AKCE J.Graphs. Combin., 1, No. 2 (2004), pp. 83- 98

OPTIMUM MULTILEVEL ORTHOGONAL ARRAYS FOR CORRELATED ERRORS

MIWAKO MISHIMA
Information and Multimedia Center, Gifu University
1-1 Yanagido, Gifu 501-1193, Japan
e-mail: miwako@info.gifu-u.ac.jp
and
MASAKAZU JIMBO
Graduate School of Information Science, Nagoya University
Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
e-mail: jimbo@is.nagoya-u.ac.jp

Communicated by: Mariko Hagita
Received : December 06, 2004

Abstract

In the case where errors are correlated depending on the closeness of the experiments, the optimality of multilevel orthogonal arrays is discussed and some advantages of linear orthogonal arrays are also given.

Keywords: Hamming distance, Hamming weight, nearest-neighbor correlation structure, orthogonal array, universally optimum, weakly universally optimum.

AMS Subject Classification: 62K05, 62K15, 05B15.

1. Preliminaries

In time series analysis or spatial data analysis, correlated data are often treated (see, for example, Bennet [1] and Upton and Fingleton [11]). Similarly, in factorial experiments, the "closeness" (resemblance) of experiments is supposed to cause the correlation between observations. Several known results on this subject are due to Cheng [3], Jimbo [4], Kiefer [5], Kiefer and Wynn [6] and Mishima *et al.* [9]. Kiefer [5] introduced a general notion of optimality. Kiefer and Wynn [6] discussed the optimality of balanced incomplete block designs and Latin square (or Latin hypercube) designs for correlated observations, and gave several constructions for the optimum designs (see also Cheng [3]). Jimbo [4] provided optimum two-level orthogonal arrays for ordinary least-square estimators of main effects under a certain correlation structure of errors. Mishima *et al.* [9] showed similar results to [4] for generalized least-square estimators of main effects. In this article, as a generalization of the study due to Mishima *et al.* [9], optimum multilevel orthogonal arrays are presented under the same covariance structures as discussed in [9].

Let $A_1, \ldots, A_m$ be m factors (treatment) with q levels each and $\Gamma = (\gamma_{ij})$ be an array of assemblies (level-combinations) of factors, where $\gamma_{ij} \in \mathbb{Z}_q$ is the level of the j th factor for the i th experiment. We assume that there are no interaction effects between these factors and consider the following model:

$$\boldsymbol{y} = \mu \mathbf{1}_N + X\boldsymbol{\alpha} + \boldsymbol{\epsilon}, \qquad \text{cov}(\boldsymbol{\epsilon}) = \Sigma. \tag{1.1}$$

In model (1.1), μ represents a general mean, $\mathbf{1}_N$ is the N-dimensional all-one column vector, and $\boldsymbol{y} = (y_1, \ldots, y_N)'$ is an N-dimensional observation vector, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_N)'$ is an N-dimensional error vector and $\boldsymbol{\alpha} = (\alpha^{(1)'}, \ldots, \alpha^{(m)'})' = (\alpha_0^{(1)}, \ldots, \alpha_{q-1}^{(1)} | \cdots | \alpha_0^{(m)}, \ldots, \alpha_{q-1}^{(m)})'$ is an mq-dimensional main effect vector, where $\alpha_{\gamma_{ij}}^{(j)}$ is the main effect for the γ_{ij} th level of the j th factor. Without loss of generality, we can assume

$$\alpha_0^{(j)} + \alpha_1^{(j)} + \cdots + \alpha_{q-1}^{(j)} = 0$$

for $j = 1, \ldots, m$. Further $X = (X^{(1)}, \ldots, X^{(m)})$ is an $N \times mq$ design matrix corresponding to an $N \times m$ array $\Gamma = (\gamma_{ij})$ of assemblies, where $X^{(j)} = (x_{ik}^{(j)})$ is an $N \times q$ submatrix of X such that

$$x_{ik}^{(j)} = \begin{cases} 1 & \text{if } \gamma_{ij} = k, \\ 0 & \text{otherwise.} \end{cases}$$

Here the first column of $X^{(j)}$ is referred to as the 0 th column. Since the parameters in model (1.1) are not independent, the normal equation can not be solved uniquely for $\boldsymbol{\alpha}$. If this is the case, we transform the model to a manageable one. Consider a $q \times q$ orthogonal matrix whose first row is $(1/\sqrt{q})(1, \ldots, 1)$ and define Q as the $(q-1) \times q$ matrix obtained by deleting the first row. Let $\beta^{(j)} = Q\alpha^{(j)}$ and let $W = (W^{(1)}, \ldots, W^{(m)})$ be the $N \times m(q-1)$ matrix such that

$$X^{(j)}\alpha^{(j)} = X^{(j)}Q'\beta^{(j)} = W^{(j)}\beta^{(j)}. \tag{1.2}$$

Then, model (1.1) can be rewritten as follows:

$$\boldsymbol{y} = \mu \mathbf{1}_N + W\boldsymbol{\beta} + \boldsymbol{\epsilon}, \qquad \text{cov}(\boldsymbol{\epsilon}) = \Sigma, \tag{1.3}$$

where $\boldsymbol{\beta} = (\beta^{(1)'}, \ldots, \beta^{(m)'})' = (\beta_1^{(1)}, \ldots, \beta_{q-1}^{(1)} | \cdots | \beta_1^{(m)}, \ldots, \beta_{q-1}^{(m)})'$ is the transformed main effect vector. The ordinary least-square estimator (OLSE) and the generalized least-square estimator (GLSE) of $\boldsymbol{\beta}$, both of which are denoted by $\hat{\boldsymbol{\beta}}$ unless otherwise stated, can be given as solutions of respective normal equations

$$\left(W'W - \frac{1}{N}W'J_NW\right)\hat{\boldsymbol{\beta}} = \left(W' - \frac{1}{N}W'J_N\right)\boldsymbol{y}$$

and

$$\begin{aligned}&\left(W'\Sigma^{-1}W - (\mathbf{1}_N'\Sigma^{-1}\mathbf{1}_N)^{-1}W'\Sigma^{-1}J_N\Sigma^{-1}W\right)\hat{\boldsymbol{\beta}} \\ &= \left(W'\Sigma^{-1} - (\mathbf{1}_N'\Sigma^{-1}\mathbf{1}_N)^{-1}W'\Sigma^{-1}J_N\Sigma^{-1}\right)\boldsymbol{y},\end{aligned}$$

where J_N is the $N \times N$ all-one matrix.

