Generalized wordtype pattern for nonregular factorial designs with multiple groups of factors

Mingyao Ai

Department of Probability and Statistics

School of Mathematical Sciences

Peking University

myai@math.pku.edu.cn

(Based on the joint work with Profs Shuyuan He and Runchu Zhang)

Contents

1. Generalized Wordtype Pattern

Wordtype Pattern

- For a regular $2^{(l_1+l_2)-k}$ design D containing l_1 Group I factors and l_2 Group II factors, let $A_{i_1,i_2}(D)$ be the number of words in the defining contrast subgroup containing i_1 Group I factors and i_2 Group II factors. Zhu (Ann, 2003) called $[A_{i_1,i_2}(D)]$ the wordtype pattern of D.
- $A_i = \sum_{i_1+i_2=i} A_{i_1,i_2}$ is just the popular wordlength pattern of D.

Generalized Wordtype Pattern

- For a factorial $(n; s_1 \cdots s_{l_1}, s_{l_1+1} \cdots s_{l_1+l_2})$-design D, let $R_I = R_{s_1} \times \cdots \times R_{s_{l_1}}$, $R_{II} = R_{s_{l_1+1}} \times \cdots \times R_{s_{l_1+l_2}}$ and $R = R_I \times R_{II}$. Following the similar notations of Xu and Wu (Ann, 2001), define

$$B_{i_1,i_2}(D) = n^{-2} \sum_{wt(u_1)=i_1, wt(u_2)=i_2} |\chi_u(D)|^2, \qquad (1)$$

  where $u = (u_1, u_2)$, $u_1 \in R_I$, $u_2 \in R_{II}$, $\chi_u(D) = \sum_{x \in D} \chi_u(x)$, the above summations are over all $u \in R = R_I \times R_{II}$ with $wt(u_1) = i_1, wt(u_2) = i_2$, and $\{\chi_u, u \in R\}$ are the given orthonormal contrasts.

- Similarly, $[B_{i_1,i_2}(D)]$ is called the generalized wordtype pattern of design D.

- $A_i(D) = \sum_{i_1+i_2=i} B_{i_1,i_2}(D)$ is just the generalized wordlength pattern.
- The generalized wordtype pattern is the MacWilliams transform of the double distance distribution, that is,

$$B_{i_1,i_2}(D) = E'_{i_1,i_2}(D).$$

2. Consulting Design Theory

2.1. Regular Symmetrical Factorial Designs

Let H_t be the regular saturated design with s^t runs. An $s^{n-(n-t)}$ design D can be considered as a set of n columns in H_t. $H_t = [D, \overline{D}]$. $\overline{D}$ is called the *complementary design* of D. Then Tang and Wu (Ann, 1996) and Suen, Chen and Wu (Ann, 1997) showed that sequentially minimizing

$$A_i(D), \;\; i = 3, \ldots, n$$

is equivalent to sequentially minimizing

$$(-1)^i A_i(\overline{D}), \;\; i = 3, \ldots, f, \tag{2}$$

where $f = L_t - n$ and $L_t = (s^t - 1)/(s - 1)$.

2.2. Regular Mixed-level Factorial Designs

$(s^r)s^n$ Factorial Designs

Consider an $(s^r)s^n$ factorial design $D = [D_0, D_T]$ in s^t runs, involving one s^r-level factor $(r \geq 2)$, grouped by the L_r s-level factors, and n s-level factors.

Wu and Zhang (Biometrika, 1993) partitioned the words of the same length of $D = [D_0, D_T]$ into two types, type 0 and type 1, depending on whether they contain any factor in D_0 and suggested the following ordering of wordlength pattern:

$$\{A_{i,0}, A_{i,1}\}_{3 \leq i \leq n+1}. \tag{3}$$

$$H_t = (D_0, D_T, D_Q)$$

is a column partition of H_t after several column permutations such that $D = [D_0, D_T]$.

$$D_R = [D_0, D_Q]$$

is called the *consulting design* of D, which corresponds to an $(s^r)s^f$ design, where $f = L_t - L_r - n$.

