k}],\ j, k = 1, ..., p.$$

Review

Example: Observations as

$$
\begin{aligned}
Y_i &= \exp(\beta_i) + \epsilon_i, \quad \beta_i \sim N(\beta, d), \quad \epsilon_i \sim N(0,1) \text{ yield} \\
\mathfrak{M}_{1,\beta} &= \frac{\exp(2\beta)}{1 + d\exp(2\beta)} \text{ and} \\
\mathfrak{M}_{2,\beta} &= \frac{\exp(2\beta)}{1 + d\exp(2\beta)} + \frac{2d^2\exp(4\beta)}{(1 + d\exp(2\beta))^2}.
\end{aligned}
$$



Design

- Experiment settings ξ_i for individuals: $x_{ij} \in \mathcal{X}$

$$\xi_i = \begin{pmatrix} x_{i1} & \dots & x_{im_i} \end{pmatrix} \in \mathcal{X}^{m_i}.$$

- Alternative: Approximate individual designs - not here!
- Population design ζ:

$$\zeta = \begin{pmatrix} \xi_1 & \dots & \xi_k \\ \omega_1 & \dots & \omega_k \end{pmatrix}, \sum_{j=1}^{k} \omega_j = 1,$$

ω_j: weight of indivdual design ξ_j in the population.

- Population design $\hat{=}$ approximate design on $\mathcal{X}^m$

Design

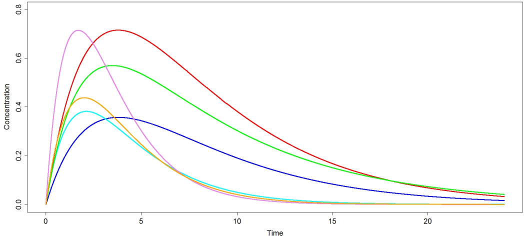
- Individual information: $\mathfrak{M}_\beta(\xi_i)$
- Population information: $\mathfrak{M}_{\beta;pop}(\zeta) = \sum_{i=1}^{k} \omega_i \mathfrak{M}_\beta(\xi_i)$
- Population design ζ^* D-optimal if and only if:

$$Tr\ \mathfrak{M}_{\beta;pop}(\zeta^*)^{-1}\mathfrak{M}_\beta(\xi) \leq p,\ \forall \xi \in \mathcal{X}^m$$

- Efficiency:

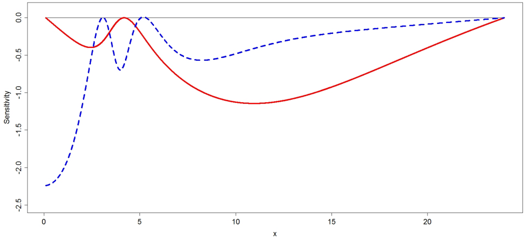
$$\delta(\zeta) := \left(\frac{\det(\mathfrak{M}_{\beta;pop}(\zeta))}{\det(\mathfrak{M}_{\beta;pop}(\zeta^*))} \right)^{\frac{1}{p}}.$$

Example [Schmelter 2007]:



- $Y_i = \eta(\beta_i, x_i)\exp(\epsilon_i)$, with $\epsilon_i \sim N(0, \sigma^2)$
- $\eta(\beta_i, x_i) := \frac{\beta_{1,i}}{\beta_{3,i}\beta_{1,i} - \beta_{2,i}}[\exp(-\frac{\beta_{2,i}}{\beta_{3,i}}x_i) - \exp(-\beta_{1,i}x_i)]$,
- $\beta_{i,k} = \beta_k \exp(b_{ik})$ with $b_{ik} \sim N(0, \sigma^2 d_k)$, $k = 1, 2, 3$,

Example: Information matters



$\mathfrak{M}_{1,\beta}$-Information:

$$\rightarrow \zeta_1 = \begin{pmatrix} (0.10) & (4.18) & (24.00) \\ 0.33 & 0.33 & 0.33 \end{pmatrix}, \delta(\zeta_2) = 0.66$$

$\mathfrak{M}_{2,\beta}$-Information:

$$\rightarrow \zeta_2 = \begin{pmatrix} (3.10) & (5.18) & (24.00) \\ 0.61 & 0.08 & 0.31 \end{pmatrix}, \delta(\zeta_1) = 0.55$$

