Some optimal criteria of model-robustness for two-level non-regular fractional factorial designs

Satoshi Aoki (Kagoshima Univ.)

JST CREST Conference in Osaka
29 June 2010

Contents

1. Introduction

- Regular designs have many merits in two-level fractional factorial designs.
  - Elegant theory based on a linear algebra over $GF(2)$ is established.
  - Many desirable properties (orthogonality, balanced etc)
  - Concepts such as resolution and aberration can be considered easily.

Only drawback: run size must be a power of 2.

$\Longrightarrow$ non-regular designs are also important.

- **designs of m factors with n runs**
- **Fundamental question: How to choose a non-regular design for given (m, n) ?**
- **Generalized Minimum Aberration criterion (Deng and Tang, 1999)**
  - **Extension of Minimum Aberration criterion for non-regular designs**
  - **MA criterion (Fries and Hunter, 1980): criterion based on the hierarchical assumption**

    Lower order interaction is more important.

    Resolution III $<$ Resolution IV $<$ Resolution V

- **GMA criterion is a natural extension of MA criterion**
- **Cheng, Deng and Tang (2002):**
  evaluate GMA criterion from the viewpoint of model-robustness

  In this study, we give an extension of Cheng, Deng and Tang (2002)

  See Aoki (2010) *AISM* **for detail.**

- **Another motivation:**
  Affinely full-dimensional factorial design
  (Aoki and Takemura, 2009)
- **Aoki and Takemura (2009):**
  We give a new class of non-regular design and investigate its property from the viewpoint of D-optimality for the main effect model.
  $\Longrightarrow$ **However, characteristic for the models including interaction effects is unknown.**

In this study, relation among the affinely full-dimensionality, GMA criterion and our proposed criterion are shown.

2. GMA criterion

- d: design of m factors with n runs
- d is shown as an $n \times m$ matrix $X(d) \in \{-1, +1\}^{n \times m}$.
- $S = \{j_1, \ldots, j_s\} \subseteq \{1, \ldots, m\}$: subset of $\{1, \ldots, m\}$
- $\mathbf{x}_S(d)$: component-wise product of the $j_1, \ldots, j_s$-th columns of $X(d)$.
  - $\mathbf{x}_S(d)$ is a column vector.
  - Write the component-wise product of two column vectors as $\odot$,
  $$\mathbf{x}_S(d) \odot \mathbf{x}_T(d) = \mathbf{x}_{S \triangle T}(d)$$
  holds, where
  $$S \triangle T = (S \cup T) \setminus (S \cap T).$$

- $|S|$: **cardinality of** S
- $j_S(d)$: **sum of the elements of** $\mathbf{x}_S(d)$
- **Define** $B_s(d)$ **characteristic as**

$$B_s(d) = \sum_{S:\ |S|=s} \left(\frac{j_S(d)}{n} \right)^2$$

for $s = 1, \ldots, m$**.**

- **Example:** 2^{5-2}_{III} **design** (ABD = ACE = I)

A	B	C	D	E
1	1	1	1	1
1	1	−1	1	−1
1	−1	1	−1	1
1	−1	−1	−1	−1
−1	1	1	−1	−1
−1	1	−1	−1	1
−1	−1	1	1	−1
−1	−1	−1	1	1

From $\mathbf{x}_{\{1,2,4\}}(d) = (1,1,1,1,1,1,1,1)'$,

$j_{\{1,2,4\}}(d) = 8$

Similarly, $j_{\{1,3,5\}}(d) = j_{\{2,3,4,5\}}(d) = 8$

$j_S(d) = 0$ **for all the other** S

Therefore we have

$B_1(d) = B_2(d) = 0$

$$B_3(d) = \left(\frac{8}{8}\right)^2 + \left(\frac{8}{8}\right)^2 = 2$$

$$B_4(d) = \left(\frac{8}{8}\right)^2 = 1$$

- **There is one-to-one relation between design d and $j_S(d),\ S \subseteq \{1,\ldots,m\}$**
  - **The relation to the indicator function is also given ($j_S(d)/n = b_S/b_\emptyset$).**

$$j_{\{1,2,4\}}(d) = j_{\{1,3,5\}}(d) = j_{\{2,3,4,5\}}(d) = 8$$

$$f(\mathbf{x}) = \frac{1}{4} + \frac{1}{4}(x_1x_2x_4 + x_1x_3x_5 + x_2x_3x_4x_5)$$

  - **For regular designs, $j_S(d)/n$ must be one of $\{-1,0,1\}$**
    ($\Longleftrightarrow$ All the nonzero coefficient equal to the constant term in the indicator function.)

