Sparse Sampling *D*-Optimal Designs

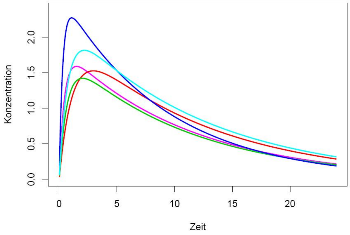
Tobias Mielke

BMBF grant SKAVOE / Otto-von-Guericke University Magdeburg

PODE-Workshop 2009 - St. Petersburg

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):

$$\begin{aligned} Y_{ij} &= \eta(x_{ij}, \beta_i) + \epsilon_{ij},\ j = 1, ..., m_i,\ \epsilon_{ij} \sim N(0, \sigma^2) \\ &= f(x_{ij})^T \beta_i + \epsilon_{ij},\ \text{in linear cases} \end{aligned}$$

- 2. stage (inter-individual variation):

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

- b_i and ϵ_{ij} are assumed to be independent.

Estimation

For linear regression functions:

- $Y_i \sim N_{m_i}(F_i\beta, \sigma^2 V_i)$, where
  - m_i - number of observations for individual i,
  - $F_i := (f(x_{i1}), ..., f(x_{im_i}))^T$ - design matrix of individual i,
  - $V_i := I_{m_i} + F_i D F_i^T$.
- $Y = F\beta + Gb + \epsilon$, where
  - $Y := (Y_1^T, ..., Y_n^T)^T$,
  - $F := (F_1^T, ..., F_n^T)^T$,
  - $G := diag((F_1, ..., F_n))$ and
  - $V := diag(V_1, ..., V_n)$.

It follows $Y \sim N_{\sum m_i}(F\beta, \sigma^2 V)$.

Estimation

- For linear regression functions:
  - $\hat{\beta} = (F^T V^{-1} F)^{-1} F^T V^{-1} Y$ is the ML-Estimator
  - $Cov(\hat{\beta}) = \sigma^2 (F^T V^{-1} F)^{-1} = \sigma^2 \mathfrak{M}^{-1}$
- Information:

$$\mathfrak{M} = \sum_{i=1}^{n} \mathfrak{M}_{ind,i} = \sum_{i=1}^{n} F_i^T V_i^{-1} F_i.$$

- For nonlinear regression functions:
  - Use of 2-stage procedures
  - Maximum likelihoood estimation

Design-Problem

- Experimental settings ξ_i for individual i: $x_{ij} \in X$ with m_{ij} observations:

$$\xi_i = \begin{pmatrix} x_{i1} & \dots & x_{ik_i} \\ m_{i1} & \dots & m_{ik_i} \end{pmatrix}, \sum_{j=1}^{k_i} m_{ik_j} = m_i.$$

- Population design ζ:

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

ξ_j: Individual design, ω_j: weight of indivdual design ξ_j in the population.

Design-Problem

Choose design, such that:

$$Cov(\hat{\beta}) = \sigma^2 (F^T V^{-1} F)^{-1} = \sigma^2 \mathfrak{M}(\zeta)^{-1}$$

is minimal. $\rightarrow$ Optimality-criteria

- *c*-optimality,
- *A*-optimality,
- *D*-optimality,
- ...

How to determine the optimal design?

$\rightarrow$ Use of equivalence theorems

Linear Regression

Assume:

- $Y_i = b_{i1} + b_{i2}x_i = f(x_i)^T b_i$, $i = 1, ..., n$, with
- $b_i \sim N_2(\beta, D)$, $D = diag(d_1, d_2)$ and $x_i \in [-1, 1]$.
- vector of regression functions is given by

$$f(x_i) = (1, x_i)^T, \; x_i \in [-1, 1]$$

- variance function:

$$\sigma^2(x_i) := f(x_i)^T D f(x_i).$$

Linear Regression

Equivalence theorem in this situation:

- A design ζ^* is *D*-optimal for estimating $\beta \in \mathbb{R}^p$ if and only if

