Optimal Designs for Individual Prediction in Population Based Models

Maryna Prus, Rainer Schwabe

22.3.2012

Contents

RCR model

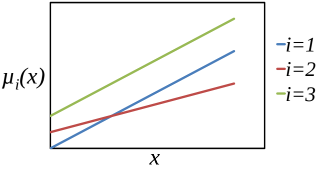
$$Y_{ij} = \mathbf{f}(x_{ij})^\top \boldsymbol{\beta}_i + \varepsilon_{ij}, \quad i = 1, .., n, \quad j = 1, .., m_i, \quad x_{ij} \in \mathcal{X}$$

- $\mathbf{f} = (f_1, .., f_p)^\top$
- $\boldsymbol{\beta}_i = (\beta_{i1}, .., \beta_{ip})^\top$, $\mathrm{E}(\boldsymbol{\beta}_i) = \boldsymbol{\beta}$ unknown, $\mathrm{Cov}(\boldsymbol{\beta}_i) = \sigma^2 \mathbf{D}$
- $\mathrm{E}(\varepsilon_{ij}) = 0$, $\mathrm{Var}(\varepsilon_{ij}) = \sigma^2$
- All $\boldsymbol{\beta}_i$ and all $\varepsilon_{i'j}$ uncorrelated

Gladitz & Pilz (1982): $\boldsymbol{\beta}$ known

Individual response:

$$\mu_i(x) = \mathbf{f}(x)^\top \boldsymbol{\beta}_i$$



Special model:

$$m_i = m \quad \& \quad x_{ij} = x_j$$

BLUE & BLUP in LMM

$$\mathbf{Y} = \mathbf{X}\beta + \mathbf{Z}\gamma + \varepsilon$$

- $\mathrm{E}(\gamma) = 0$, $\mathrm{Cov}(\gamma) = \mathbf{G}$, $\mathbf{G}$ regular
- $\mathrm{E}(\varepsilon) = 0$, $\mathrm{Cov}(\varepsilon) = \mathbf{R}$, $\mathbf{R}$ regular
- γ and ε uncorrelated
- β unknown

$$\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = \underbrace{\begin{pmatrix} \mathbf{X}^\top\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^\top\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^\top\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^\top\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{pmatrix}^{-1}}_{=:\mathbf{C}} \begin{pmatrix} \mathbf{X}^\top\mathbf{R}^{-1}\mathbf{Y} \\ \mathbf{Z}^\top\mathbf{R}^{-1}\mathbf{Y} \end{pmatrix}$$

$$\mathbf{C} = \mathrm{Cov}\begin{pmatrix} \hat{\beta} \\ \hat{\gamma} - \gamma \end{pmatrix}$$

Henderson (1975)

Model adaptation I

- $Y_{ij} = \mathbf{f}(x_j)^\top \boldsymbol{\beta} + \mathbf{f}(x_j)^\top \mathbf{b}_i + \varepsilon_{ij}$

$$\mathbf{b}_i := \boldsymbol{\beta}_i - \boldsymbol{\beta}$$

- $\mathbf{Y}_i = \mathbf{F}\boldsymbol{\beta} + \mathbf{F}\mathbf{b}_i + \boldsymbol{\varepsilon}_i$

$$\mathbf{Y}_i = \begin{pmatrix} Y_{i1} \\ \vdots \\ Y_{im} \end{pmatrix} \quad \boldsymbol{\varepsilon}_i = \begin{pmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{im} \end{pmatrix} \quad \mathbf{F} = \begin{pmatrix} \mathbf{f}(x_1)^\top \\ \vdots \\ \mathbf{f}(x_m)^\top \end{pmatrix}$$

- $\mathbf{Y} = (\mathbf{1}_n \otimes \mathbf{F})\boldsymbol{\beta} + (\mathbf{I}_n \otimes \mathbf{F})\mathbf{b} + \boldsymbol{\varepsilon}$

$$\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_n \end{pmatrix} \quad \boldsymbol{\varepsilon} = \begin{pmatrix} \boldsymbol{\varepsilon}_1 \\ \vdots \\ \boldsymbol{\varepsilon}_n \end{pmatrix} \quad \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \vdots \\ \mathbf{b}_n \end{pmatrix}$$