- $B_s(d),\ s = 1, \ldots, m$ **has the information of the aberration of the design** d**.**
  - $B_1(d) = 0$ **if the levels are equireplicated for each factor.**
  - $B_2(d) = 0$ **for the orthogonal designs.**
  - $B_3(d) = 0$ **for regular designs of the resolution IV.**
  - $B_3(d) = B_4(d) = 0$ **for regular designs of the resolution V.**

- **GMA criterion (Tang and Deng, 1999)**
  sequentially minimize $B_1(d), B_2(d), \ldots, B_m(d)$**.**
- **Affinely full-dimensional factorial design**
  (Aoki and Takemura, 2009)
  $|j_S(d)/n| < 1$ **for all** $S \subset \{1, \ldots, m\}$**.**

- **Relation between GMA criterion and affinely full-dimensionality**

From the definition $B_s(d) = \sum_{S:\ |S|=s} \left(\frac{j_S(d)}{n} \right)^2$**,**

restriction $\left| \frac{j_S(d)}{n} \right| < 1$ **coincides with minimizing** $B_s(d), s = |S|$ **to some extent.**

  - **GMA: minimizing** $B_1(d), B_2(d), \ldots$ **sequentially**
  - **aff. full-dim: minimizing** $B_1(d), B_2(d), \ldots$ **simultaneously**

We investigate this relation from the viewpoint of the model-robustness.

3. Optimal criteria for model-robustness

- **We are interested in robust designs when the interaction effects cannot be ignored.**
- **Situation considered:**
  - **all the main effects are of primary interest and their estimates are required.**
  - **there are f active two-factor interaction and g active three-factor interaction effects, but which of two- and three-factor interactions are active is unknown.**
  - **all the four- and higher-factor interactions are negligible.**

- $\mathcal{P}$**: set of all the subsets of the size two from** $\{1,\ldots,m\}$**, i.e.,** $\mathcal{P}=\{\{1,2\},\{1,3\},\ldots,\{m-1,m\}\}$**.**

$$|\mathcal{P}|=\binom{m}{2}=F$$

- $\mathcal{Q}$**: set of all the subsets of the size three from** $\{1,\ldots,m\}$**, i.e.,** $\mathcal{Q}=\{\{1,2,3\},\{1,2,4\},\ldots,\{m-2,m-1,m\}\}$

$$|\mathcal{Q}|=\binom{m}{3}=G$$

- $\mathcal{F}\subset\mathcal{P},\mathcal{G}\subset\mathcal{Q}$**: active two- and three–factor interaction effects.**

$$|\mathcal{F}|=f,\quad |\mathcal{G}|=g$$

- $\mathcal{F}, \mathcal{G}$ **are unknown. However, it is natural to restrict the models to satisfy the following hierarchical assumption:**

  Definition $\mathcal{F}$ **and** $\mathcal{G}$ **are called hierarchically consistent if**

$$(i_1, i_2, i_3) \in \mathcal{G} \implies (i_1, i_2), (i_1, i_3), (i_2, i_3) \in \mathcal{F}.$$

- **For given $\mathcal{F}, \mathcal{G}$, we consider the linear model**

$$\mathbf{y} = \mu \mathbf{1}_n + X(d)\beta_1 + Y_{\mathcal{F}}(d)\beta_2 + Z_{\mathcal{G}}(d)\beta_3 + \varepsilon.$$