Mukerjee and Wu (Sinica, 2001) showed that sequentially minimizing $\{A_{3,0}(D), A_{3,1}(D), A_{4,0}(D), A_{4,1}(D)\}$ is equivalent to sequentially minimizing

$$-G_3(D_R), -G_3(D_Q), G_4(D_R), G_4(D_Q),$$

where $G_i(Q) = (s-1)^{-1}\#\{\beta : wt(\beta) = i, Q\beta = 0\}$ $[= A_i(Q)]$.

Ai and Zhang (Statist Papers, 2005) further showed that sequentially minimizing $\{A_{i,0}(D), A_{i,1}(D)\}$ for $i = 3, \ldots, n+1$ is equivalent to sequentially minimizing

$$\{(-1)^i[A_{i,0}(D_R) + A_{i,1}(D_R)],\ (-1)^i A_{i,0}(D_R)\}_{3\le i\le f+1}. \tag{4}$$

$(s^r)^2 s^n$ Factorial Designs

Consider an $(s^r)^2 s^n$ factorial design $D = [D_{01}, D_{02}, D_t]$ in s^t runs, involving two s^r-level factor $(r \geq 2)$, grouped by the mutually exclusive $2L_r$ s-level factors, and n s-level factors. The ordering of wordlength pattern is as follows:

$$\{A_{i,0}, A_{i,1}, A_{i,2}\}_{3\leq i\leq n+2}. \tag{5}$$

$H_t = (D_{01}, D_{02}, D_T, D_Q)$ is a column partition of H_t after several column permutations such that $D = [D_{01}, D_{02}, D_T]$.

$$D_R = [D_{01}, D_{02}, D_Q]$$

is called the *consulting design* of D, which corresponds to an $(s^r)^2 s^f$ design, where $f = L_t - 2L_r - n$.

Ai and Zhang (Statist Papers, 2005) showed that sequentially minimizing

$$\{A_{i,0}(D), A_{i,1}(D), A_{i,2}(D)\},\ i = 3, \ldots, n+2,$$

is equivalent to sequentially minimizing

$$\Big\{(-1)^i \sum_{u=0}^{2} A_{i,u}(D_R), (-1)^i[2A_{i,0}(D_R) + A_{i,1}(D_R)], (-1)^i A_{i,0}(D_R)\Big\}_{3\le i\le f+2.} \qquad (6)$$

2.3. Blocked Regular Mixed-level Factorial Designs

For a blocked regular $(s^{n-(n-t)} : s^k)$-design, Zhang and Park (JSPI, 2000) and Ai and Zhang (Canad J Statist, 2004) suggested the ordering of wordlength pattern as:

$$A_3^t, A_2^b, A_4^t, A_5^t, A_3^b, A_6^t, \ldots. \qquad (7)$$

Consider a blocked regular $((s^r)s^n : s^k)$-design $D = [D_B, D_0, D_T]$ in s^k blocks.

$H_t = [D_B, D_0, D_T, D_Q]$ is a column partition of H_t after several column permutations.

$$D_R = [D_B, D_0, D_Q]$$

is called the *consulting design* of D, which corresponds to an $((s^r)s^f : s^k)$-design, where $f = L_t - L_k - L_r - n$.

Ai and Zhang (JSPI, 2004) showed that sequentially minimizing the first six terms

$$A^t_{3,0}(D), A^t_{3,1}(D), A^b_{2,0}(D), A^b_{2,1}(D), A^t_{4,0}(D), A^t_{4,1}(D),$$

is equivalent to sequentially minimizing the following six terms of D_R:

$$\begin{aligned}
&-[A^t_{3,0}(D_R) + A^t_{3,1}(D_R) + A^b_{2,0}(D_R) + A^b_{2,1}(D_R)],\\
&-[A^t_{3,0}(D_R) + A^b_{2,0}(D_R)],\\
&[A^b_{2,0}(D_R) + A^b_{2,1}(D_R)],\ \ A^b_{2,0}(D_R),\\
&[A^t_{4,0}(D_R) + A^t_{4,1}(D_R) + A^b_{3,0}(D_R) + A^b_{3,1}(D_R)],\\
&[A^t_{4,0}(D_R) + A^b_{3,0}(D_R)].
\end{aligned} \tag{8}$$

Note that Chen and Cheng (Ann, 1999) considered blocked regular two-level designs and suggested a new mixed ordering.