Information Approximation

Remember:

$$Y_i = \eta(\beta_i, \xi_i) + \epsilon_i,\ \beta_i \sim N(\beta, \sigma^2 D),\ \epsilon_i \sim N(0, \sigma^2 I_{m_i})$$

log-Likelihood function:

$$\begin{aligned} l(\beta; y_i) &:= \log(f_{Y_i}(y_i)),\ \text{with} \\ f_{Y_i}(y_i) &:= \int_{\mathbb{R}^p} \frac{1}{c_1} \exp[-\frac{1}{2\sigma^2}\tilde{l}(\beta_i, \beta; y_i)] d\beta_i \text{ and} \\ \tilde{l}(\beta_i, \beta; y_i) &:= (y_i - \eta(\beta_i, \xi_i))^T (y_i - \eta(\beta_i, \xi_i)) \\ &\quad + (\beta_i - \beta)^T D^{-1} (\beta_i - \beta). \end{aligned}$$

We search:

$$\mathfrak{M}_\beta = E(\frac{\partial l(\beta; Y_i)}{\partial \beta} \frac{\partial l(\beta; Y_i)}{\partial \beta^T}).$$

Information Approximation

Alternative representation:

- With $E_{y_i}(\beta_i) := E(\beta_i | Y_i = y_i)$:

$$\frac{\partial l(\beta; y_i)}{\partial \beta} = \frac{1}{\sigma^2} D^{-1}(E_{y_i}(\beta_i) - \beta).$$

- For $Var_{y_i}(\beta_i) = Var(\beta_i | Y_i = y_i)$ follows:

$$\begin{aligned} \mathfrak{M}_\beta &= \frac{1}{\sigma^2} D^{-1} Var(E_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} \\ &= \frac{1}{\sigma^2} D^{-1} - \frac{1}{\sigma^2} D^{-1} E(Var_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} \end{aligned}$$

$\rightarrow$ Approximation of conditional moments.

Remark

More general inter-individual model:

Differentiable function $g : \mathbb{R}^{p_1} \to \mathbb{R}^p$:

$$\beta_i = g(\beta) + b_i, \text{ with } \beta \in \mathbb{R}^{p_1}, \; b_i \sim N_p(0, \sigma^2 D)$$

With the $(p \times p_1)$-Jacobi-matrix $G(\beta)$ follows:

$$\begin{aligned} \mathfrak{M}_\beta &= G(\beta)^T \frac{1}{\sigma^2} D^{-1} Var(E_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} G(\beta) \\ &= G(\beta)^T \left(\frac{1}{\sigma^2} D^{-1} - \frac{1}{\sigma^2} D^{-1} E(Var_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2}\right) G(\beta)^T \end{aligned}$$

Information Approximation

Approximation of conditional moments:

- Quadrature rules
- Simulations
  - $\rightarrow$ dependence on experimental settings?
  - $\rightarrow$ computational intensive.
- Laplace approximation
  - $\rightarrow$ accuracy?

Information Approximation

Laplace approximation:

- Problem in nonlinear mixed models:

$$E(E_{Y_i}(\beta_i)E_{Y_i}(\beta_i)^T)$$

- Approximation of $E_{y_i}(\beta_i)$ given y_i:

$$\begin{aligned} E_{y_i}(\beta_i) &= \frac{\int \beta_i \exp(-\frac{\tilde{l}(\beta_i,\beta;y_i)}{2\sigma^2})d\beta_i}{\int \exp(-\frac{\tilde{l}(\beta_i,\beta;y_i)}{2\sigma^2})d\beta_i}, \text{ with} \\ \tilde{l}(\beta_i,\beta;y_i) &:= (y_i-\eta(\beta_i,\xi_i))^T(y_i-\eta(\beta_i,\xi_i)) \\ &\quad +(\beta_i-\beta)^T D^{-1}(\beta_i-\beta). \end{aligned}$$

- Laplace approximation of both integrals. [Tierney and Kadane (1986)]

Information Approximation

Laplace approximation:

- Idea: Approximation of

$$f_{Y_i}(y_i) \quad = \quad \int \frac{1}{c_1} \exp(-\frac{\tilde{l}(\beta_i, \beta; y_i)}{2\sigma^2}) d\beta_i$$

  where $\tilde{l}(\beta_i, \beta; y_i)$ is minimial.