There are many types of "optimality" criteria for the evaluation of the efficiency of $\hat{\boldsymbol{\beta}}$. Among them, we adopt two simple and sufficient criteria called *universally optimum* and *weakly universally optimum* that were formalized by Kiefer [5] and Kiefer and Wynn [6]. The criterion of universally optimum includes A-, D-, E-optimality as its special cases and the criterion of weakly universally optimum includes A- and E-optimality (see [5] and [6]). Here we cite two propositions concerning those criterion that are needed for the discussion in the rest of this article. For an $N \times m$ array Γ and its corresponding $N \times m(q-1)$ matrix W, define $D(\Gamma)$ and $C(\Gamma)$ respectively by

$$D(\Gamma) = \text{cov}(\,\hat{\boldsymbol{\beta}}\,) = \begin{cases} (W'W - \frac{1}{N}W'J_NW)^{-1}(W' - \frac{1}{N}W'J_N)\Sigma \\ \qquad (W - \frac{1}{N}J_NW)(W'W - \frac{1}{N}W'J_NW)^{-1} \\ \hfill \text{for the OLSE } \hat{\boldsymbol{\beta}}, \\ (W'\Sigma^{-1}W - (\mathbf{1}_N'\Sigma^{-1}\mathbf{1}_N)^{-1}W'\Sigma^{-1}J_N\Sigma^{-1}W)^{-1} \\ \hfill \text{for the GLSE } \hat{\boldsymbol{\beta}} \end{cases} \tag{1.4}$$

and $C(\Gamma) = D(\Gamma)^{-1}$. Note that $C(\Gamma)$ for the OLSE $\hat{\boldsymbol{\beta}}$ becomes the so-called C-matrix (information matrix) when $\text{cov}(\,\boldsymbol{\epsilon}\,) = \sigma^2 I_N$.

Let Ξ be the set of $N \times m$ arrays Γ by which the main effect $\boldsymbol{\beta}$ is estimable by the ordinary least-square method, or the generalized least-square method.

Proposition 1.1 (Kiefer [5]). *An array $\Gamma^* \in \Xi$ is universally optimum relative to Ξ if*

(a) $C(\Gamma^*) = cI_{m(q-1)}$ *for some constant c and*

(b) $\text{tr}(C(\Gamma^*)) = \max_{\Gamma \in \Xi} \text{tr}(C(\Gamma))$,

where I_n is the $n \times n$ identity matrix and tr(C) *denotes the trace of a matrix C.*

Proposition 1.2 (Kiefer and Wynn [6]). *An array $\Gamma^* \in \Xi$ is weakly universally optimum relative to Ξ if*

(a)′ $D(\Gamma^*) = cI_{m(q-1)}$ *for some constant c and*

(b)′ $\text{tr}(D(\Gamma^*)) = \min_{\Gamma \in \Xi} \text{tr}(D(\Gamma))$.

Note that if Γ^* is universally optimum relative to Ξ, then it is weakly universally optimum as well. It is also known that if Γ^* is (weakly) universally optimum for the estimation of $\boldsymbol{\alpha}$, then it is (weakly) universally optimum for the estimation of $\boldsymbol{\beta}$, and the converse is also true. So, we treat only the transformed model (1.3) throughout this article.

An $N \times m$ array $\Gamma = (\gamma_{ij})$, where $\gamma_{ij} \in \mathbb{Z}_q$, is called an *orthogonal array* of size N, m constraints, q levels and strength 2, denoted by OA$\,(N, m, q, 2)$, if any $N \times 2$ submatrix of Γ contains every ordered pair from $\mathbb{Z}_q \times \mathbb{Z}_q$ with the same frequency N/q^2 (for a more general definition, see Beth *et al.* [2] and Raghavarao

[10]). Without loss of generality, we may assume that the first row of Γ is the all-zero vector. For the $N \times m(q-1)$ matrix W corresponding to Γ, since

$$W' J_N = O \quad \text{and} \quad W'W = \frac{N}{q} I_{m(q-1)} \tag{1.5}$$

hold, where O is the all-zero matrix, the covariance matrix (1.4) for the OLSE of $\boldsymbol{\beta}$ is reduced to

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \frac{q^2}{N^2} W' \Sigma W.$$

In the case where cov $(\boldsymbol{\epsilon}) = \sigma^2 I_N$, it is well known that an OA $(N, m, q, 2)$ is universally optimum among the arrays by which $\boldsymbol{\beta}$ is estimable (see [5]). Our objective is to find an optimum array Γ not only for the case where $\mathrm{cov}(\epsilon) = \sigma^2 I_N$ but also for a more general class of covariance structures given in the following sections. The optimality of orthogonal arrays for the OLSE of $\boldsymbol{\beta}$ is discussed in Section 2. In Section 3, some advantages of linear orthogonal arrays are shown for the OLSE and the GLSE of $\boldsymbol{\beta}$. In Section 4, optimum orthogonal arrays for a nearest-neighbor covariance structure are presented by using the result in Section 3. Section 5 provides the covariance matrices of main effect $\boldsymbol{\beta}$ in the case of the complete factorial designs under the covariance structures discussed in the previous sections.

2. Optimum orthogonal arrays for OLSE

Let $\gamma_i = (\gamma_{i1}, \ldots, \gamma_{im})$ be the i th row of an $N \times m$ array Γ, where $\gamma_{ij} \in \mathbb{Z}_q$. For the i th and the k th rows of Γ, the number of j such that $\gamma_{ij} \neq \gamma_{kj}$ is called the *Hamming distance* between γ_i and γ_k, denoted by $d(\gamma_i, \gamma_k)$, and the number of non-zero coordinates of γ_i is called the *Hamming weight* of γ_i, denoted by $w(\gamma_i)$.