Remark: Similar result for blocked regular $((s^{r_1})(s^{r_2})s^n : s^k)$-designs.

2.4. Blocked Nonregular Factorial Designs

For unblocked symmetrical case, $H = (D, \overline{D})$. Xu and Wu (Ann, 2001) showed that sequentially minimizing $A_i(D),\ i = 3, \ldots, n$ is equivalent to sequentially minimizing $(-1)^i A_i(\overline{D}),\ i = 3, \ldots, f$.

Let H be a saturated $OA(N, s^p, 2)$. $H = (D_T, D_B, D_C)$ is a column partition of H after several column permutations such that D_T consists of the n treatment factors and D_B consists of the r independent block columns. Thus the blocked $(N, s^n : s^r)$-design D can be viewed as the matrix (D_T, D_B). The matrix $D_R = (D_C, D_B)$ corresponding to a blocked $(N, s^{p-n-r} : s^r)$-design is called the blocked *consulting design* of D in H.

$$W_1(D) = (A_1^b(D), A_3^t(D), A_2^b(D), A_4^t(D), A_5^t(D), A_3^b(D), \ldots),$$

Ai and Zhang (Canad J Statist, 2004) showed that sequentially minimizing the components of the combined GWP $W_1(D)$ of D is equivalent to sequentially minimizing the following components of D_R:

$$\{-A_1^b(D_R), -A_3(D_R), A_2^b(D_R), A_4(D_R), -A_5(D_R), -[A_3^t(D_R) + A_3^b(D_R)], A_6(D_R), \ldots\}. \quad (9)$$

Note that the above general rule *no longer* holds for blocked nonregular mixed-level designs. Nevertheless, the following weak result can be obtained from Xu (Sinica, 2003):

$$A_3^t(D) = -A_3(D_R) + \text{constant}. \quad (10)$$

2.5. Designs with Multiple Groups of Factors

Let H be a saturated $OA(N, s_1^{l_1} s_2^m, 2)$.

$$H = (D_1, D_2, D_3)$$

is a column partition of H after several column permutations such that $D = [D_1, D_2]$.

$$D_R = [D_1, D_3]$$

corresponding to a new $(N; s_1^{l_1}, s_2^{m-l_2})$-design is called the *consulting design* of D in H. Let

$$\theta_{i,j}(T, n, m, s) = s^{-m} \sum_{k=0}^{m} P_i(T-k; n, s) P_k(j; m, s).$$

Then

$$B_{j_1,j_2}(D) = \text{constant} + \sum_{k_1=0}^{l_1} \sum_{k_2=0}^{j_2} c_{j_1,j_2;k_1,k_2} B_{k_1,k_2}(D_R), \tag{11}$$

where

$$c_{j_1,j_2;k_1,k_2} = s_1^{-l_1} \sum_{i=0}^{l_1} \theta_{j_2,k_2}((N - s_1 i)s_2^{-1}, l_2, m - l_2, s_2) \\ P_{j_1}(i; l_1, s_1) P_i(k_1; l_1, s_1).$$

3. Selection of Optimal Single Arrays

A symmetrical single array $D = (D_1, D_2)$ with N runs, in which the first l_1 group I factors are the noise factors and the rest l_2 group II factors are the control factors, is a factorial $(N; s^{l_1}, s^{l_2})$-design. Similar to Wu and Zhu (Technometrics, 2003), we assume that all effects with order ≥ 3 are negligible.

Define the index vector $J = (J_1, J_2, J_3, J_4, J_5, J_6)$, where $J_1 = B_{1,2}(D) + B_{2,1}(D) + B_{2,2}(D)$, $J_2 = 3B_{0,3}(D) + 3B_{1,3}(D) + B_{1,2}(D)$, $J_3 = B_{2,1}(D) + 3B_{3,1}(D) + 3B_{3,0}(D)$, $J_4 = B_{0,4}(D)$, $J_5 = B_{2,2}(D)$, and $J_6 = B_{4,0}(D)$. Then the generalized minimum J-aberration (GMJA) criterion is to sequentially minimize J_i for $i = 1, \ldots, 6$.