- Second order Taylor expansion around β_i^* with

$$\beta_i^* := \underset{\beta_i \in \mathbb{R}^p}{argmin} \{\tilde{l}(\beta_i, \beta; y_i)\}$$

$\rightarrow$ Linear part vanishes $\rightarrow$ normal density.

$$f_{Y_i}(y_i) \quad \approx \quad \frac{1}{c_2(y_i, \beta)} \exp(-\frac{\tilde{l}(\beta_i^*, \beta; y_i)}{2\sigma^2}).$$

- Dependence of β_i^* on y_i.

Information Approximation

- For approximating

$$f_{\beta_i|Y_i=y_i} = \frac{\phi_{Y_i|\beta_i}(y_i)\phi_{\beta_i}(\beta_i)}{f_{Y_i}(y_i)}$$

- Apply similar approximation to the numerator

$$\Rightarrow \beta_i|_{Y_i=y_i} \overset{app.}{\sim} N(\beta_i^*, \sigma^2 M_{\beta_i^*}^{-1}), \text{ with}$$

$$M_{\beta_i^*} := \frac{1}{2}\frac{\partial^2 \tilde{l}(\beta_i, \beta; y_i)}{\partial\beta_i\partial\beta_i^T}|_{\beta_i=\beta_i^*}.$$

- Problems:
  - Dependence of β_i^* on y_i
  - Dependence of $M_{\beta_i^*}^{-1}$ on y_i

Information Approximation

- Alternative: Just approximation of $\eta(\beta_i, \xi_i)$:

$$\eta(\beta_i, \xi_i) \approx \eta(\hat{\beta}_i, \xi_i) + F_{\hat{\beta}_i}(\beta_i - \hat{\beta}_i).$$

  in $\tilde{l}(\beta_i, \beta; y_i)$ with some $\hat{\beta}_i \in \mathbb{R}^p$.

- Same approximation to numerator yields

$$\begin{aligned} \Rightarrow \beta_i|_{Y_i=y_i} &\overset{app.}{\sim} N(\mu(y_i, \hat{\beta}_i, \beta), \sigma^2 \mathbf{M}_{\hat{\beta}_i}^{-1}), \text{ with} \\ \mathbf{M}_{\hat{\beta}_i} &:= F_{\hat{\beta}_i}^T F_{\hat{\beta}_i} + D^{-1} \\ \mu(y_i, \hat{\beta}_i, \beta) &:= \mathbf{M}_{\hat{\beta}_i}^{-1}(F_{\hat{\beta}_i}^T(y_i - \eta(\hat{\beta}_i, \xi_i)) + F_{\hat{\beta}_i}\hat{\beta}_i) + D^{-1}\beta). \end{aligned}$$

- Appropriate choice of $\hat{\beta}_i$?

Information Approximation

- For $\hat{\beta}_i = \beta_i^*$ minimizing $\tilde{l}(\beta_i, \beta; y_i)$:

$$\Rightarrow \beta_i|_{Y_i=y_i} \overset{app.}{\sim} N(\beta_i^*, \sigma^2 \mathbf{M}_{\beta_i^*}^{-1}).$$

Problem: Dependence on y_i.

- For $\hat{\beta}_i = \beta$:

$$\Rightarrow \beta_i|_{Y_i=y_i} \overset{app.}{\sim} N(\beta + \mathbf{M}_{\beta}^{-1} F_{\beta}^T (y_i - \eta(\beta, \xi_i)), \sigma^2 \mathbf{M}_{\beta}^{-1}).$$

- For $\hat{\beta}_i = \beta_i$:

$$\Rightarrow \beta_i|_{Y_i=y_i} \overset{app.}{\sim} N(\mu(y_i, \beta_i, \beta), \sigma^2 \mathbf{M}_{\beta_i}^{-1}).$$

Information Approximation

Fisher information approximations in $\hat{\beta}_i = \beta$:

- With $V_\beta := I_{m_i} + F_\beta D F_\beta^T$:
- Conditional variance:

$$\begin{aligned}\mathfrak{M}_\beta &= \frac{1}{\sigma^2} D^{-1} - \frac{1}{\sigma^2} D^{-1} E(Var_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} \\ &\approx \frac{1}{\sigma^2} F_\beta^T V_\beta^{-1} F_\beta =: \mathfrak{M}_{1,\beta}\end{aligned}$$

- Conditional expectation:

$$\begin{aligned}\mathfrak{M}_\beta &= \frac{1}{\sigma^2} D^{-1} Var(E_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} \\ &\approx \frac{1}{\sigma^4} F_\beta^T V_\beta^{-1} Var(Y_i) V_\beta^{-1} F_\beta =: \mathfrak{M}_{3,\beta}.\end{aligned}$$

Information Approximation

Fisher information approximation in $\hat{\beta}_i = \beta_i$:

- With $V_{\beta_i} := I_{m_i} + F_{\beta_i} D F_{\beta_i}^T$:
- Conditional variance:

$$\mathfrak{M}_{4,\beta_i} \;:=\; \frac{1}{\sigma^2} F_{\beta_i}^T V_{\beta_i}^{-1} F_{\beta_i}.$$

$\rightarrow$ consider here $\mathfrak{M}_{4,\beta} := \frac{1}{\sigma^2} E(F_{\beta_i}^T V_{\beta_i}^{-1} F_{\beta_i})$.

Some other possible approximations:

- Nonlinear heteroscedastic normal model:

$$\mathfrak{M}_{2,\beta} \;:=\; \frac{1}{\sigma^2} F_{\beta}^T V_{\beta}^{-1} F_{\beta} + \frac{1}{2} S.$$

- Quasi-Information:

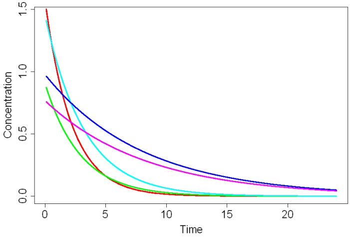
$$\mathfrak{M}_{5,\beta} := F_{\beta,QL}^T Var(Y_i)^{-1} F_{\beta,QL}$$

Information Approximation

Summary: With $V_{\beta_i} := I_{m_i} + F_{\beta_i} D F_{\beta_i}^T$:

$$
\begin{aligned}
\mathfrak{M}_{1,\beta} &:= \frac{1}{\sigma^2} F_\beta^T V_\beta^{-1} F_\beta \\
\mathfrak{M}_{2,\beta} &:= \frac{1}{\sigma^2} F_\beta^T V_\beta^{-1} F_\beta + \frac{1}{2} S \\
\mathfrak{M}_{3,\beta} &:= \frac{1}{\sigma^4} F_\beta^T V_\beta^{-1} Var(Y_i) V_\beta^{-1} F_\beta \\
\mathfrak{M}_{4,\beta} &:= \frac{1}{\sigma^2} E(F_{\beta_i}^T V_{\beta_i}^{-1} F_{\beta_i}) \\
\mathfrak{M}_{5,\beta} &:= F_{\beta,QL}^T Var(Y_i)^{-1} F_{\beta,QL}
\end{aligned}
$$

Example 3:

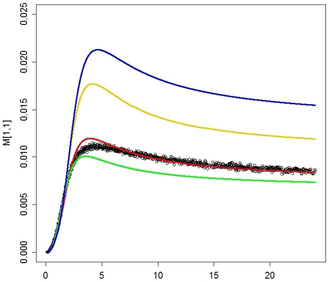


- Log-Concentration:

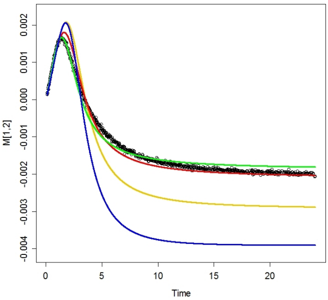
$$Y_i \quad = \quad -\beta_{2i} - x_i exp(\beta_{1i} - \beta_{2i}) + \epsilon_i, \;\; i = 1, ..., N$$

