Robustness of Complete Diallel cross designs with a Single Missing Observation

Yong-Man Kwon and Jang-Jae Lee[†]

Abstract

The reduction of efficiency of missing observations on complete diallel cross designs are examined. we studies robustness of optimal block designs for estimating general combining ability against loss of missing observations in diallel cross. A-efficiencies suggest that these designs are fairly robust. Simple g-inverses may be found for the information matrices of the line effects which allow evaluation of expressions for the variances of the differences between the pairs of line effects with missing observations. we numerically calculate the reduction of efficiency for estimating general combining ability against loss of missing observations in diallel cross.

Key words : Nested Balanced Incomplete Block Design, Complete Diallel Cross(CDC), Informat-ion Matrix, General Inverse, Robustness

1. Introduction

Diallel crosses as mating designs are commonly used to study the genetic properties of inbred lines in animal and plant breeding experiments. Suppose there are p inbred lines and let a cross between lines i and j be denoted by (i, j), $i < j = 1, 2..., p$. Let n_c denote the total number of distinct crosses in the experiment. Our interest lies in comparing the parents with respect to their general combining ability(gca) parameters. The complete diallel cross(CDC) involves all possible crosses among the p parental lines with $n_c = p(p\text{-}1)/2$. as discussed in detail by Griffing who referred to it as type mating design[1]. Gupta and Kageyama gave a method of constructing balanced block designs for CDCs using the nested balanced incomplete block(BIB) designs of Preece[2-3]. When a complete design is used, situations sometimes arise when one or two observations are lost for some reason not necessarily associated with either the response or the particular lines involved in the crosses. The robustness of block designs has been investigated extensively, for example Most, John, and others have considered the loss of one or more treat-

Department of computer science and statistics, Chosun University, Gwangju

ments[4-5]. The main measures of robustness were variance balance of the resulting design and average variance of pairwise treatment differences. Das and Kageyama also derived the information matrix for treatment effects when a block is missing from a Youden square design[6]. Precott and Mansson investigate the efficiency of pair-wise treatment comparisons in incomplete block experiments when a complete block is missed[7].

In this paper, we use a similar approach to investigate robustness of optimal block designs for estimating general combining ability effects against loss of observations in diallel cross. For single missing observation, there is only one distinct type of resulting design. The form of the information matrix and a suitable generalized inverse are derived theoretically. From this g-inverse it is possible to obtain expressions for the variances of all lines comparisons, and consequently the average of these variances. Theoretical results are difficult to obtain when there are two or more observations missing and even if these were derived they would be complicated and hard to interpret.

2. Optimal Complete Diallel Crosses Block Design

An incomplete block design for a diallel cross with constant block size k is taken to imply that the design

necessarily involves n_c treatments. This implication results in each cross being replicated k times, where r denotes the constant replication number for the design. It has also been suggested that a complete block design containing all the n_c crosses in each block is the most efficient design. Thus, we start with an incomplete block design D in p treatments, b blocks, constant replication number r, and constant block size $k' = 2k$. Then k pairs of parents are formed in each block which provide k crosses of the block. Obviously, the pairing should result in a complete diallel cross.

The pairing of the treatments within b blocks will result in a complete diallel cross if an irreducible incomplete block design, say D_n, with block size 2 is nested within D in the sense of Preece[3].

Furthermore, for the sake of optimality, design D is also taken to be a balanced incomplete block design. In some cases the above pairing may yield replications of a complete diallel cross. In that case will be $\lambda_n > 1$ replications of an irreducible balanced incomplete block design with block size 2. The method of construction is illustrated below in Example 1.

Example 1. Let D be a balanced incomplete block design with $p = 5$ treatments in $b = 5$ blocks of size $k' = 2k = 4$, replication number $r = 4$ of each treatment and number of concurrences $\lambda = 3$ of any two treatments, having blocks {1,2,3,4}, {1,2,3,5}, {1,2,4,5}, {1,3,4,5}, {2,3,4,5}.

These blocks are further divided into sub-blocks {(1, 2), (3, 4)}, {(1, 3), (2, 5)}, {(1, 5), (2, 4)}, {(1, 4), (3, 5)}, {(2, 3), (4, 5)}. For instance, the first block {(1, 2), (3, 4)} contains two crosses (1, 2), (3, 4) which, in fact, constitute two blocks of D_n. There is a total of 10 crosses in the experiment since each cross is replicated only once.

Series 1. Cyclically developing the initial block {1, 2t}, {2, 2t-1},..., {t, t+1}, mod (2t+1), yields an optimal design for $p = 2t + 1$ lines in $2t + 1$ blocks, there being t crosses in each block.

Another series of optimal designs can be obtained when p is even. For even p, take D as $p-1$ replications of the complete block in p treatments. Then there exists a resolvable balanced incomplete block design D_n with block size 2 and $\lambda_n = 1$ which is nested within D. Using the plan of D_n as given by Singh & Hinkelmann yields

the following series[8].

