Efficiencies of designs in the simple linear regression model with correlated errors

Point of departure: We consider designs \( \xi \) with \( n \) experimental units. Each unit consists of \( q \) design points \( (x_1, \ldots, x_q) \). They are the same for all units.

Model for \( n = 1 \):

\[
y = X_0 \beta + \epsilon
\]

with \( y = \begin{bmatrix} y_1 \\ \vdots \\ y_q \end{bmatrix}, X_0 = \begin{bmatrix} 1 & x_1 \\ \vdots \\ 1 & x_q \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_q \end{bmatrix}, \epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_q \end{bmatrix}. \]

The vector \( \epsilon \) is normally distributed with \( \text{E}(\epsilon) = 0 \) and \( \text{Cov}(\epsilon) = \Sigma(\sigma^2, \lambda) \).

Here \( \Sigma(\sigma^2, \lambda) = \begin{bmatrix} \lambda & \cdots & \lambda \xi & \cdots & \lambda \xi q -1 \xi & \cdots & \lambda \xi q -1 \xi \\ \vdots \\ \lambda & \cdots & \lambda \xi & \cdots & \lambda \xi q -1 \xi & \cdots & \lambda \xi q -1 \xi \\ \vdots \\ 1 & \cdots & \lambda & \cdots & \lambda \xi & \cdots & \lambda \xi q -1 \xi \\ \cdots \\ 1 & \cdots & 1 & \cdots & \lambda & \cdots & \lambda \xi q -1 \xi \end{bmatrix} \) with \( \lambda \in [0, 1] \).

Correlations depend on the distances of the corresponding design points.

Model for \( n \geq 1 \):

\[
y = [y_1, \ldots, y_q] = [y_1, \ldots, y_q]^T, \text{Cov}(\epsilon) = \Pi_n \otimes \Sigma(\sigma^2, \lambda) =: W, \text{design matrix: } \Pi_n \otimes X =: X. \]

Errors within one unit are correlated. Errors between different units are independent.

Parameters \( \beta_0 \) (intercept) and \( \beta_q \) (slope) are of main interest, \( \sigma^2 \) (error variance) and \( \lambda \) (correlation parameter) are nuisance parameters.

D-criterion and D-optimality:

\[
D(\xi) = (\text{det}(I_1 - g))^{1/2} \text{ with information matrix } I_1.
\]

\[
eff(\xi) = D(\xi)/D(\xi^*), \text{ in which design } \xi^* \text{ is D-optimal.}
\]

In what follows:

- \( q = 3, x_j = (0, 1, 1) \), \( x_1 = 0 \) and \( x_3 = 1 \), because the D-optimal designs observe at 0 and 1 (see [1]).

- How to choose \( x_2 = d \).

Results in Dette et al. (2008):

- \( d \) and \( d - 1 \) lead to the same D-criterion.
- For \( \lambda \to 1 \): \( d = 1/2 \) is D-optimal.
- For \( \lambda \in (0, 1) \): only locally D-optimal designs exist, see Figure 1.

![Figure 1](image1)

Two-step maximum likelihood estimation:

Likelihood function:

\[
L_d(\beta, \sigma^2, \lambda) = \frac{1}{(2\pi)^n/2 |\text{det}(W)|} \exp(-1/2 (y - \hat{X} \hat{\beta})^T W^{-1} (y - \hat{X} \hat{\beta})).
\]

- Step 1:
  1. Determine \( \hat{\beta}_{\text{CLS}} \) with \( \lambda = 0 \). Determine residuals \( \hat{\epsilon} \).
  2. Replace \( y - \hat{X} \hat{\beta} \) with \( \hat{\epsilon} \Rightarrow L_d(\sigma^2, \lambda) \). Determine \( \sigma^2 \) and \( \lambda \) by ML-estimators.

- Step 2:
  3. Determine \( \hat{\beta}_{\text{CLS}} \) with \( \lambda \) out of (2). Determine residuals \( \hat{\epsilon} \).

- Repetition of two-step procedure:
  4. Repeat (2) and (3) until estimates converge. Thereby use in (2) residuals of (3) instead of (1).


Simulation study:

\( n \in \{2, 50\}, \lambda \in \{0, 0.25, 0.5, 0.75, 0.9\}, d \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6\} \), 10000 datasets for all combinations.

New results:

- \( d \in [0, 1] \):
  1. \( d \in [0, 1] \): \( \text{eff}(\xi) \geq 0.7071 \) (\( \lambda \in [0, 1] \)).
  2. \( d \in [0, 1] \): \( \text{eff}(\xi) \geq 0.59 \) (\( \lambda \in [0.2, 1] \)).
  3. \( d \in [0, 1] \): \( \text{lim}_{\lambda \to 0} \text{eff}(\xi) = 1 \).
  4. \( d = 0 \): min \( \text{eff}(\xi) = 0.7071 \).
  5. \( d = 0.5 \): \( \text{eff}(\xi) \geq 0.58 \) (\( \lambda \in [0, 1] \)).

- \( d \in \{1/8, 0.0, 1, \ldots, 8\} \):
  1. \( d = 0 \): \( \text{min \text{eff}(\xi) = 0.7493} \).
  2. \( d = 1/8 \): \( \text{min \text{eff}(\xi) = 0.9220} \).
  3. \( d = 1/4 \): \( \text{min \text{eff}(\xi) = 0.9014} \).
  4. \( d = 3/8 \): \( \text{min \text{eff}(\xi) = 0.8750} \).
  5. \( d = 1/2 \): \( \text{min \text{eff}(\xi) = 0.8660} \).

Results:

First line: empirical D-criterion (blue = locally D-optimal design (empirical))

Second line: theoretical D-criterion (red = locally D-optimal design (nuisance parameters are known))

<table>
<thead>
<tr>
<th>( d )</th>
<th>( n = 2 )</th>
<th>( n = 50 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0.2</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2</td>
<td>0.3</td>
<td>0.5</td>
</tr>
<tr>
<td>0.3</td>
<td>0.4</td>
<td>0.5</td>
</tr>
<tr>
<td>0.4</td>
<td>0.5</td>
<td>0.6</td>
</tr>
<tr>
<td>0.5</td>
<td>0.6</td>
<td>0.7</td>
</tr>
<tr>
<td>0.6</td>
<td>0.7</td>
<td>0.8</td>
</tr>
<tr>
<td>0.7</td>
<td>0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>0.8</td>
<td>0.9</td>
<td>1.0</td>
</tr>
<tr>
<td>0.9</td>
<td>1.0</td>
<td>1.1</td>
</tr>
</tbody>
</table>

- Empirical D-criterion is close to the theoretical D-criterion.

- Only locally D-optimal designs exist.

- Results for known nuisance parameters can be transferred to the case of unknown parameters even if the number of experimental units is small.

References