Considering models such as (1.1) and (1.3), we usually assume that cov $(\boldsymbol{\epsilon}) = \sigma^2 I_N$. However, the real covariance structure might be different from this. As a likely correlation structure, we suppose that the closeness of two experiments causes some correlation between the corresponding two errors and measure the closeness of experiments by the Hamming distance between two rows in an array Γ of assemblies. If

$$\ell = \min\{d(\gamma_i, \gamma_j) : \text{any two distinct rows } \gamma_i \text{ and } \gamma_j \text{ of } \Gamma\},$$

then Γ is called an array with *minimum distance* ℓ. Now, suppose that the real covariance structure is as follows:

Covariance structure I:

$$\mathrm{cov}(\epsilon_i, \epsilon_k) \begin{cases} = \sigma^2 & \text{if } i = k, \\ \geq 0 & \text{if } d(\gamma_i, \gamma_k) \leq \ell, \\ = 0 & \text{if } d(\gamma_i, \gamma_k) > \ell, \end{cases}$$

where $\ell \leq \lfloor m(q-1)/q \rfloor$ ($\lfloor a \rfloor$ indicates the greatest integer not exceeding a).

When we do not know the real covariance structure, we usually adopt the OLSE as an estimate of the main effect $\boldsymbol{\beta}$. In this section, among the optimum arrays of assemblies under the covariance structure $\text{cov}(\boldsymbol{\epsilon}) = \sigma^2 I_N$, we find an array which is still optimum under Covariance structure I.

Let Ξ be the set of OA $(N, m, q, 2)$.

Theorem 2.1. *Under Covariance structure I, if $\Gamma^* \in \Xi$ is an array with minimum distance $\ell + 1$, then Γ^* is weakly universally optimum relative to Ξ and for the OLSE $\hat{\boldsymbol{\beta}}$,*

$$\text{cov}(\hat{\boldsymbol{\beta}}) = \frac{q}{N}\sigma^2 I_{m(q-1)}$$

holds without reference to Covariance structure I.

Proof. Let W be the $N \times m(q-1)$ matrix corresponding to $\Gamma \in \Xi$ as defined in Section 1. It follows from (1.5) that

$$\text{cov}(\hat{\boldsymbol{\beta}}) = \frac{q^2}{N^2} W' \Sigma W = \frac{q^2}{N^2} \sum_{i,k} \text{cov}(\epsilon_i, \epsilon_k) w_i' w_k,$$

where $w_i = (w_{i1}^{(1)}, \dots, w_{i,q-1}^{(1)} | \cdots | w_{i1}^{(m)}, \dots, w_{i,q-1}^{(m)})$ is the ith row of $W = (W^{(1)}, \dots, W^{(m)})$. Note that the jth block $(w_{i1}^{(j)}, \dots, w_{i,q-1}^{(j)})$ of w_i is represented by one of columns in Q which is defined by (1.2). Taking account of $Q'Q = I_q - (1/q)J_q$ and $\text{tr}(w_i' w_k) = \text{tr}(w_k w_i') = (m - d(\gamma_i, \gamma_k))(q-1)/q - d(\gamma_i, \gamma_k)/q$, we have

$$\begin{aligned}
\text{tr}(\text{cov}(\hat{\boldsymbol{\beta}})) &= \text{tr}(D(\Gamma)) \\
&= \frac{q^2}{N^2} \sum_{i,k} \text{cov}(\epsilon_i, \epsilon_k) \text{tr}(w_i' w_k) \\
&= \frac{q(q-1)}{N} m\sigma^2 \\
&\quad + \frac{q^2}{N^2} \sum_{i \neq k} \text{cov}(\epsilon_i, \epsilon_k) \left(\frac{q-1}{q}(m - d(\gamma_i, \gamma_k)) - \frac{1}{q} d(\gamma_i, \gamma_k) \right) \\
&\geq \frac{q(q-1)}{N} m\sigma^2
\end{aligned}$$

and equality holds if and only if $d(\gamma_i^*, \gamma_k^*) > \ell$ for any two rows $\gamma_i^*, \gamma_k^* \in \Gamma^*$. Thus Proposition 1.2 completes the proof. □

Let q be a prime power. For $m \geq n$, take m distinct n-dimensional column vectors $g_1, \dots, g_m$ from the vector space GF $(q)^n$ and let $G = (g_1, \dots, g_m)$. Furthermore, let Γ be the $q^n \times m$ array obtained by arranging q^n m-dimensional row vectors θG $(\theta \in \text{GF}(q)^n)$. If the set $\{\theta G : \theta \in \text{GF}(q)^n\}$ is an n-dimensional linear subspace of GF $(q)^m$, then Γ is called a *linear* OA $(q^n, m, q, 2)$ generated from G and G is called a *generator matrix*. It should be mentioned that constructing nonlinear orthogonal arrays is not that simple as the number m of constraints (factors) gets larger. In what follows, we treat linear orthogonal arrays only. The next theorem is useful to construct an array Γ which satisfies the minimum distance condition in Theorem 2.1.

Theorem 2.2. *Let* $G = (I_n|K)$ *be an* $n\times m$ *matrix over* GF(q) *for some* $n\times(m-n)$ *matrix* K *and let* Γ^* *be a linear* OA $(N, m, q, 2)$ *generated from* G. *If any* ℓ *distinct column vectors of* $H = (-K'|I_{m-n})$ *are linearly independent, then the array* Γ^* *is weakly universally optimum under Covariance structure I.*

Proof. The theorem is a direct consequence of a well-known result of coding theory: A linear code with a parity check matrix H has minimum distance at least $\ell+1$ if any ℓ distinct column vectors of H are linearly independent (see, for example, MacWilliams and Sloane [7]). □

3. Advantages of linear orthogonal arrays

Covariance structure I describes that there are no correlation between the errors if the corresponding experiments are a certain Hamming distance apart, and Theorem 2.2 gives a construction for optimum arrays under Covariance structure I by using linear orthogonal arrays. In this section, considering a more general class of covariance structures such that

$$\text{cov}(\epsilon_i, \epsilon_k) = \begin{cases} \sigma^2 & \text{if } i = k, \\ \sigma^2 \rho_{d(\gamma_i,\gamma_k)} & \text{if } i \neq k, \end{cases} \tag{3.1}$$

we discuss some other advantages of linear orthogonal arrays when we estimate the OLSE and the GLSE of $\boldsymbol{\beta}$. A sufficient condition for the covariance matrices of the OLSE and the GLSE to coincide is obtained as well.