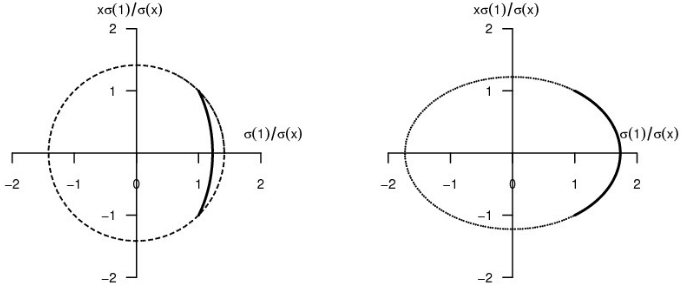
$$g_{\zeta^*}(x) := \frac{f(x)^T \mathfrak{M}(\zeta^*)^{-1} f(x)}{\sigma^2(x)} \leq p \text{ for all } x \in X.$$

- $g_{\zeta^*}(x) = p$, for $x \in \zeta^*$.

Linear Regression

Geometric Interpretation:

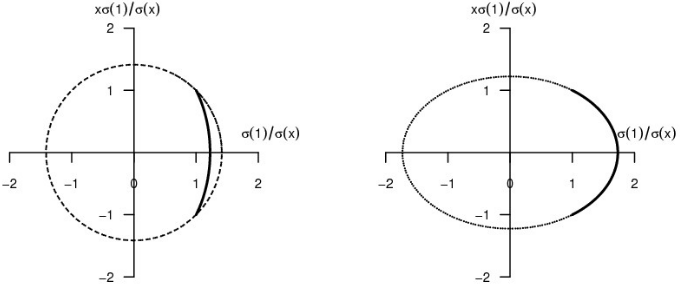
- Transformation $t_1(x) = \frac{\sigma(1)}{\sigma(x)}$ and $t_2(x) = \frac{x\sigma(1)}{\sigma(x)}$
- Design locus $T = \{(t_1(x), t_2(x)); x \in [-1, 1]\}$



Linear Regression

D-optimal design:

- For $d_1 \geq d_2$: $\zeta_{x^*,-x^*}$ is *D*-optimal, where $x^* = 1$.
- For $d_1 < d_2$: $\zeta_{x^*,-x^*}$ is *D*-optimal, where $x^* = \sqrt{d_1/d_2}$.



Quadratic Regression

Assume:

- $Y_{ij} = b_{i1} + b_{i2}x_{ij} + b_{i3}x_{ij}^2 + \epsilon_{ij}$, with
- 2 observations per individual,
- $b_i \sim N_3(\beta, \sigma^2 D_k)$, $k = 1, 2, 3$ where $D_1 = diag(d_1, 0, 0)$, $D_2 = diag(0, d_2, 0)$ or $D_3 = diag(0, 0, d_3)$,
- $X = [-1, 1]$ and
- $\epsilon_{ij} \sim N(0, \sigma^2)$.

Quadratic Regression

- 2 observations per individual in points $x_i, y_i \in X$.

$$\begin{aligned} \Rightarrow \mathfrak{M}_{ind}(\xi_i) &= F_i^T V_i^{-1} F_i, \text{ with} \\ F_i &= F_{(x_i,y_i)} = \begin{pmatrix} 1 & x_i & x_i^2 \\ 1 & y_i & y_i^2 \end{pmatrix} \text{ and} \\ V_i &= V_{(x_i,y_i)} = I_2 + F_i D F_i^T . \end{aligned}$$

- Invariance considerations yield:

$$\mathfrak{M}_{pop}(\zeta) = \sum_{i=1}^{k} \omega_i \mathfrak{M}_{ind}(\xi_i) = \begin{pmatrix} a & 0 & b \\ 0 & c & 0 \\ b & 0 & d \end{pmatrix} .$$

Quadratic Regression

Multivariate equivalence theorem:

- The design ζ^* is *D*-optimal,
- The design ζ^* minimizes $\max_\xi Tr\,(\mathfrak{M}(\zeta^*)^{-1}\mathfrak{M}(\xi))$,
- $\max_\xi Tr\,(\mathfrak{M}(\zeta^*)^{-1}\mathfrak{M}(\xi)) = p$.