Series 2. For even p there exists the following optimal mating design having $p-1$ blocks with $p/2$ crosses in each block.

$$\left\{ (j+1, p-1-j+1), (l, p), \tfrac{1}{2}p-1 \right\} \ (l = 0, 1,..., p-2).$$

Where, addition is done mod($p-1$) and lines are coded as 0, 1,..., $p-1$, p).

The two series provide optimal designs for just one replication of the complete diallel cross. Their optimal design for $p = 5$ uses a balanced incomplete block design with parameters $n_c = 10$, $b = 15$, $r = 6$, $k' = 4$, $\lambda = 2$. In this design each of the 10 crosses is replicated 6 times while Example 1 provides an optimal design using only one replication.

3. A Single Missing Observation in Complete Diallel Crosses Block Design

3.1. CDC design when one observation is lost

To illustrate the process of examining the effects of losing an observation from a diallel cross design, we shall consider the situation where a single cross is lost from the CDC design produced from a nested BIB design constructed using the method described by Gupta and Kageyama[2]. Without loss of generality, assume that the first k lines occurred in the affected block, and that the cross between lines 1 and 2 is lost.

3.1.1 p is even

The information matrix for line effects may now be expressed as

$$C = \begin{bmatrix} (p-2)I_2 - \left\{ \dfrac{2(p-2)}{p} \right\} J_2 & -\left(1-\dfrac{4}{p}\right)J_{2, p-2} \\ -\left(1-\dfrac{4}{p}\right)J_{p-2, 2} & (p-2)I_{p-2} - \left\{ \dfrac{p^2-4p+8}{p(p-2)} \right\} J_{p-2} \end{bmatrix}$$

(1)

A particular g-inverse, C^-, may be found by adding $(1-4/p)J_p$ to C and inverting the resulting non-singular matrix to get

$$C^- = \begin{bmatrix} \left(\dfrac{1}{p-2}\right)I_2 - \left\{\dfrac{1}{(p-2)(p-4)}\right\}J_2 & 0_{2,p-2} \\ 0_{p-2,2} & \left(\dfrac{1}{p-2}\right)I_{p-2} - \left\{\dfrac{2}{(p-2)^2(p-4)}\right\}J_{p-2} \end{bmatrix}$$

$$(2)$$

There are several situations to consider when determining the variances of the differences between the pairs of line effects. Different variances are given for pairs of lines consisting of: (1) the two lines in the missing cross, or the two lines not in the missing cross (2) one of the lines in the missing cross and another line in the affected block. We denote the variances of the differences between the line effects for these three situations by T_1 and T_2 respectively.

In case (1), for the two lines in the missing cross, for the two lines not in the missing cross, the variance of the difference is unchanged at

$$T_1 = \frac{2\sigma^2}{p-2} \tag{3}$$

For the worst case, case (2), when comparing one line of the missing cross and a line not in the missing cross, the variance is increased to

$$T_2 = \frac{(2p^2 - 11p + 16)\sigma^2}{(p-2)^2(p-4)} \tag{4}$$

The CDC design was constructed by Gupta and Kageyama from a nested BIB design with $p = 6$ lines[2]. If any individual cross is lost from one of the blocks of this CDC design, the two variances for the differences between pairs of line effects are $T_1 = 0.5\ \sigma^2$, $T_2 = 0.6875\ \sigma^2$. The loss of efficiency is greatest, at just over 15%, for the comparison of one of the lines in the missing cross and another line from the affected block.

To illustrate the process of examining the effects of losing an observation from a diallel cross design, we shall consider the situation where a single cross is lost from the CDC design produced from a nested BIB design constructed using the method described by Gupta and Kageyama[2]. Without loss of generality, assume that the first k lines occurred in the affected block, and that the cross between lines 1 and 2 is lost

3.1.1 p is odd
The information matrix for line effects may now be expressed as

$$\begin{bmatrix} \left\{\dfrac{p(p-3)}{p-1}\right\}I_2 - \left\{\dfrac{2(p-3)}{p-1}\right\}J_2 & -\left(\dfrac{p-5}{p-1}\right)J_{2,p-3} & -\left(\dfrac{p-3}{p-1}\right)J_{2,1} \\ -\left(\dfrac{p-5}{p-1}\right)J_{p-3,2} & \left\{\dfrac{p(p-3)}{p-1}\right\}I_{p-3} - \left\{\dfrac{2(p-3)}{p-1}\right\}J_{2p-3} & -\left(\dfrac{p-3}{p-1}\right)J_{p-3,1} \\ -\left(\dfrac{p-3}{p-1}\right)J_{1,2} & -\left(\dfrac{p-3}{p-1}\right)J_{1,p-3} & p-3 \end{bmatrix}$$

$$(5)$$

A particular g-inverse, C^-, may be found by adding $\{(p-3)/(p-1)\}J_p$ to C and inverting the resulting non-singular matrix to get

$$C^- = \begin{bmatrix} a_1^{-1}I_2 + a_4J_2 & a_3J_{2,p-3} & 0_{2,1} \\ a_3J_{p-3,2} & a_1^{-1}I_2 + a_4J_2 & 0_{1,p-3,1} \\ 0_{1,2} & 0_{1,p-3} & \dfrac{p-1}{p(p-3)} \end{bmatrix} \tag{6}$$