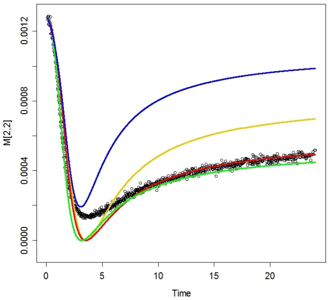
- - $\epsilon_i \sim N(0, \sigma^2)$, $\xi_i = (x_i) \in [t_{min}, t_{max}]$,
  - $\beta_i = (\beta_{1i}, \beta_{2i})^T \sim N(\beta, \sigma^2 D)$,
  - $\beta = (\beta_1, \beta_2)^T$, $D = diag(d_1, d_2)$.

Example 3



- $\mathfrak{M}_{1,\beta}$: Red
- $\mathfrak{M}_{2,\beta}$: Blue
- $\mathfrak{M}_{3,\beta}$: Yellow
- $\mathfrak{M}_{5,\beta}$: Green





Example 3

Information matrices $M_{1,\beta}(\beta)$ or $M_{5,\beta}(\beta)$:

For every pair of experimental settings $x_1, x_2 \in [t_{min}, t_{max}]$ with

$$\alpha(x_1, x_2) := \sigma^2 + cov(Y(x_1), Y(x_2)) = 0,$$

the population design

$$\zeta^1 = \begin{pmatrix} (x_1) & (x_2) \\ 0.5 & 0.5 \end{pmatrix}$$

is D-optimal in the case of one allowed observation per individual.

Example 3

How to see this:

- Equivalence theorem for D-optimality: ζ optimal

$$g_\zeta(\xi) := Tr\, \mathfrak{M}_{\beta;pop}(\zeta^*)^{-1}\mathfrak{M}_\beta(\xi) \leq p,\ \forall \xi \in \mathcal{X}^m$$

- With general nonsingular $M(\zeta) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ and $\xi = x$:

$$\tilde{g}_\zeta(x) := F_\beta(x) \begin{pmatrix} c & -b \\ -b & a \end{pmatrix} F_\beta(\xi)^T - 2(ac - b^2)V_\beta(x) \leq 0$$

- 2 support points, weight: $\omega = 0.5$.
- For

$$\zeta = \begin{pmatrix} (x_1) & (x_2) \\ 0.5 & 0.5 \end{pmatrix}$$

$$\Rightarrow \tilde{g}_\zeta(x) = \frac{\alpha(x_1, x_2)}{V(x_1)V(x_2)}(x - x_2)(x - x_1).$$

Example 3

$\mathfrak{M}_{1,\beta}$-Information:

$$\rightarrow \zeta_1 = \begin{pmatrix} (1.13) & (11.98) \\ 0.5 & 0.5 \end{pmatrix}, \delta_1(\zeta_1) = 1.0000$$

$\mathfrak{M}_{2,\beta}$-Information:

$$\rightarrow \zeta_2 = \begin{pmatrix} (0.96) & (\ 6.00) \\ 0.47 & 0.53 \end{pmatrix}, \delta_1(\zeta_2) = 0.9895$$

$\mathfrak{M}_{3,\beta}$-Information:

$$\rightarrow \zeta_3 = \begin{pmatrix} (1.70) & (24.00) \\ 0.5 & 0.5 \end{pmatrix}, \delta_1(\zeta_3) = 0.9794$$

$\mathfrak{M}_{5,\beta}$-Information:

$$\rightarrow \zeta_5 = \begin{pmatrix} (0.93) & (11.99) \\ 0.5 & 0.5 \end{pmatrix}, \delta_1(\zeta_5) = 0.9949$$

