

Series of Incomplete Row-Column Designs with Two Units per Cell

Anindita Datta¹, Seema Jaggi¹, Cini Varghese¹ and Eldho Varghese¹

Abstract

Here, two series of incomplete row-column designs with two units per cell have been developed that are structurally complete, i.e. all the cells corresponding to the intersection of row and column receive two distinct treatments. Properties of these classes of designs have been studied and the methods result in designs in which the elementary contrasts of treatment effects are estimated with same variance.

1. Introduction

Row-column designs are used for controlling heterogeneity in the experimental material in two directions. Most of the row-column designs developed in the literature have one unit corresponding to the intersection of row and column. Row-column designs with more than one unit per cell are used when the number of treatments is substantially large with limited number of replicates. For example, (Bailey and Monod, 2001), to conduct an experiment for comparing 4 treatments using 4 plants with leaves at 2 different heights, the following row-column design with complete rows and columns having two units per cell can be used:

	Plants			
Leaf Height	1 2	2 3	3 4	4 1
	3 4	4 1	1 2	2 3

These designs are termed as semi-Latin squares in the literature. An $(n \times n)/k$ semi-Latin square is an arrangement of nk symbols (treatments) in an $(n \times n)$ square array such that

¹ ICAR-Indian Agricultural Statistics Research Institute, Library Avenue, New Delhi-110 012, India. Tel.: +91-11-25847284, E-mail address: seema.iasri@outlook.com, seema@iasri.res.in

each row-column intersection contains k symbols and each symbol occurs once in each row and each column. Trojan squares are a special class of semi-Latin squares based on sets of mutually orthogonal superimposed Latin squares and have been shown to be maximally efficient for pair-wise treatment comparisons in the plots-within-blocks stratum (Bailey, 1992). Following is an example of a Trojan square design of size $(4 \times 4)/2$ for 8 treatments constructed by superimposing two mutually orthogonal Latin squares of size 4, one with 1, 2, 3 and 4 treatments and the other with 5, 6, 7 and 8 treatments.

	Columns			
Rows	1 5	2 6	3 7	4 8
	2 7	1 8	4 5	3 6
	3 8	4 7	1 6	2 5
	4 6	3 5	2 8	1 7

This arrangement could be extended for 12 treatments of size $(4 \times 4)/3$ by superimposing the third orthogonal Latin square of size 4 but no further Trojan extension is possible, there being only three mutually orthogonal Latin squares of size 4.

Complete Trojan squares of size $(n \times n)/k$ have n^2 blocks of size k and require n replicates of nk treatments. Sometimes, design or cost constraints make complete Trojan squares impossible and then Incomplete Trojan squares of size $[(n-1) \times n]/k$ or of size $[n \times (n-1)]/k$ can be useful. Such incomplete Trojan squares can be constructed by omitting any complete row or any complete column from any Trojan design of size $(n \times n)/k$. Trojan squares were first discussed by Harshbarger and Davis (1952) but then it was named as Latinized Near Balanced Rectangular Lattices having $k = n-1$. Later, Darby and Gilbert (1958) discussed the general case for $k < n$ and introduced the name Trojan square designs where $k > 2$. However, all designs of the Latinized Rectangular Lattice type are now commonly described as Trojan squares for any $1 < k < n$. Williams (1986) generalized the notion and called semi-Latin squares as Latinized incomplete-block designs. Andersen and Hilton (1980) called semi-Latin squares as $(1, 1, k)$ Latin rectangles.

Preece and Freeman (1983) discussed the combinatorial properties of semi-Latin squares and related designs. Bailey (1988) discussed further construction for a range of semi-Latin and Trojan square designs. Bailey (1992) gave methods of constructing a range of semi-Latin and Trojan square designs, studied their efficiencies and showed that the Trojan squares are the optimal choice of semi-Latin squares for pair-wise comparisons of treatment means. These are particularly suitable for crop research experiments either in field or in the glasshouse. Trojan squares are normally the best choice of semi-Latin squares for crop research (Edmondson, 1998). Bedford and Whitaker (2001) have given several methods of construction of semi-Latin squares.