Where,

$$a_3 = \frac{p(p-3)}{p-1}, \quad a_2 = \frac{8(p-1)^2}{p^3(p-2)(p-3)^4 - 8p(p-1)(p-3)^{2'}}$$

$$a_3 = \frac{2p\{(p-3)a_1a_2 + 1\}}{p(p-2)(p-3)^2a_1}, \quad a_4 = \frac{(p-1)(2a_3+1)}{p(p-2)(p-3)}$$

Different variances are given for pairs of lines consisting of: (1) the two lines in the missing cross, or two lines in the affected block but not part of the missing cross, or two lines not in the affected block; (2) one of the lines in the missing cross and a line in the affected block; (3) a line in the affected block (but not in the missing cross) and a line not in the block and (4) a line in the missing cross and a line not in the affected block. We denote the variances of the differences between the line effects for these four situations by T_1, T_2, T_3 and T_4, respectively.

In case (1), for the two lines in the missing cross, or for two other lines in the affected block, or for two lines not in the affected block, the variance of the difference is unchanged at

$$T_1 = \frac{2(p-1)}{p(p-3)}\sigma^2 \tag{7}$$

In case (2), when comparing one line of the missing cross and another line in the affected block, the variance is increased to

$$T_2 = (2a_1^{-1} + a_2 + a_4 - 2a_3)\sigma^2 \qquad (8)$$

The variance of a comparison between a line from the affected block, but not in the lost cross, and a line not in the affected block in case (3) is

$$T_3 = (2a_1^{-1} + a_2)\sigma^2 \qquad (9)$$

Finally For the worst case (4), the variance of a comparison between one of the lines in the missing cross and a line which is not in the affected block, increases to

$$T_4 = (2a_1^{-1} + a_4)\sigma^2 \qquad (10)$$

If any individual cross is lost from one of the blocks of this CDC design with $p = 7$ lines, the four variances for the differences between pairs of line effects are $T_1 = 0.4286\sigma^2$, $T_2 = 0.4383\sigma^2$, $T_3 = 0.4318\sigma^2$, $T_4 = 0.4805\sigma^2$.

4. Final Remarks

We have seen from the theoretical results for a single cross missing from a CDC design and from the numerical results for missing crosses that these designs are reasonably robust overall. However, consideration of the variances of individual line differences indicates that in some cases these comparisons may be seriously affected by the missing data. For even p in some of the examples with a single cross missing, reductions in average efficiencies were as much as 3-15%, and reductions in maximum efficiencies were as much as 6-27%, the worst cases being the loss of a cross between one of the lines in the missing cross and another line in the affected block. For odd p in some of the examples with a single cross missing, reductions in average efficiencies were as much as 2-4%, and reductions in maximum

efficiencies were as much as 4-11%, the worst cases being the loss of the cross between one of the lines in the missing cross and a line which is not in the affected block.

Acknowledgments

References

[1] B. Griffing, "A generalized treatment of the use of diallel crosses in quantitative inheritance", Heredity, Vol. 10, pp. 31-50, 1956.

[2] S. Gupta and S. Kageyama, "Optimal complete diallel crosses", Biometrika, Vol. 81, No. 2, pp. 420-424, 1994.

[3] D.A. Preece, "Nested balanced incomplete block designs", Biometrika, Vol. 54, pp. 479-486, 1967.

[4] B.M. Most, and B.M., "Resistance of balanced incomplete block designs", Ann. Statist, Vol. 3, No. 5, pp. 1149-1162, 1975.

[5] P.W.M. John, "Robustness of balanced incomplete block design", Ann. Statist, Vol. 4, No. 5, pp. 960-962, 1976.

[6] A. Das, and S. Kageyama, "Robustness of BIB and extended BIB designs against the unavailability of any number of observations in a block", Comput. Statist. Data Anal, Vol. 14, No. 3, pp. 343-358, 1992.

[7] R.A. Mansson and P. Prescott, "Missing observations in Youden square designs", Comput. Statist. Data Anal, Vol. 40, pp. 329-338, 2002.

[8] M. Singh and K. Hinkelmann, "On generation of efficient partial diallel crosses plans", Biometrika, Vol. 32, pp. 177-187, 1990.