  - **y: $n \times 1$ vector of observations**
  - μ**: mean parameter**
  - $X(d)$**: $n \times m$ matrix**
  - β_1**: $m \times 1$ vector of the main effect**
  - $Y_{\mathcal{F}}(d)$**: $n \times f$ matrix consisting of the columns $\mathbf{x}_S(d), S \in \mathcal{F}$**
  - β_2**: $f \times 1$ vector of active two-factor interaction**
  - $Z_{\mathcal{G}}(d)$**: $n \times g$ matrix consisting of the columns $\mathbf{x}_S(d), S \in \mathcal{G}$**
  - β_3**: $g \times 1$ vector of active three-factor interaction**
  - ε**: random vector satisfying $E(\varepsilon) = \mathbf{0}, var(\varepsilon) = \sigma^2 I_n$**

- **Write** $X_{\mathcal{F},\mathcal{G}} = [\mathbf{1}_n \,\vdots\, X(d) \,\vdots\, Y_{\mathcal{F}}(d) \,\vdots\, Z_{\mathcal{G}}(d)]$.
- **The information matrix of** d:

$$M_{\mathcal{F},\mathcal{G}}(d) = \frac{1}{n} X_{\mathcal{F},\mathcal{G}}(d)' X_{\mathcal{F},\mathcal{G}}(d)$$
$$= \begin{bmatrix} 1 & \frac{1}{n}\mathbf{1}_n' X(d) & \frac{1}{n}\mathbf{1}_n' Y_{\mathcal{F}}(d) & \frac{1}{n}\mathbf{1}_n' Z_{\mathcal{G}}(d) \\ \frac{1}{n} X(d)'\mathbf{1}_n & \frac{1}{n} X(d)'X(d) & \frac{1}{n} X(d)'Y_{\mathcal{F}}(d) & \frac{1}{n} X(d)'Z_{\mathcal{G}}(d) \\ \frac{1}{n} Y_{\mathcal{F}}(d)'\mathbf{1}_n & \frac{1}{n} Y_{\mathcal{F}}(d)'X(d) & \frac{1}{n} Y_{\mathcal{F}}(d)'Y_{\mathcal{F}}(d) & \frac{1}{n} Y_{\mathcal{F}}(d)'Z_{\mathcal{G}}(d) \\ \frac{1}{n} Z_{\mathcal{G}}(d)'\mathbf{1}_n & \frac{1}{n} Z_{\mathcal{G}}(d)'X(d) & \frac{1}{n} Z_{\mathcal{G}}(d)'Y_{\mathcal{F}}(d) & \frac{1}{n} Z_{\mathcal{G}}(d)'Z_{\mathcal{G}}(d) \end{bmatrix}$$

- **If the model $\mathcal{F}, \mathcal{G}$ is known, we can rely on various optimal criteria based on $M_{\mathcal{F},\mathcal{G}}(d)$ to choose d.**

For example, D-optimal design is to maximize $\det M_{\mathcal{F},\mathcal{G}}(d)$.

- **For unknown $\mathcal{F},\mathcal{G}$, we consider the average performance over all possible combinations of f two-factor and g three-factor interactions.**
- **This corresponds to consider the expectation $D_{f,g} = E_p[\det M_{\mathcal{F},\mathcal{G}}(d)]$ for the uniform distribution**

  $$p(\mathcal{F},\mathcal{G}) = \begin{cases} \text{Const}, & \textbf{if } \mathcal{F} \textbf{ and } \mathcal{G} \textbf{ are hierarchically consistent} \\ 0, & \textbf{otherwise} \end{cases}$$

  over $2^{\mathcal{P}}, 2^{\mathcal{Q}}$.

  We call this $D_{f,g}$-optimal criterion.

- **However, it is difficult to evaluate** $\det M_{\mathcal{F},\mathcal{G}}(d)$**.**
  We consider minimizing $\mathrm{tr}(M_{\mathcal{F},\mathcal{G}}(d))^2$ **instead of maximizing** $\det M_{\mathcal{F},\mathcal{G}}(d)$**.**

  It is known that this is a good surrogate (Cheng, 1996).