Dharmalingam (2002) gave an application of Trojan square designs and used it to obtain partial triallel crosses. Edmondson (2002) constructed generalized incomplete Trojan square designs, denoted by $(m \times n)/k$ where m denotes the number of replicates of nk treatments, based on a set of k cyclic generators.

There existed three optimal $(4 \times 4)/4$ semi-Latin squares (Bailey and Chigbu, 1997) for sixteen treatments in blocks of size four. Since these squares do not have the same concurrences, there was a need for distinguishing one square from the others and determining the most preferred square in a given context. Chigbu (2003) obtained the best of the three optimal $(4 \times 4)/4$ semi-Latin squares by finding and comparing the variances of elementary contrasts of treatments for the squares.

Jaggi et al. (2010) defined Generalized Incomplete Trojan-Type Designs and developed method of constructing these designs. Varghese and Jaggi (2011) obtained generalized row-column designs and showed their application in obtaining mating plans. Datta et al. (2014) obtained some methods of constructing row-column designs with multiple units per cell that are structurally incomplete i.e. corresponding to the intersection of any row and column, there is at least one cell which does not contain any treatment. Datta et al. (2015) developed row-column designs with multiple units per cell with equal/ unequal cell sizes. Jaggi et al. (2016) obtained another series of generalized incomplete Trojan-type designs for number of treatments $v = sm + 1$.

Most of the methods available in the literature are for complete rows and complete columns. Here, two methods of constructing row-column designs with two units per cell in incomplete rows or columns are obtained that are balanced for estimating elementary contrasts of treatment effects.

2. Experimental Setup and Information Matrix

We consider a row-column design with v treatments arranged in m rows, n columns and in each row-column intersection (i.e. cells) there are k units or plots resulting in total mnk experimental units or observations. The following four-way classified model with treatments, rows, columns and cells as the four classifications, is considered:

$$\mathbf{Y} = \mathbf{X}_1\theta_1 + \mathbf{X}_2\theta_2 + \mathbf{e},$$

where

$$\mathbf{X}_1 = [\Delta'] , \quad \mathbf{X}_2 = [\mathbf{1} \quad \mathbf{D}'_1 \quad \mathbf{D}'_2 \quad \mathbf{D}'_3]'$$

$\theta_1 = (\boldsymbol{\tau})$ is the vector of parameters of interest and $\theta_2 = (\mu \quad \boldsymbol{\beta} \quad \boldsymbol{\gamma} \quad \boldsymbol{\eta})$ is the vector of nuisance parameters. \mathbf{Y} is a $mnk \times 1$ vector of observations, μ is the grand mean, $\mathbf{1}$ is the $mnk \times 1$ vector of ones, Δ' is $mnk \times v$ matrix of observations versus treatments, $\boldsymbol{\tau}$ is a $v \times 1$ vector of treatment effects, \mathbf{D}'_1 is $mnk \times m$ matrix of observations versus rows, $\boldsymbol{\beta}$ is $m \times 1$ vector of row effects, \mathbf{D}'_2 is $mnk \times n$ matrix of observations versus columns, $\boldsymbol{\gamma}$ is $n \times 1$ vector of column effects, \mathbf{D}'_3 is $mnk \times mn$ matrix of observations versus cells, $\boldsymbol{\eta}$ is $mn \times 1$ vector of cell effects and \mathbf{e} is $mnk \times 1$ vector of random errors with $E(\mathbf{e}) = 0$ and $D(\mathbf{e}) = \sigma^2 \mathbf{I}$.