Now, define $N \times N$ matrices $D_l = (d_{ik}^{(l)})$ by

$$d_{ik}^{(l)} = \begin{cases} 1 & \text{if } d(\gamma_i, \gamma_k) = l, \\ 0 & \text{otherwise} \end{cases}$$

for $l = 0, \dots, m$. The matrix D_l is called the l th *adjacency matrix* of Γ. Then, with D_l $(l = 0, \dots, m)$, we rewrite the covariance structure (3.1) as follows:

$$\text{cov}(\boldsymbol{\epsilon}) = \Sigma = \sigma^2 \left(I_N + \rho_0(D_0 - I_N) + \sum_{l=1}^{m} \rho_l D_l \right), \tag{3.2}$$

where ρ_l is the correlation coefficient for two experiments of the Hamming distance l. Normally, it is natural to assume that $\rho_0 \geq \rho_1 \geq \cdots \geq \rho_m \geq 0$. If Γ is a linear orthogonal array, then there are no repeated rows. Thus the 0th adjacency matrix of Γ is given by $D_0 = I_N$. This means that $d(\gamma_i, \gamma_k) = 0$ holds if and only if $i = k$. By setting $\rho_0 = 1$, we can further reduce the covariance matrix (3.2) to

$$\text{cov}(\boldsymbol{\epsilon}) = \Sigma = \sigma^2 \sum_{l=0}^{m} \rho_l D_l. \tag{3.3}$$

In a linear orthogonal array Γ, the number of rows with Hamming distance l from a given row γ_i is constant, not depending on the choice of γ_i. Thus $D_l\mathbf{1}_N = \delta_l\mathbf{1}_N$ holds, where δ_l is the number of rows with Hamming weight l in Γ. Hence $\Sigma\mathbf{1}_N = \delta\mathbf{1}_N$ holds for some constant δ, which implies that $\Sigma^{-1}\mathbf{1}_N = \delta^{-1}\mathbf{1}_N$, since $\mathbf{1}_N = \Sigma^{-1}\Sigma\mathbf{1}_N = \delta\Sigma^{-1}\mathbf{1}_N$. This fact simplifies the covariance matrix (1.4) for the GLSE of $\boldsymbol{\beta}$ as follows:

$$\text{cov}(\hat{\boldsymbol{\beta}}) = (W'\Sigma^{-1}W)^{-1}. \tag{3.4}$$

For any given $m\times m$ permutation matrix P_1, if there exists an $N\times N$ permutation matrix P_2 such that $\Gamma P_1 = P_2\Gamma$, then Γ is said to be *invariant* with respect to any column permutations.

Lemma 3.1. *Let Γ be a linear* OA $(N,m,q,2)$ *which is invariant with respect to any column permutations and $A=(a_{ij})$ be an $N\times N$ matrix whose (i,j) th entry a_{ij} depends only on the Hamming distance between the i th and the j th rows of Γ. If there exists A^{-1}, then the entries of A^{-1} also depend only on the Hamming distance between the rows of Γ.*

Proof. We transform every row of Γ by adding a given vector γ_h. Since the set of rows of Γ is a linear subspace of $\text{GF}(q)^m$, there exists exactly one k such that $\gamma_i+\gamma_h=\gamma_k$ for each i. This means that the transformation induces a permutation on the subspace. Let P_1 be the $N\times N$ matrix representing such a permutation. Then P_1 exchanges any pair of rows γ_i and γ_k as long as $\gamma_i+\gamma_h=\gamma_k$ is satisfied. It is easy to show that $P_1AP_1'=A$, since $d(\gamma_i+\gamma_h,\gamma_j+\gamma_h)=d(\gamma_i,\gamma_j)$. Similarly, let P_2 be an $N\times N$ matrix induced by a column permutation of Γ. Then this type of permutation also satisfies $P_2AP_2'=A$.

Suppose that $d(\gamma_i,\gamma_j)=d(\gamma_k,\gamma_l)$, which implies that $a_{ij}=a_{kl}$. By making a product of the two types of permutations P_1 and P_2, we have a permutation P which exchanges the (i,j) th and the (k,l) th entries of A. It follows from $PA'P=A$ and $P'P=PP'=I_N$ that

$$A(PA^{-1}P') = (PAP')(PA^{-1}P') = PAA^{-1}P' = I_N$$

and whence $PA^{-1}P'=A^{-1}$. This means that the (i,j) th and the (k,l) th entries of A^{-1} are equal. Therefore A^{-1} is also a matrix whose (i,j) th entry depends only on the Hamming distance of the i th and the j th rows of Γ. □

Now, we clarify the relation between the entries of Γ and those of W. For convenience of the following description, we first define several notations. Let $\langle\Gamma\rangle$ be the set of rows of an OA $(N,m,q,2)$ Γ, let $\langle\Gamma\rangle^+=\langle\Gamma\rangle\setminus\{(0,\ldots,0)\}$ and define the equivalence relation $\sim$ by

$$\gamma_1\sim\gamma_2 \overset{def}{\Longleftrightarrow} \gamma_1=a\gamma_2 \ \text{ for an } a\in\text{GF}(q)\setminus\{0\},\ \ \gamma_1,\gamma_2\in\langle\Gamma\rangle^+.$$

If Γ is a linear orthogonal array, then $\langle\Gamma\rangle^+$ is divided into $(N-1)/(q-1)$ equivalence classes each of which is of size $q-1$. Let $\mathcal{S}$ be the set of the $(N-1)/(q-1)$ equivalence classes of $\langle\Gamma\rangle^+$. Since any vector in an identical class has the same Hamming weight, we denote it by $w(S)$ for a class $S\in\mathcal{S}$, instead of $w(s)$ for a vector

$s \in S$, unless otherwise specified. Label the elements of GF(q) as r_i for $i \in \mathbb{Z}_q$, and let $E = \{e_i : i \in \mathbb{Z}_q\}$ for the q-dimensional unit vectors e_i, *i.e.*, for the vectors e_i with the ith coordinate 1 and the rest all zero. We define a one-to-one mapping φ from GF(q) to E such that $e_i = \varphi(r_i)$. In the case where $r_i + r_j = r_k$ over GF(q), there exists a certain $q \times q$ permutation matrix T_{r_j} such that $e_i T_{r_j} = e_k$.