Problem: $\mathrm{Cov}(\mathbf{b}) = \sigma^2 \mathbf{I}_n \otimes \mathbf{D}$ may be singular

Model adaptation II

$\text{rank}(\mathbf{D}) = q,\ q \leq p$

$$\Rightarrow \quad \exists\ \mathbf{H} \quad \text{with} \quad \mathbf{D} = \mathbf{H}\mathbf{H}^\top \quad \& \quad \text{rank}(\mathbf{H}) = q$$

$\mathbf{c}_i := (\mathbf{H}^\top\mathbf{H})^{-1}\mathbf{H}^\top\mathbf{b}_i$

$$\Rightarrow \quad \mathbf{b}_i = \mathbf{H}\mathbf{c}_i \quad \& \quad \text{E}(\mathbf{c}_i) = 0,\ \text{Cov}(\mathbf{c}_i) = \sigma^2\mathbf{I}_q$$

- $\mathbf{Y}_i = \mathbf{F}\beta + \mathbf{F}\mathbf{H}\mathbf{c}_i + \varepsilon_i$
- $\mathbf{Y} = \underbrace{(\mathbf{1}_n \otimes \mathbf{F})}_{\mathbf{X}}\beta + \underbrace{(\mathbf{I}_n \otimes (\mathbf{F}\mathbf{H}))}_{\mathbf{Z}}\mathbf{c} + \varepsilon \qquad \mathbf{c} = \left(\mathbf{c}_1^\top, \ldots, \mathbf{c}_n^\top\right)^\top$

$$\Rightarrow \quad \mathbf{R} = \sigma^2\mathbf{I}_{nm} \qquad \mathbf{G} = \sigma^2\mathbf{I}_{nq}$$

Individual prediction

$$\begin{aligned}\hat{\boldsymbol{\beta}}_i &= \hat{\boldsymbol{\beta}} + \mathbf{H}\hat{\mathbf{c}}_i \\ &= \mathbf{D}((\mathbf{F}^\top\mathbf{F})^{-1}+\mathbf{D})^{-1}\hat{\boldsymbol{\beta}}_{i;\mathrm{ind}} + (\mathbf{F}^\top\mathbf{F})^{-1}((\mathbf{F}^\top\mathbf{F})^{-1}+\mathbf{D})^{-1}\hat{\boldsymbol{\beta}}\end{aligned}$$

$$\hat{\boldsymbol{\beta}} = (\mathbf{F}^\top\mathbf{F})^{-1}\mathbf{F}^\top\bar{\mathbf{Y}} \quad \& \quad \hat{\boldsymbol{\beta}}_{i;\mathrm{ind}} := (\mathbf{F}^\top\mathbf{F})^{-1}\mathbf{F}^\top\mathbf{Y}_i$$

$$\bar{\mathbf{Y}} = \tfrac{1}{n}\sum_{i=1}^n \mathbf{Y}_i$$

MSE Matrix for predictors $(\hat{\boldsymbol{\beta}}_1^\top, ..., \hat{\boldsymbol{\beta}}_n^\top)^\top$:

$$\sigma^2\left(\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top\right)\otimes\left(\mathbf{D}-\mathbf{D}((\mathbf{F}^\top\mathbf{F})^{-1}+\mathbf{D})^{-1}\mathbf{D}\right) + \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top\otimes(\mathbf{F}^\top\mathbf{F})^{-1}\right)$$

for regular $\mathbf{D}$ simplifies to

$$\sigma^2\left(\left(\mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^\top\right)\otimes(\mathbf{F}^\top\mathbf{F}+\mathbf{D}^{-1})^{-1} + \frac{1}{n}\left(\mathbf{1}_n\mathbf{1}_n^\top\right)\otimes(\mathbf{F}^\top\mathbf{F})^{-1}\right)$$