The information matrix of row-column design with multiple units per cell for treatment effects is obtained as

$$\begin{aligned} \mathbf{C} = \mathbf{R}_\tau - \{ & (\mathbf{N}_1\mathbf{K}_{11}\mathbf{N}'_1 + \mathbf{N}_2\mathbf{K}_{21}\mathbf{N}'_1 + \mathbf{N}_3\mathbf{K}_{31}\mathbf{N}'_1 + \mathbf{N}_1\mathbf{K}_{12}\mathbf{N}'_2 + \mathbf{N}_2\mathbf{K}_{22}\mathbf{N}'_2) \\ & + (\mathbf{N}_3\mathbf{K}_{23}\mathbf{N}'_2 + \mathbf{N}_1\mathbf{K}_{13}\mathbf{N}'_3 + \mathbf{N}_2\mathbf{K}_{23}\mathbf{N}'_3 + \mathbf{N}_3\mathbf{K}_{33}\mathbf{N}'_3) \} \end{aligned} \quad \dots(2.1)$$

where

$$\mathbf{K}_{11} = \mathbf{K}_\beta^- + \left[\mathbf{K}_\beta^- (\mathbf{M}_1\mathbf{K}_{22}\mathbf{M}'_1 + \mathbf{M}_2\mathbf{K}_{23}\mathbf{M}'_1 + \mathbf{M}_1\mathbf{K}_{23}\mathbf{M}'_1 + \mathbf{M}_2\mathbf{K}_{33}\mathbf{M}'_2) \mathbf{K}_\beta^- \right]$$

$$\mathbf{K}_{12} = -\mathbf{K}_\beta^- (\mathbf{M}_1\mathbf{K}_{22} + \mathbf{M}_2\mathbf{K}_{23})$$

$$\mathbf{K}_{13} = \mathbf{K}_\beta^- (\mathbf{M}_1\mathbf{K}_{23} + \mathbf{M}_2\mathbf{K}_{33})$$

$$\mathbf{K}_{22} = \mathbf{K}_\gamma^- + \mathbf{K}_\gamma^- \mathbf{M}_3 (\mathbf{K}_\eta - \mathbf{M}'_3 \mathbf{K}_\gamma^- \mathbf{M}_3) \mathbf{K}_\gamma^-$$

$$\mathbf{K}_{23} = -\mathbf{K}_\gamma^- \mathbf{M}_3 (\mathbf{K}_\eta - \mathbf{M}'_3 \mathbf{K}_\gamma^- \mathbf{M}_3) \mathbf{K}_\gamma^-$$

$$\mathbf{K}_{33} = (\mathbf{K}_\eta - \mathbf{M}'_3 \mathbf{K}_\gamma^- \mathbf{M}_3) \mathbf{K}_\gamma^-$$

\mathbf{R}_t is the diagonal matrix of replication of treatments, \mathbf{K}_β is the diagonal matrix of row-sizes, \mathbf{K}_γ is the diagonal matrix of column-sizes and \mathbf{K}_η is the diagonal matrix of cell-sizes. \mathbf{N}_1 is the incidence matrix of treatments versus rows, \mathbf{N}_2 is the incidence matrix of treatments versus columns, \mathbf{N}_3 is the incidence matrix of treatments versus cells, \mathbf{M}_1 is the incidence matrix of rows versus columns, \mathbf{M}_2 is the incidence matrix of rows versus cells and \mathbf{M}_3 is the incidence matrix of columns versus cells.

The $v \times v$ matrix \mathbf{C} is symmetric, non-negative definite with zero row and column sums.

3. Methods of Construction

We present here two methods of constructing row-column designs with two units per cell. The first method is for odd number of treatments and the second is for even number.

Method 3.1: For $v = 2t + 1$ ($t > 1$), obtain the following initial column having two units per cell:

1	$2t + 1$
2	$2t$
3	$2t - 1$
.	.
.	.
.	.
t	$2t - (t - 2)$

Develop $2t$ more columns horizontally from the initial column by adding $1, 2, \dots, 2t$ consecutively reducing mod $(2t+1)$. The resulting design is a row-column design with two units per cell and with $m = t$ rows of size $2(2t+1)$, $n = (2t+1)$ columns of size $2t$, $k = 2$ and $r = 2t$ replications. The design is complete row-wise and column-wise it is incomplete. The information matrix of the design for treatment effects is obtained from (2.1) as

$$\mathbf{C} = (t + 0.5)\mathbf{I} - 0.5\mathbf{J}.$$

It is seen that all the elementary treatment contrasts are estimated with same variance. This method would give designs for all odd number of treatments.