Lemma 3.2. *Let $F = \{f_i : f_i = e_i Q',\ e_i \in E,\ i \in$ GF$(q)\}$, where Q is the $(q-1) \times q$ matrix as defined in Section 1. If $r_i + r_j = r_k$ over* GF(q), *then $f_i Q T_{r_j} Q' = f_k$ holds.*

Proof. Since $Q'Q = I_q - (1/q)J_q$ and $J_q Q' = O$, we have $Q'QT_{r_j}Q' = T_{r_j}Q'$. On the other hand, from the fact that $e_i T_{r_j} = e_k$, if $r_i + r_j = r_k$ over GF(q), then $e_i T_{r_j} Q' = e_k Q'$ holds. Thus we have $f_i Q T_{r_j} Q' = f_k$. □

Lemma 3.3. *Let Γ be a linear* OA$(N, m, q, 2)$ *and $A = (a_{ij})$ be an $N \times N$ matrix, where a_{ij} depends only on the Hamming distance between the ith and jth rows of Γ, represented by $a_{ij} = \upsilon_{d(\gamma_i,\gamma_j)}$. Suppose that W is the $N \times m(q-1)$ matrix corresponding to Γ as defined by* (1.2)*. Then*

$$AW = W\left(\upsilon_0 I_{m(q-1)} + \sum_{S \in \mathcal{S}} \upsilon_{w(S)} \Lambda_S\right)$$

holds, where $\Lambda_S = diag\,(\underbrace{\lambda_1, \ldots, \lambda_1}_{q-1}; \cdots; \underbrace{\lambda_m, \ldots, \lambda_m}_{q-1})$ with

$$\lambda_j = \begin{cases} q-1 & \text{if } s_j = 0, \\ -1 & \text{if } s_j \neq 0 \end{cases}$$

for an arbitrary vector $s = (s_1, \ldots, s_m)$ chosen from an equivalence class $S \in \mathcal{S}$.

Proof. Define $N \times N$ matrices $P_{\gamma_j} = (p_{ik}^{(j)})$ by

$$p_{ik}^{(j)} = \begin{cases} 1 & \text{if } \gamma_i + \gamma_j = \gamma_k, \\ 0 & \text{otherwise.} \end{cases}$$

Since $\langle\Gamma\rangle$ is a linear subspace of GF$(q)^m$, for any pair of distinct vectors $\gamma_i, \gamma_k \in \langle\Gamma\rangle$, there exists a unique $\gamma_j \in \langle\Gamma\rangle$ satisfying $\gamma_i + \gamma_j = \gamma_k$ over GF$(q)^m$. Then P_{γ_j} can be regarded as a permutation matrix which transforms γ_i to γ_k. In this case, $d(\gamma_i, \gamma_k) = w(\gamma_j)$ and whence A can be represented as

$$A = \sum_{\gamma \in \langle\Gamma\rangle} \upsilon_{w(\gamma)} P_\gamma = \upsilon_0 I_N + \sum_{\gamma \in \langle\Gamma\rangle^+} \upsilon_{w(\gamma)} P_\gamma,$$

which immediately leads to

$$AW = \left(\upsilon_0 I_N + \sum_{\gamma \in \langle\Gamma\rangle^+} \upsilon_{w(\gamma)} P_\gamma\right) W. \tag{3.5}$$

Now, let

$$\mathcal{Q} = \begin{pmatrix} Q & & & \\ & Q & & \\ & & \ddots & \\ & & & Q \end{pmatrix}$$

be the $m(q-1) \times mq$ block-diagonal matrix. Then it follows from Lemma 3.2 that

$$P_s W = W \mathcal{Q}\, \mathrm{diag}(T_{s_1}, \ldots, T_{s_m}) \mathcal{Q}',$$

where T_{s_j} is the $q \times q$ permutation matrix as defined just before Lemma 3.2. Further let

$$\mathrm{diag}(T_{S(1)}, \ldots, T_{S(m)}) = \sum_{(s_1,\ldots,s_m)\in S} \mathrm{diag}(T_{s_1}, \ldots, T_{s_m})$$

be an $mq \times mq$ block-diagonal matrix. By noting that $\sum_{r_j \in \mathrm{GF}(q)} T_{r_j} = J_q$ and $T_0 = I_q$, we have

$$T_{S(j)} = \begin{cases} (q-1)I_q & \text{if } s_j = 0, \\ J_q - I_q & \text{if } s_j \neq 0. \end{cases}$$

It is easy to see that

$$Q T_{S(j)} Q' = \begin{cases} (q-1)I_{q-1} & \text{if } s_j = 0, \\ -I_{q-1} & \text{if } s_j \neq 0, \end{cases}$$

since $QQ' = I_{q-1}$ and $QJ_q = O$. Then we obtain

$$\begin{aligned} \sum_{\gamma \in \langle \Gamma \rangle^+} \upsilon_{w(\gamma)} P_\gamma W &= \sum_{S \in \mathcal{S}} \sum_{s \in S} \upsilon_{w(s)} P_s W \\ &= \sum_{S \in \mathcal{S}} \upsilon_{w(S)} W \mathcal{Q}\, \mathrm{diag}(T_{S(1)}, \ldots, T_{S(m)}) \mathcal{Q}' \qquad (3.6) \\ &= W \sum_{S \in \mathcal{S}} \upsilon_{w(S)} \Lambda_S. \end{aligned}$$

Thus (3.5) and (3.6) complete the proof. □

By virtue of Lemma 3.3, we can state the following.

Theorem 3.1. *Let* Γ *be a linear* OA $(N, m, q, 2)$ *and* W *be the* $N \times m(q-1)$ *matrix corresponding to* Γ. *Under covariance structure* (3.3),

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \frac{q}{N}\sigma^2 \left(I_{m(q-1)} + \sum_{S \in \mathcal{S}} \rho_{w(S)} \Lambda_S \right) \qquad (3.7)$$

holds for the OLSE of $\boldsymbol{\beta}$.