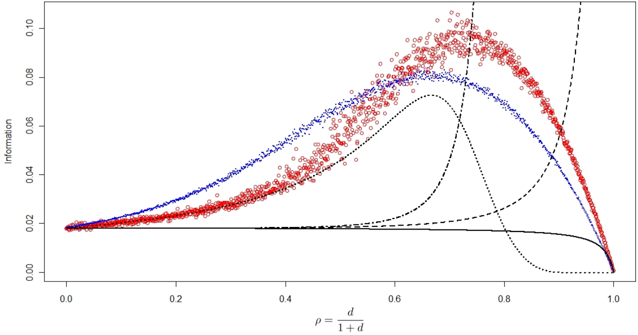
Example 1

Example: Observations as

$$
\begin{aligned}
Y_i &= \exp(\beta_i) + \epsilon_i, \;\; \beta_i \sim N(\beta, d), \;\; \epsilon_i \sim N(0,1) \text{ yield} \\
\mathfrak{M}_{1,\beta} &= \frac{\exp(2\beta)}{1 + d\exp(2\beta)}, \\
\mathfrak{M}_{2,\beta} &= \frac{\exp(2\beta)}{1 + d\exp(2\beta)} + \frac{2d^2\exp(4\beta)}{(1 + d\exp(2\beta))^2}, \\
\mathfrak{M}_{3,\beta} &= \frac{\exp(2\beta)(\exp(2\beta + d)(\exp(d) - 1) + 1)}{(1 + d\exp(2\beta))^2} \text{ and} \\
\mathfrak{M}_{5,\beta} &= \frac{\exp(2\beta + d)}{\exp(2\beta + d)(\exp(d) - 1) + 1}.
\end{aligned}
$$

Example 1

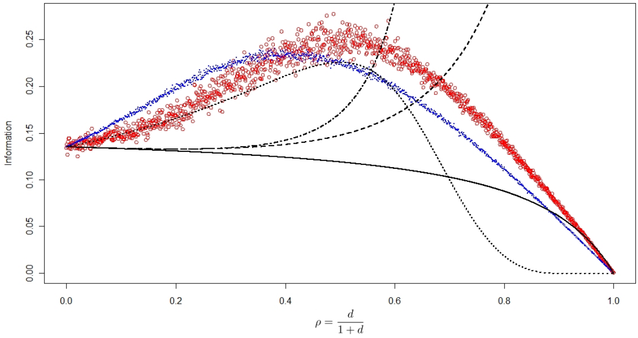
$$Y_i = \exp(\beta_i) + \epsilon_i;\ \beta_i \sim N(-2, d = \frac{\rho}{1-\rho});\ \epsilon_i \sim N(0, 1)$$



- Red: Simulated Fisher information; Blue: $\mathfrak{M}_{4,\beta}$
- $\mathfrak{M}_{1,\beta}$: solid, $\mathfrak{M}_{2,\beta}$: dashed, $\mathfrak{M}_{3,\beta}$: dot-dash, $\mathfrak{M}_{5,\beta}$: dotted

Example 1

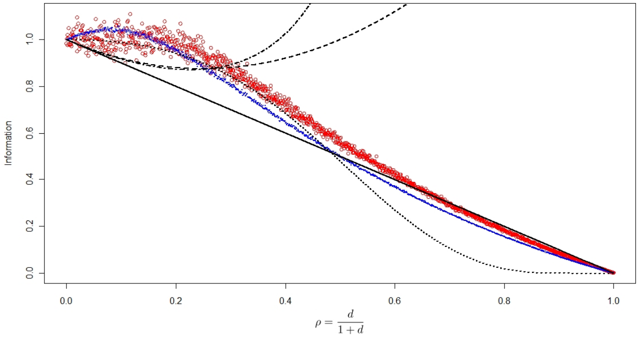
$$Y_i = \exp(\beta_i) + \epsilon_i;\ \beta_i \sim N(-1, d = \frac{\rho}{1-\rho});\ \epsilon_i \sim N(0,1)$$



- Red: Simulated Fisher information; Blue: $\mathfrak{M}_{4,\beta}$
- $\mathfrak{M}_{1,\beta}$: solid, $\mathfrak{M}_{2,\beta}$: dashed, $\mathfrak{M}_{3,\beta}$: dot-dash, $\mathfrak{M}_{5,\beta}$: dotted