An $(N; 2^{l_1}, 2^{l_2})$ single array D has GMJA within the class of designs derived from Hadamard matrices of order N if and only if its consulting design D_R is the unique $(N; 2^{l_1}, 2^{N-1-(l_1+l_2)})$-design that sequentially minimizes the first i $(1 \le i \le 6)$ components in the following sequence:

$$\begin{aligned}
&\sum_{j+k=3} jB_{j,k}(D_R) + [B_{2,2}(D_R) + 3B_{3,1}(D_R) + 6B_{4,0}(D_R)],\\
&-\sum_{j+k=3} (3+2j)B_{j,k}(D_R) - \sum_{j+k=4} 3jB_{j,k}(D_R),\\
&-[B_{2,1}(D_R) + 3B_{3,0}(D_R)] - [3B_{3,1}(D_R) + 12B_{4,0}(D_R)],\\
&\sum_{j+k=3} B_{j,k}(D_R) + \sum_{j+k=4} B_{j,k}(D_R),\\
&[B_{2,1}(D_R) + 3B_{3,0}(D_R)] + [B_{2,2}(D_R) + 3B_{3,1}(D_R) + 6B_{4,0}(D_R)],\\
&B_{4,0}(D_R).
\end{aligned} \tag{12}$$

4. An Illustration

As an illustration, Tables 2-4 only tabulates GMJA single arrays derived from a specific Hadamard matrix of order 16, that is, Hall's $OA(16, 2^{15}, 2)$ of type III given in Appendix 7B of Wu and Hamada (2000), which is shown in Table 1. Note that the columns Col.(C) and Col.(N) list the control and noise factor columns, respectively. For comparison, the last column $(J_1, J_2, J_3, J_4, J_5, J_6)_R$ presents the aliasing index vectors of minimum J-aberration regular single arrays in Table C.2 of Wu and Zhu (Technometrics, 2003).

Table 1: Hall's $OA(16, 2^{15}, 2)$ of type III

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
−	−	−	−	−	−	−	−	−	−	−	−	−	−	−
−	−	−	−	−	−	−	+	+	+	+	+	+	+	+
−	−	−	+	+	+	+	−	−	−	−	+	+	+	+
−	−	−	+	+	+	+	+	+	+	+	−	−	−	−
−	+	+	−	−	+	+	−	−	+	+	−	−	+	+
−	+	+	−	−	+	+	+	+	−	−	+	+	−	−
−	+	+	+	+	−	−	−	−	+	+	+	+	−	−
−	+	+	+	+	−	−	+	+	−	−	−	−	+	+
+	−	+	−	+	−	+	−	+	−	+	−	+	−	+
+	−	+	−	+	−	+	+	−	+	−	+	−	+	−
+	−	+	+	−	+	−	−	+	+	−	−	+	+	−
+	−	+	+	−	+	−	+	−	−	+	+	−	−	+
+	+	−	−	+	+	−	−	+	+	−	+	−	−	+
+	+	−	−	+	+	−	+	−	−	+	−	+	+	−
+	+	−	+	−	−	+	−	+	−	+	+	−	+	−
+	+	−	+	−	−	+	+	−	+	−	−	+	−	+

Table 2: GMJA $(16; 2^1, 2^{l_2})$ single arrays from $OA(16, 2^{15}, 2)$ in Table 1