Theorem 3.2. *Besides the assumption of Theorem 3.1, assume that the (i,j) th entry of Σ^{-1} depends only on the Hamming distance between the i th and the j th rows of Γ for any i and j. Then,* (3.7) *also holds for the GLSE of $\boldsymbol{\beta}$.*

Proof. Since $\Sigma = \sigma^2 \sum_{l=0}^{m} \rho_l D_l$ for the adjacency matrices D_l $(l = 0, \ldots, m)$ of Γ, it follows from Lemma 3.3 that

$$\Sigma W = \sigma^2 W \left(I_{m(q-1)} + \sum_{S \in \mathcal{S}} \rho_{w(S)} \Lambda_S \right) \tag{3.8}$$

under covariance structure (3.3). Similarly, from the assumption of Σ^{-1} and Lemma 3.3, we have

$$\Sigma^{-1} W = \frac{1}{\sigma^2} W \left(\nu_0 I_{m(q-1)} + \sum_{S \in \mathcal{S}} \nu_{w(S)} \Lambda_S \right) \tag{3.9}$$

for some constants ν_l such that $\Sigma^{-1} = (1/\sigma^2) \sum_{l=0}^{m} \nu_l D_l$. With the second equation of (1.5), (3.4) becomes

$$\text{cov}(\hat{\boldsymbol{\beta}}) = \frac{q}{N} \sigma^2 \left(\nu_0 I_{m(q-1)} + \sum_{S \in \mathcal{S}} \nu_{w(S)} \Lambda_S \right)^{-1}.$$

Since $(\Sigma W)'(\Sigma^{-1} W) = W' \Sigma \Sigma^{-1} W = (N/q) I_{m(q-1)}$, it follows from (3.8) and (3.9) that

$$\left(\nu_0 I_{m(q-1)} + \sum_{S \in \mathcal{S}} \nu_{w(S)} \Lambda_S \right)^{-1} = I_{m(q-1)} + \sum_{S \in \mathcal{S}} \rho_{w(S)} \Lambda_S,$$

which completes the proof. □

Remark. If Γ is a linear orthogonal array which is invariant with respect to any column permutations, then Lemma 3.1 guarantees that Σ^{-1} satisfies the condition of Theorem 3.2.

4. Optimum orthogonal arrays for the nearest-neighbor correlation structure

In this section, we consider the covariance structure that Kiefer and Wynn dealt with as a "nearest-neighbor" (NN) correlation structure in [6]. This is a special case of (3.3) such that $\rho_0 = 1$, $\rho_1 = \rho$ and $\rho_2 = \cdots = \rho_m = 0$.

Covariance structure II:

$$\text{cov}(\boldsymbol{\epsilon}) = \Sigma = \sigma^2 (I_N + \rho D_1).$$

Let Γ be a linear OA $(q^n, m, q, 2)$, where $n \leq m$. Without loss of generality, we can assume that there are $t(q-1)$ rows with Hamming weight 1. Then the linear subspace $\langle\Gamma\rangle$ can be regarded as a direct sum of the t-dimensional linear space $\mathrm{GF}(q)^t = \langle\Gamma_1\rangle$ and an $(n-t)$-dimensional linear subspace $\langle\Gamma_2\rangle$ of the linear space $\mathrm{GF}(q)^{m-t}$, *i.e.*,

$$\langle\Gamma\rangle = \langle\Gamma_1\rangle \oplus \langle\Gamma_2\rangle. \tag{4.1}$$

Note that the minimum distance of Γ_2 is at least 2. Then the $N \times m(q-1)$ matrix W corresponding to Γ is written by

$$W = (\mathbf{1}_{q^{n-t}} \otimes W_1 | W_2 \otimes \mathbf{1}_{q^t}), \tag{4.2}$$

where $\otimes$ is the direct product, W_1 is the $q^t \times t(q-1)$ matrix corresponding to Γ_1 and W_2 is the $q^{n-t} \times (m-t)(q-1)$ matrix corresponding to Γ_2.

Let $\bar{D}_0 (= I_{q^t}), \bar{D}_1, \cdots, \bar{D}_t$ be the adjacency matrices of Γ_1 and let $V = \bar{D}_0 + \rho\bar{D}_1$. From (4.1) and (4.2), it is easy to see that both Γ and W can be divided into q^{n-t} blocks with q^t rows each such that the covariance of any two rows contained in distinct blocks is 0, *i.e.*,

$$\mathrm{cov}(\boldsymbol{\epsilon}) = \Sigma = \sigma^2 \mathrm{diag}(\underbrace{V, \ldots, V}_{q^{n-t}}). \tag{4.3}$$

Lemma 4.1. *Let Γ be a linear $\mathrm{OA}(q^n, m, q, 2)$ satisfying (4.1), W be the $q^n \times m(q-1)$ matrix of form (4.2) corresponding to Γ and $B = (b_{ij})$ be a $q^t \times q^t$ matrix, where b_{ij} depends only on the Hamming distance between the ith and the jth rows of Γ_1, represented by $b_{ij} = \upsilon_{d(\gamma_i,\gamma_j)}$. Then, for a $q^n \times q^n$ block-diagonal matrix $A = \mathrm{diag}(\underbrace{B, \ldots, B}_{q^{n-t}})$,*

$$AW = W \mathrm{diag}(\underbrace{\lambda, \ldots, \lambda}_{t(q-1)}, \underbrace{\kappa, \ldots, \kappa}_{(m-t)(q-1)})$$

holds, where

$$\lambda = \kappa - q \sum_{l=1}^{t} \upsilon_l \binom{t-1}{l-1} (q-1)^{l-1} \quad \textit{and} \quad \kappa = \sum_{l=0}^{t} \upsilon_l \binom{t}{l} (q-1)^l. \tag{4.4}$$

Proof. Assuming that W is divided into q^{n-t} submatrices of size $q^t \times m(q-1)$, say, $U^{(1)}, \ldots, U^{(q^{n-t})}$, AW can be written as

$$AW = \mathrm{diag}(B, \ldots, B) \begin{pmatrix} U^{(1)} \\ \vdots \\ U^{(q^{n-t})} \end{pmatrix} = \begin{pmatrix} BU^{(1)} \\ \vdots \\ BU^{(q^{n-t})} \end{pmatrix}.$$