Example 1

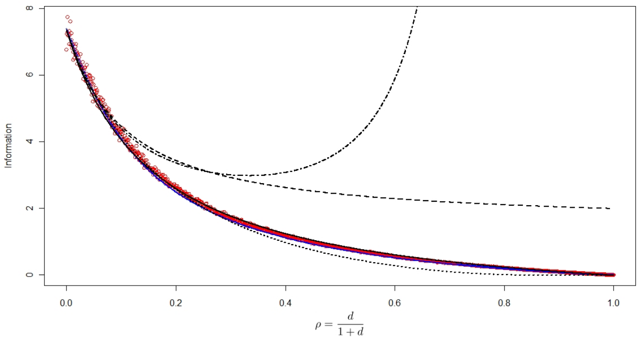
$$Y_i = \exp(\beta_i) + \epsilon_i;\ \beta_i \sim N(0, d = \frac{\rho}{1-\rho});\ \epsilon_i \sim N(0,1)$$



- Red: Simulated Fisher information; Blue: $\mathfrak{M}_{4,\beta}$
- $\mathfrak{M}_{1,\beta}$: solid, $\mathfrak{M}_{2,\beta}$: dashed, $\mathfrak{M}_{3,\beta}$: dot-dash, $\mathfrak{M}_{5,\beta}$: dotted

Example 1

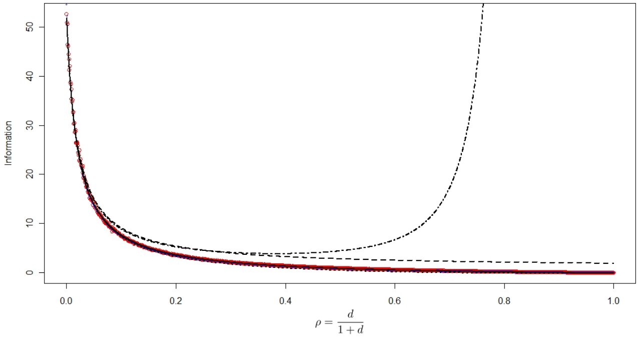
$$Y_i = \exp(\beta_i) + \epsilon_i;\ \beta_i \sim N(1, d = \frac{\rho}{1-\rho});\ \epsilon_i \sim N(0, 1)$$



- Red: Simulated Fisher information; Blue: $\mathfrak{M}_{4,\beta}$
- $\mathfrak{M}_{1,\beta}$: solid, $\mathfrak{M}_{2,\beta}$: dashed, $\mathfrak{M}_{3,\beta}$: dot-dash, $\mathfrak{M}_{5,\beta}$: dotted

Example 1

$$Y_i = \exp(\beta_i) + \epsilon_i;\ \beta_i \sim N(2, d = \frac{\rho}{1-\rho});\ \epsilon_i \sim N(0,1)$$



- Red: Simulated Fisher information; Blue: $\mathfrak{M}_{4,\beta}$
- $\mathfrak{M}_{1,\beta}$: solid, $\mathfrak{M}_{2,\beta}$: dashed, $\mathfrak{M}_{3,\beta}$: dot-dash, $\mathfrak{M}_{5,\beta}$: dotted

Example 1

Remember:

$$
\begin{aligned}
Var(\beta_i) &= E(Var_{Y_i}(\beta_i)) + Var(E_{Y_i}(\beta_i)) \\
\mathfrak{M}_\beta &= \frac{1}{\sigma^2} D^{-1} Var(E_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2} \\
&= \frac{1}{\sigma^2} D^{-1} - \frac{1}{\sigma^2} D^{-1} E(Var_{Y_i}(\beta_i)) D^{-1} \frac{1}{\sigma^2}
\end{aligned}
$$

Some explanation:

- $d \to \infty \quad \Rightarrow \mathfrak{M}_\beta \to 0.$
- $d \to 0 \Rightarrow$ Nonlinear regression

$$\Rightarrow \mathfrak{M}_\beta \to \frac{1}{\sigma^2} F_\beta^T F_\beta$$

Conclusions/Outlook

- Conclusions:
  - Motivation of model approximation
  - Different information $\rightarrow$ different design
  - Big influence of inter-individual variance on the accuracy of approximations
- Outlook:
  - More insight needed on:
    - appropriateness of the approximations
    - approximation in β_i or β_i^*
  - Inclusion of variance parameters.

Thank you for your attention!

Further informations needed ($M_{6,\beta}$,...)?