l_2	Col.(C)	Col.(N)	$(J_1, J_2, J_3, J_4, J_5, J_6)$	$(J_1, J_2, J_3, J_4, J_5, J_6)_R$
2	2,4	1	0 0 0 0 0 0	0 0 0 0 0 0
3	2,4,8	1	0 0 0 0 0 0	0 0 0 0 0 0
4	8,10,13,14	1	0 0 0 0 0 0	0 0 0 0 0 0
5	2,8,10,13,14	1	0 6 0 0 0 0	0 6 0 0 0 0
6	2,4,8,10,13,14	1	0 12 0 1 0 0	0 12 0 6 0 0
7	2,4,6,8,10,12,15	1	0 21 0 3 0 0	0 21 0 18 0 0
8	2-4,6,8,10,12,15	1	1 31 0 5 0 0	4 31 0 30 0 0
9	2-6,8,10,12,15	1	2 44 0 9 0 0	8 44 0 54 0 0
10	2-6,8-10,13,14	1	3 60 0 16 0 0	12 60 0 96 0 0
11	2,4,6,8-15	1	4 79 0 26 0 0	16 79 0 156 0 0
12	2-12,14	1	5 107 0 38 0 0	20 107 0 228 0 0
13	2-14	1	6 138 0 55 0 0	24 138 0 330 0 0

Table 3: GMJA $(16; 2^2, 2^{l_2})$ single arrays from $OA(16, 2^{15}, 2)$ in Table 1

l_2	Col.(C)	Col.(N)	$(J_1, J_2, J_3, J_4, J_5, J_6)$	$(J_1, J_2, J_3, J_4, J_5, J_6)_R$
2	4,8	1,2	0 0 0 0 0 0	0 0 0 0 0 0
3	2,5,8	1,4	0 0 0 0 0 0	0 0 0 0 0 0
4	8,11,12,15	1,10	0 0 0 1 0 0	4 0 1 6 0 0
5	1,6-9	2,4	1 6 0 1 1 0	8 6 0 6 2 0
6	1-4,6,7	8,10	2 12 0 3 2 0	12 12 0 18 3 0
7	1-7	8,10	3 21 0 7 3 0	16 21 1 42 3 0
8	1-7,9	8,10	5 41 0 7 3 0	24 41 1 42 3 0
9	1,6-13	2,4	7 64 0 10 3 0	32 64 1 60 3 0
10	2-7,9,10,12,15	1,8	9 93 0 16 3 0	40 93 1 96 3 0
11	2,3,5-12,14	1,4	12 125 0 25 4 0	52 125 1 150 4 0
12	2,3,5-14	1,4	15 163 0 38 5 0	64 163 1 228 5 0

Table 4: GMJA $(16; 2^3, 2^{l_2})$ single arrays from $OA(16, 2^{15}, 2)$ in Table 1

l_2	Col.(C)	Col.(N)	$(J_1, J_2, J_3, J_4, J_5, J_6)$	$(J_1, J_2, J_3, J_4, J_5, J_6)_R$
2	7,9	1,2,8	0 0 0 0 0 0	0 0 0 0 0 0
3	11,13,14	1,8,12	0 3 0 0 0 0	0 3 3 0 0 0
4	9,11,13,14	1,10,12	1 6 0 0 1 0	8 7 3 0 1 0
5	4,9,11,13,14	1,8,10	3 13 0 0 2 0	16 14 3 0 2 0
6	2-4,6,11,13	1,10,12	5 23 0 1 3 0	24 27 3 0 3 0
7	2-7,9	1,8,14	7 43 0 3 3 0	36 43 5 18 3 0
8	2-7,11,13	1,10,12	10 65 0 5 5 0	52 63 5 30 5 0
9	2-7,9,13,15	1,8,14	14 91 0 9 7 0	68 91 6 54 7 0
10	2-7,9,10,13,15	1,8,14	18 129 0 15 9 0	84 129 6 90 9 0
11	2,3,5-7,9,10,12-15	1,4,8	24 168 0 25 12 0	108 168 6 150 12 0

It can be seen that all the single arrays in Tables 2-4 have less or no more GMJA than the corresponding regular single arrays. Moreover, the discrepancy between the two index vectors becomes large as the numbers of control and noise factors increase.

Acknowledgements

- Many thanks go to my PhD supervisor, Prof Runchu Zhang, and my Postdoc mentor, Prof Shuyuan He, for their continuous encouragement.
- This work was partially supported by National Natural Science Foundation of China (Grant Nos. 10571093, 10471127).

Thank You All for Your Attention!