For our case:

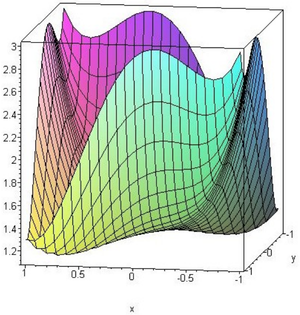
- $g_{\zeta^*}(x,y) := Tr\ F_{(x,y)}\mathfrak{M}_{pop}(\zeta^*)^{-1}F_{(x,y)}^T V_{(x,y)}^{-1} \leq 3$

Quadratic Regression

Optimal designs are of the structure:

1. For $D = D_1 = diag(d_1, 0, 0)$ with $\alpha_{d_1} \in (-1, 1)$ and $\omega_{d_1} \in (0, 1)$:

$$\zeta^*_{d_1} = \begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 1 \\ \alpha_{d_1} \end{pmatrix} & \begin{pmatrix} -1 \\ -\alpha_{d_1} \end{pmatrix} \\ \omega_{d_1} & \frac{1}{2}(1 - \omega_{d_1}) & \frac{1}{2}(1 - \omega_{d_1}) \end{pmatrix},$$

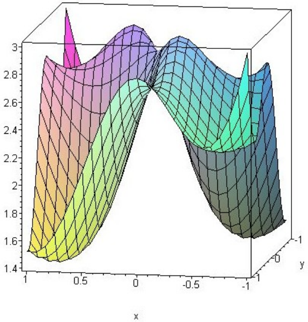


Quadratic Regression

Optimal designs are of the structure:

2. For $D = D_2 = diag(0, d_2, 0)$ with $\alpha_{d_2} \in [0, 1)$ and $\omega_{d_2} \in (0, 1)$:

$$\zeta^*_{d_2} = \begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} \alpha_{d_2} \\ -\alpha_{d_2} \end{pmatrix} \\ 1 - \omega_{d_2} & \omega_{d_2} \end{pmatrix}$$

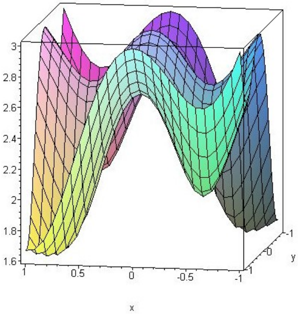


Quadratic Regression

Optimal designs are of the structure:

3. For $D = D_3 = diag(0, 0, d_3)$:

$$\zeta^*_{d_3} = \begin{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} & \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ 2/3 & 1/3 \end{pmatrix}$$



Quadratic Regression

Optimal designs are of the structure:

4. For $D = diag(d_1, d_2, d_3)$ the design can be constructed numerically:
   - Determine a regular initial design ζ_0.
   - Calculate the maxima $\xi^* := (x^*, y^*)$ of the sensitivity function $g_{\zeta_0}(x, y)$.
   - Determine the matrix $\mathfrak{M}(\zeta_k) := \omega\mathfrak{M}(\zeta_{k-1}) + (1-\omega)\mathfrak{M}(\xi^*)$ and maximize its determinant with respect to ω. Set $\zeta_k := \omega\zeta_{k-1} + (1-\omega)\xi^*$.
   - When no $(x, y) \in X^2$ exists with $g_{\zeta_k}(x, y) > 3$ then the design ζ_k is D-optimal.

Summary

- Estimation of population parameters
  - $\longrightarrow$ Reliable in linear models.
  - $\longrightarrow$ Reliable in nonlinear models?
- Construction of *D*-optimal population designs using the equivalence theorem
  - $\longrightarrow$ Design-structure depends on the variance in the parameter vector.
  - $\longrightarrow$ Generalization for nonlinear models?

Thank you for your attention!