Individual design

Experimental settings $x_1, \ldots, x_m$ not necessarily all distinct

- Distinct settings: $x_1, \ldots, x_k$
- Numbers of replications: $m_1, \ldots, m_k$; $\sum_{j=1}^{k} m_j = m$

Individual design:

$$\xi = \begin{pmatrix} x_1, \ldots, x_k \\ w_1, \ldots, w_k \end{pmatrix}$$

$$w_j = \frac{m_j}{m} \quad \& \quad \sum_{j=1}^{k} w_j = 1$$

Notation:

$$\mathbf{M}(\xi) := \sum_{j=1}^{k} w_j \mathbf{f}(x_j)\mathbf{f}(x_j)^\top = \frac{1}{m}\mathbf{F}^\top\mathbf{F} \quad \& \quad \boldsymbol{\Delta} := m\mathbf{D}$$

IMSE-criterion

$$\mathrm{IMSE}_{\mathrm{pred}}(\xi) := \int \mathrm{E}_\xi \left(\sum_{i=1}^{n} \left(\hat{\mu}_i(x) - \mu_i(x)\right)^2 \right) \nu(\mathrm{d}x)$$

$$\hat{\mu}_i(x) = \mathbf{f}(x)^\top \hat{\boldsymbol{\beta}}_i$$

$$\mathrm{IMSE}_{\mathrm{pred}}(\xi) = \frac{\sigma^2}{m} \operatorname{tr}\left(\left(\mathbf{M}(\xi)^{-1} + (n-1)\left(\boldsymbol{\Delta} - \boldsymbol{\Delta}(\mathbf{M}(\xi)^{-1} + \boldsymbol{\Delta})^{-1}\boldsymbol{\Delta}\right)\right)\mathbf{V}\right)$$

$$\mathbf{V} := \int \mathbf{f}(x)\mathbf{f}(x)^\top \nu(\mathrm{d}x)$$

For regular **D** this simplifies to

$$\mathrm{IMSE}_{\mathrm{pred}}(\xi) = \frac{\sigma^2}{m} \operatorname{tr}\left(\left(\mathbf{M}(\xi)^{-1} + (n-1)(\mathbf{M}(\xi) + \boldsymbol{\Delta}^{-1})^{-1}\right)\mathbf{V}\right)$$

weighted sum of IMSE-criterion in fixed effects model and Bayesian IMSE

Optimality condition for approximate design

ξ^* IMSE-optimal if and only if

$$(n-1)\,\mathbf{f}(x)^\top \mathbf{M}(\xi^*)^{-1}\left(\mathbf{M}(\xi^*)^{-1}+\mathbf{\Delta}\right)^{-1}\mathbf{\Delta}\mathbf{V}\mathbf{\Delta}\left(\mathbf{M}(\xi^*)^{-1}+\mathbf{\Delta}\right)^{-1}\mathbf{M}(\xi^*)^{-1}\mathbf{f}(x)$$
$$+\,\mathbf{f}(x)^\top \mathbf{M}(\xi^*)^{-1}\mathbf{V}\mathbf{M}(\xi^*)^{-1}\mathbf{f}(x)$$
$$\leq \operatorname{tr}\left(\left((n-1)\mathbf{\Delta}\left(\mathbf{M}(\xi^*)^{-1}+\mathbf{\Delta}\right)^{-1}\mathbf{M}(\xi^*)^{-1}\left(\mathbf{M}(\xi^*)^{-1}+\mathbf{\Delta}\right)^{-1}\mathbf{\Delta}+\mathbf{M}(\xi^*)^{-1}\right)\mathbf{V}\right)$$