Let $\omega_i = (\omega_{i1}, \ldots, \omega_{i,(m-t)(q-1)})$ be the i th row of W_2 . Then it follows from (4.2) that $U^{(i)} = (W_1 | \mathbf{1}_{q^t} \otimes \omega_i)$ and thus

$$BU^{(i)} = B(W_1|J)\mathrm{diag}(\underbrace{1,\ldots,1}_{t(q-1)}, \omega_{i1}, \ldots, \omega_{i,(m-t)(q-1)})$$

holds, where J is the $q^t \times (m-t)(q-1)$ all-one matrix. Then, from Lemma 3.3, we have

$$B(W_1|J) = (W_1|J)\left(\upsilon_0 I_{t(q-1)} + \sum_{\bar{S}\in\bar{\mathcal{S}}} \upsilon_{w(\bar{S})}\Lambda_{\bar{S}}\right),$$

where, without loss of generality, $\bar{S}$ is regarded as an equivalence class of $\langle\Gamma_1|O\rangle^+$, $\bar{\mathcal{S}}$ is the set of equivalence classes $\bar{S}$ and

$$\Lambda_{\bar{S}} = \mathrm{diag}(\underbrace{\lambda_1,\ldots,\lambda_1}_{q-1}; \cdots ; \underbrace{\lambda_m,\ldots,\lambda_m}_{q-1})$$

with

$$\lambda_j = \begin{cases} q-1 & \text{if } \bar{s}_j = 0, \\ -1 & \text{if } \bar{s}_j \neq 0 \end{cases}$$

for an arbitrary vector $\bar{s} = (\bar{s}_1, \ldots, \bar{s}_m)$ chosen from an equivalence class $\bar{S}$. Since $\langle\Gamma_1\rangle$ consists of all the t -dimensional vectors, the number of rows with Hamming weight l in $\langle\Gamma_1\rangle$ is $\binom{t}{l}$, and among them, there are exactly $\binom{t-1}{l-1}$ rows such that the i th coordinate is not zero for each i . By taking account of this fact and that $\bar{\mathcal{S}}$ consists of $q-1$ equivalence classes, we obtain

$$\upsilon_0 I_{t(q-1)} + \sum_{\bar{S}\in\bar{\mathcal{S}}} \upsilon_{w(\bar{S})}\Lambda_{\bar{S}} = \mathrm{diag}(\lambda,\ldots,\lambda,\kappa,\ldots,\kappa),$$

where λ and κ are calculated as (4.4). This means that

$$BU^{(i)} = U^{(i)}\mathrm{diag}(\lambda,\ldots,\lambda,\kappa,\ldots,\kappa)$$

and consequently $AW = W\mathrm{diag}(\lambda,\ldots,\lambda,\kappa,\ldots,\kappa)$. □

If Γ is a linear OA $(N,m,q,2)$, then the set of adjacency matrices $D_0(= I_N), D_1, \ldots, D_m$ becomes a so-called *Hamming scheme*. The following lemma is quite useful to evaluate λ and κ of (4.4) under Covariance structure II.

Lemma 4.2 (Mishima and Jimbo [8]). *For the Hamming scheme* $\{D_0, \ldots, D_m\}$ *on* GF $(q)^m$ *, let* $(D_0 + \rho D_1)^{-1} = \sum_{i=l}^m \upsilon_i D_i$ *. Then*

$$\sum_{i=k}^m \upsilon_i \binom{m-k}{i-k}(q-1)^{i-k} = \frac{k!(-\rho)^k}{\prod_{j=0}^k (1 + m\rho(q-1) - j\rho q)}$$

holds for $k = 0, \ldots, m$.

Theorem 4.1. *Under Covariance structure II, if* Γ *is a linear* OA $(N,m,q,2)$ *satisfying* (4.1)*, then* $C(\Gamma)$ *is a diagonal matrix with*

$$\frac{N}{\sigma^2 q}\cdot\frac{1}{1+t\rho(q-1)-\rho q} \quad \textit{and} \quad \frac{N}{\sigma^2 q}\cdot\frac{1}{1+t\rho(q-1)}$$

as $t(q-1)$ *and* $(m-t)(q-1)$ *of the* $m(q-1)$ *diagonal entries, respectively, where* t *is the number of vectors with Hamming weight* 1 *among* $\langle\Gamma\rangle$ *. Furthermore,* $C(\Gamma)$ *for the GLSE of* $\boldsymbol{\beta}$ *coincides with that for the OLSE of* $\boldsymbol{\beta}$ *.*

Proof. Let W be the $N\times m(q-1)$ matrix of form (4.2) corresponding to Γ . Then Σ is represented by (4.3) under Covariance structure II and we have

$$\Sigma^{-1}=\frac{1}{\sigma^2}\mathrm{diag}(V^{-1},\ldots,V^{-1}).$$

Since the (i,j) th entry of V depends only on the Hamming distance of the i th and the j th rows of Γ_1 and $\langle\Gamma_1\rangle$ is the linear space $\mathrm{GF}(q)^t$, it is invariant with respect to any column permutations. Hence, Lemma 3.1 guarantees that the (i,j) th entry of V^{-1} also depends only on the Hamming distance of the i th and the j th rows of Γ_1 . It follows from Lemmas 4.1 and 4.2 that

$$\Sigma W=\sigma^2 W\mathrm{diag}(\lambda,\ldots,\lambda,\kappa,\ldots,\kappa)$$

with

$$\lambda=\frac{1}{1+t\rho(q-1)-\rho q} \quad \text{and} \quad \kappa=\frac{1}{1+t\rho(q-1)}.$$

Again by Lemma 4.1 and $(\Sigma W)'(\Sigma^{-1}W)=(N/q)I_{m(q-1)}$, we have

$$\Sigma^{-1}W=\frac{1}{\sigma^2}W\mathrm{diag}(\lambda^{-1},\ldots,\lambda^{-1},\kappa^{-1},\ldots,\kappa^{-1}).$$

Then,

$$\begin{aligned}\mathrm{cov}(\hat{\boldsymbol{\beta}}_G)=(W'\Sigma^{-1}W)^{-1}&=\frac{q}{N}\sigma^2\mathrm{diag}(\lambda,\ldots,\lambda,\kappa,\ldots,\kappa)\\&=\frac{q^2}{N^2}W'\Sigma W=\mathrm{cov}(\hat{\boldsymbol{\beta}}_O)\end{aligned}$$

holds for the GLSE $\hat{\boldsymbol{\beta}}_G$ and the OLSE $\hat{\boldsymbol{\beta}}_O$ of $\boldsymbol{\beta}$. This completes the proof. □

Let Ξ_L be the set of linear OA $(N,m,q,2)$ by which the main effect $\boldsymbol{\beta}$ is estimable.