- **Since all the diagonal elements of** $M_{\mathcal{F},\mathcal{G}}(d)$ **is** 1**, minimizing** $E_p[\mathrm{tr}(M_{\mathcal{F},\mathcal{G}}(d))^2]$ **is equivalent to minimizing**

  $$S^2_{f,g} = E_p\left[\sum (\textbf{off-diagonal elements of } M_{\mathcal{F},\mathcal{G}}(d))^2\right]$$

  We call this as $S^2_{f,g}$**-criterion.**

- **It is also difficult to derive a closed-form expression of $S^2_{f,g}$ in general.**
  In this study, expressions of $S^2_{f,g}$ for some specific settings are shown.
- **Approach: evaluate the off-diagonal elements of the information matrix, respectively.**

$$
\begin{aligned}
S^2_{f,g} &= 2B_1(d) + 2B_2(d) + 2E_p\left[\frac{1}{n^2}\sum_{S\in\mathcal{F}}(j_S(d))^2\right] + 2E_p\left[\frac{1}{n^2}\sum_{i=1}^{m}\sum_{S\in\mathcal{F}}(j_{\{i\}\triangle S}(d))^2\right] \\
&+E_p\left[\frac{1}{n^2}\sum\sum_{\substack{S,T\in\mathcal{F}\\ S\neq T}}(j_{S\triangle T}(d))^2\right] + 2E_p\left[\frac{1}{n^2}\sum_{S\in\mathcal{G}}(j_S(d))^2\right] + 2E_p\left[\frac{1}{n^2}\sum_{i=1}^{m}\sum_{S\in\mathcal{G}}(j_{\{i\}\triangle S}(d))^2\right] \\
&+2E_p\left[\frac{1}{n^2}\sum_{S\in\mathcal{F}}\sum_{T\in\mathcal{G}}(j_{S\triangle T}(d))^2\right] + E_p\left[\frac{1}{n^2}\sum\sum_{\substack{S,T\in\mathcal{G}\\ S\neq T}}(j_{S\triangle T}(d))^2\right]
\end{aligned}
$$

- **Point:** $S^2_{f,g}$ **is expressed as a linear combination of** $B_1(d), B_2(d), \ldots, B_6(d)$ **if the support of** $p(\mathcal{F}, \mathcal{G})$ **is symmetric over** $\{1, \ldots, m\}$**.**

- **Theorem** $S_{f,0}^2$ **is expressed as** $S_{f,0}^2 = \sum_{s=1}^{4} a_s B_s(d)$**,**

  where

$$
\begin{aligned}
a_1 &= 2\left(1 + \frac{f(m-1)}{F}\right), \\
a_2 &= 2\left(1 + \frac{f}{F} + \frac{f(f-1)}{F(F-1)}(m-2)\right), \\
a_3 &= \frac{6f}{F}, \quad a_4 = \frac{6f(f-1)}{F(F-1)}.
\end{aligned}
$$

  We have $a_1 > a_2 > a_3 > a_4$ **for** $m > 3$**.**

- **This is an optimal criterion for the case that three-factor interaction is negligible and there are** f **active two-factor interactions.**

- **Theorem** $S^2_{F,g}$ **is expressed as** $S^2_{F,g} = \sum_{s=1}^{6} a_s B_s(d)$**,**

  where

$$
\begin{aligned}
a_1 &= 2m + \frac{g(m-1)(m-2)}{G},\\
a_2 &= 2m + \frac{2g(m-2)}{G} + \frac{g(g-1)(m-2)(m-3)}{G(G-1)},\\
a_3 &= 6 + \frac{2g}{G} + \frac{6g(m-3)}{G},\\
a_4 &= 6 + \frac{8g}{G} + \frac{6g(g-1)(m-4)}{G(G-1)}, \quad a_5 = \frac{20g}{G}, \quad a_6 = \frac{20g(g-1)}{G(G-1)}
\end{aligned}
$$

  We have $a_1 > a_2 > a_3 > a_4 > a_5 > a_6$ **for** $m > 5$**.**

- **This is an optimal criterion for the case that all the two-factor interactions are active and there are** g **active three-factor interactions.**