for all $x \in \mathcal{X}$

For regular $\mathbf{D}$ this simplifies to

$$\mathbf{f}(x)^\top\left((n-1)(\mathbf{M}(\xi^*)+\mathbf{\Delta}^{-1})^{-1}\mathbf{V}(\mathbf{M}(\xi^*)+\mathbf{\Delta}^{-1})^{-1}+\mathbf{M}(\xi^*)^{-1}\mathbf{V}\mathbf{M}(\xi^*)^{-1}\right)\mathbf{f}(x)$$
$$\leq \operatorname{tr}\left(\left((n-1)(\mathbf{M}(\xi^*)+\mathbf{\Delta}^{-1})^{-1}\mathbf{M}(\xi^*)(\mathbf{M}(\xi^*)+\mathbf{\Delta}^{-1})^{-1}+\mathbf{M}(\xi^*)^{-1}\right)\mathbf{V}\right)$$

For any support point x_j of ξ^* equality holds

Special case: random intercepts

- $f_1(x) \equiv 1$
- $\mathbf{D} = d_1 \mathbf{e}_1 \mathbf{e}_1^\top, \; \mathbf{e}_1 = (1, 0, ..., 0)^\top$

Objective function:

$$\mathrm{IMSE}_{\mathrm{pred}}(\xi) = \sigma^2 \left(\frac{1}{m} \operatorname{tr} \left(\mathbf{M}(\xi)^{-1} \mathbf{V} \right) + \frac{(n-1) d_1}{1 + m \, d_1} \nu(\mathcal{X}) \right)$$

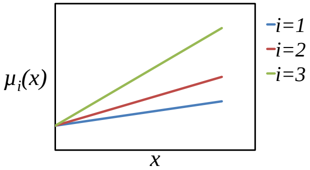
IMSE-optimal design in fixed effects model remains IMSE-optimal for prediction

Example: straight line regression

$$Y_{ij} = \beta_{i1} + \beta_{i2} x_j + \varepsilon_{ij}, \qquad x_{ij} \in [0, 1]$$

Only slope is random

- $\beta_{i1} \equiv \beta_1$
- $\mathbf{D} = d_2 \mathbf{e}_2 \mathbf{e}_2^\top$

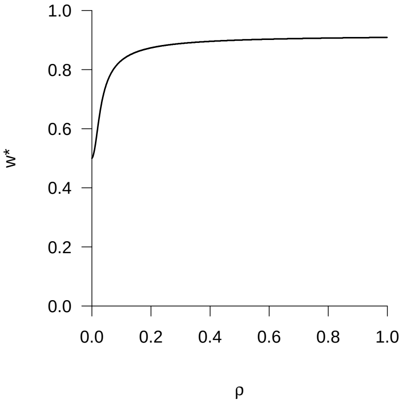


Individual design:

$$\xi^* = \begin{pmatrix} 0 & 1 \\ 1 - w^* & w^* \end{pmatrix}$$

Optimal weight

- $n = 100$
- $m = 10$
- $\rho = d_2/(1 + d_2)$



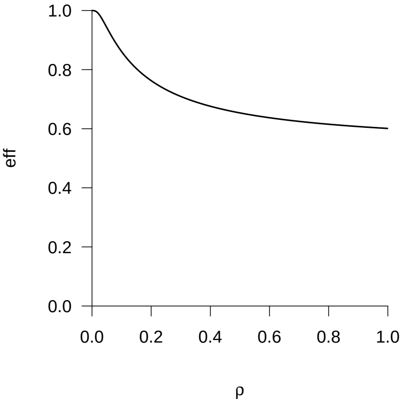
For small values of d_1

$$w^*(d_1) \approx w^* \quad \& \quad \textit{Bayesian optimal design:} \quad w_B^* = 1$$

Efficiency of equi-replicated design

- $n = 100$
- $m = 10$
- Equi-replicated design:

$$w = 0.5$$



Efficiency of Bayesian optimal design for small values of d_1: $eff(\xi_B^*) = 0$

Outlook

- Exact design
- Different designs for different individuals
- Other design criteria
- ...

Thank you for your attention!

OTTO VON GUERICKE UNIVERSITÄT MAGDEBURG