Theorem 4.2. *Let* $m\geq 3$ *and* Γ^* *be a linear* OA $(N,m,q,2)$ *with minimum distance at least* 2 *. Then, under Covariance structure II, for the OLSE and the GLSE of* $\boldsymbol{\beta}$ *,* Γ^* *is*

(i) *universally optimum relative to* Ξ_L *for* $0 \le \rho < 1 - q/(m(q-1))$ *, and*

(ii) *weakly universally optimum relative to* Ξ_L *for* $0 \le \rho (< 1)$ *.*

Proof. For an array $\Gamma \in \Xi_L$, assume that the number t of rows with Hamming weight 1 is at least 1.

(i) To prove (i) of the theorem, we need only show $\text{tr}(C(\Gamma)) < \text{tr}(C(\Gamma^*))$ for any $\Gamma \in \Xi_L$ when $0 \le \rho < 1 - q/(m(q-1))$ and $t \ge 1$. Let W be the $N \times m(q-1)$ matrix corresponding to Γ . From Theorem 4.1, we have

$$\text{tr}(C(\Gamma)) = \frac{N(q-1)}{q\sigma^2}\left(\frac{t}{1+t\rho(q-1)-\rho q} + \frac{m-t}{1+t\rho(q-1)}\right).$$

In particular, when $t = 0$,

$$\text{tr}(C(\Gamma^*)) = \frac{N(q-1)}{q\sigma^2} m.$$

Then we have only to prove that

$$\frac{t}{1+t\rho(q-1)-\rho q} + \frac{m-t}{1+t\rho(q-1)} < m$$

for $0 \le \rho < 1 - q/(m(q-1))$ and $t \ge 1$. When $t = 1$, it is true if $\rho < 1 - q/(m(q-1))$, and when $t \ge 2$, it is always true for any $\rho \ge 0$. Hence (i) is immediately established by Proposition 1.1.

(ii) In a manner similar to case (i), we have

$$\text{tr}(D(\Gamma)) = \frac{q(q-1)\sigma^2}{N}(t(1+t\rho(q-1)-\rho q) + (m-t)(1+t\rho(q-1)))$$

and

$$\text{tr}(D(\Gamma^*)) = \text{tr}(C(\Gamma^*)^{-1}) = \frac{q(q-1)\sigma^2}{N} m.$$

Then $\text{tr}(D(\Gamma)) - \text{tr}(D(\Gamma^*)) > 0$ holds for $t \ge 1$ and $0 \le \rho(< 1)$, since $m \ge 3$. Hence (ii) follows from Proposition 1.2. □

Remark. Theorem 2.1 proves Theorem 4.2(ii) under a more general covariance structure but only for the OLSE of $\boldsymbol{\beta}$.

5. Complete factorial designs

Last of all, we show covariance matrices of the OLSE and the GLSE of $\boldsymbol{\beta}$ under the covariance structures discussed in the previous sections when we adopt the complete factorial designs as Γ . In the case of complete factorial designs with m factors, $\langle \Gamma \rangle$ will be the linear space $\text{GF}(q)^m$. This means that Γ is invariant with respect to any column permutations. Hence by applying Lemmas 3.1, 4.1 and Theorem 3.2 with $t = m = n$, we have the following immediately.

Theorem 5.1. *For the complete factorial design with* m *factors, under covariance structure* (3.3)*, the covariance matrix for the OLSE of* $\boldsymbol{\beta}$ *is congruent with that for the GLSE, which is given by*

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \frac{q}{N}\sigma^2 \left(\sum_{l=0}^{m} \rho_l \binom{m}{l} (q-1)^l - q \sum_{l=1}^{m} \rho_l \binom{m-1}{l-1} (q-1)^{l-1} \right).$$

Corollary 5.2. *For the complete factorial design with* m *factors, under Covariance structure II, the covariance matrix for the OLSE of* $\boldsymbol{\beta}$ *is congruent with that for the GLSE, which is given by*

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}) = \frac{\sigma^2}{q^{m-1}}(1 + \rho m(q-1) - \rho q) I_{m(q-1)}.$$

Acknowledgment

The work was supported in part by Ministry of Education, Culture, Sports, Science and Technology of Japan, Grant-in-Aid for Young Scientists (B)1570023 to M. Mishima and by Japan Society for the Promotion of Science, Grant-in-Aid for Scientific Research (B)15340035 to M. Jimbo.

References

[1] R.J. Bennet, Spatial Time Series, Pion, 1979.

[2] T. Beth, D. Jungnickel and H. Lenz, Design Theory (2nd ed.), Vol.1 and 2, Cambridge University Press, Cambridge, 1999.

[3] C.S. Cheng, Construction of optimal balanced incomplete block designs for correlated observations. *Ann. Statist.*, **11** (1983), 240–246.

[4] M. Jimbo, Robustness of orthogonal arrays for correlated errors. Proceedings of the 2nd Japan-China Symposium on Statistics, Kyusyu University, 1986.

[5] J. Kiefer, Construction and optimality of generalized Youden designs II, in: J.N. Srivastava Ed., A Survey of Statistical Designs and Linear Models, North Holland, Amsterdam (1975), 333–353.

[6] J. Kiefer and H.P. Wynn, Optimum balanced incomplete block and Latin square designs for correlated observation, *Ann. Statist.*,**9** (1981), 737-757.

[7] F.J. MacWilliams and N. J. A. Sloane, The Theory of Error-Correcting Codes, North-Holland, New York, 1977.

[8] M. Mishima and M. Jimbo, A series of identities for the coefficients of inverse matrices on a Hamming scheme, *Discrete Math.*,**156** (1996), 285–290.

[9] M. Mishima, M. Jimbo and S. Shirakura, On the optimality of orthogonal arrays in case of correlated errors, *J. Statist. Plann. Inf.*,**88** (2000), 319–338.

[10] D. Raghavarao, Constructions and Combinatorial Problems in Design of Experiments, Wiley, New York, 1971.

[11] G.J.G. Upton and B. Fingleton, Spatial data analysis by example, Vol.1, Point pattern and quantitative data, John Wiley & Sons, Chichester, 1985.