- **Theorem** $S^2_{3,1}$ **is expressed as** $S^2_{3,1} = \sum_{s=1}^{4} a_s B_s(d)$**,**

  where

$$
\begin{aligned}
a_1 &= 2\left(1 + \frac{9}{m}\right), \\
a_2 &= 2\left((m-1) + \frac{4(m-2)}{G}\right), \\
a_3 &= \frac{2(3m-5)}{G}, \quad a_4 = \frac{8}{G}
\end{aligned}
$$

  We have $a_2 > a_1 > a_3 > a_4$ **for** $m > 3$**.**

- **This is an optimal criterion for the case that there is one active three-factor interaction and three active two-factor interactions included in the three-factor interaction hierarchically.**

- **From the coefficients of $S^2_{f,g} = \sum_s a_s B_s(d)$, the relations among the GMA criterion, aff. full-dim design and the robust designs are shown.**

- **For example of $S^2_{F,g}$, since $a_1 > \cdots > a_6$ holds for $m \geq 5$, there is a consistency between the GMA criterion and the $S^2_{F,g}$-criterion for $m \geq 5$ to some extent.**

- **On the other hand, $a_1 < a_2$ holds for $S^2_{3,1}$. Therefore $S^2_{3,1}$-criterion puts the importance on the orthogonality rather than the equireplicated design.**

4. Empirical studies

- **We consider the $S^2_{f,0}$- and $S^2_{3,1}$-optimal design for 5 factor designs with 12 runs.**
- **An important point:**
  the existence of Hadamard matrices

  It is easy to construct designs with $B_1(d) = B_2(d) = 0$.

  We are interested in the existence of the robust resigns satisfying $B_1(d) \neq 0$ or $B_2(d) \neq 0$.

- **Result: $S^2_{f,0}$-, $S^2_{3,1}$-optimal design**

1	1	1	1	1
1	1	−1	−1	−1
1	−1	1	1	−1
1	−1	1	−1	1
1	−1	−1	1	1
−1	1	1	1	−1
−1	1	1	−1	1
−1	1	−1	1	1
−1	−1	1	1	1
−1	−1	1	−1	−1
−1	−1	−1	1	−1
−1	−1	−1	−1	1

- **For $S^2_{f,0}$-, $S^2_{3,1}$-optimal designs (d_s),**

$$B_1(d_s) = 0.138889, \;\; B_2(d_s) = 0,$$
$$B_3(d_s) = 0.27778, \;\; B_4(d_s) = 0.5556$$

holds. On the other hand, for Hadamard design (d_h),

$$B_1(d_h) = B_2(d_h) = 0,$$
$$B_3(d_h) = 1.1111, \quad B_4(d_h) = 0.5556$$

holds, which is GMA optimal design.

	$S^2_{1,0}$	$S^2_{2,0}$	$S^2_{3,0}$	$S^2_{4,0}$	$S^2_{5,0}$	$S^2_{3,1}$
d_s	0.5556	0.9074	1.3333	1.8333	2.4074	1.7778
d_h	0.6667	1.4074	2.2222	3.1111	4.0741	2.6667

- **All these optimal designs are also affinely full-dimensional.**

  Therefore the simple strategy that "choose 12 points from regular designs" is bad from the both optimality.

5. Summary

- **We formalize a method to construct a robust two-level fractional factorial designs.**
- **It is true that the assumption that the experimenters only have an information on the number of the interactions in the true model seems unnatural in actual situations. However, we think that the $S^2_{f,g}$ values for small f, g can be used to evaluate the model-robustness. Here we regard f and g as the degree of contamination of interactions.**

- **Though the calculations of $S^2_{f,g}$ for large f, g will be rather complicated, they are indeed based on a simple counting.**
- **In application, there are also many important situation that the support of $p(\mathcal{F}, \mathcal{G})$ is asymmetric.**
- **For the cases of asymmetric support, the evaluation of $S^2_{f,g}$ will be very difficult since $S^2_{f,g}$ cannot be expressed as a linear combination of $B_s(d)$.**