Example 3.1.1: Let $t = 3$, so $v = 7$. The contents of the initial column are obtained as follows:

1	7
2	6
3	5

Developing this column by adding 1,2,...,6 reducing mod 7 would result in the following row-column design in three rows of size 14, 7 columns of size 6 with 2 units per cell and replication of each treatment being 6:

	Columns													
Rows	1	7	2	1	3	2	4	3	5	4	6	5	7	6
	2	6	3	7	4	1	5	2	6	3	7	4	1	5
	3	5	4	6	5	7	6	1	7	2	1	3	2	4

The information matrix for estimating treatment effects is $\mathbf{C} = 3.5\mathbf{I} - 0.5\mathbf{J}$.

Example 3.1.2: For $t = 4$, the contents of the initial column for $v = 9$ are obtained as follows:

1	9
2	8
3	7
4	6

The row-column design in 4 rows of size 18, 9 columns of size 8 and 8 replications is constructed as given.

	Columns																	
Rows	1	9	2	1	3	2	4	3	5	4	6	5	7	6	8	7	9	8
	2	8	3	9	4	1	5	2	6	3	7	4	8	5	9	6	1	7
	3	7	4	8	5	9	6	1	7	2	8	3	9	4	1	5	2	6
	4	6	5	7	6	8	7	9	8	1	9	2	1	3	2	4	3	5

Here, the design is complete row-wise with each treatment occurring twice and column-wise it is incomplete.

Method 3.2: For v even, obtain the following initial column with 2 units per cell:

1	v
v	2
2	$v-1$
$v-1$	3
.	.
.	.
.	.
$v - (\frac{v}{2} - 2)$	$v - \frac{v}{2}$
$\frac{v}{2}$	$\frac{v}{2} + 1$

Develop $\frac{v}{2} - 1$ more columns horizontally from the initial column by adding $1, 2, \dots, (\frac{v}{2} - 1)$ consecutively reducing mod v . The resulting design is a row-column design with 2 units per cell in incomplete rows and complete columns. The parameters of the design are v , $m = (v-1)$ rows of size v , $n = \frac{v}{2}$ columns of size $2(v-1)$, $k = 2$ and $r = (v-1)$. From (2.1), the information matrix for treatment effects is obtained as

$$C = \frac{v}{2} \mathbf{I} - 0.5 \mathbf{J}.$$

Thus, all the elementary contrasts of treatment effects are estimated with same variance.

Example 3.2.1: For $v = 8$, following is a row-column design with cells containing 2 units in 7 rows of size 8 each and 4 columns of size 14 each:

	Columns							
Rows	1	8	2	1	3	2	4	3
	8	2	1	3	2	4	3	5
	2	7	3	8	4	1	5	2
	7	3	8	4	1	5	2	6
	3	6	4	7	5	8	6	1
	6	4	7	5	8	6	1	7
	4	5	5	6	6	7	7	8

The canonical efficiency factor of the above two series of designs is obtained as

$$E = \frac{H}{r}, \text{ where } H = \left(\frac{1}{v-1} \sum_{i=1}^{v-1} \lambda_i^{-1} \right)^{-1},$$

λ_i are the eigen-values of \mathbf{C} - matrix of the designs obtained and r is the number of replications of the treatments. It is assumed that σ^2 is same for the developed design and the orthogonal design to which it is compared. The canonical efficiency factor of the developed designs was worked out and was found to be fairly good.

Acknowledgements

The authors are grateful to the Editor and the reviewers for the constructive comments that have led to considerable improvement in the paper.

References

- [1] Andersen, L.D. and Hilton, A.J.W. (1980): Generalized Latin rectangles I: Constructions and decomposition, *Discrete Maths.*, **31**, 125-152.
- [2] Bailey, R.A. (1988): Semi Latin squares, *J. Statist. Plan. Inf.*, **18**, 299-312.
- [3] Bailey, R.A. (1992): Efficient semi-Latin squares, *Statistica Sinica*, **2**, 413-437.
- [4] Bailey, R.A. and Chigbu, P.E. (1997): Enumeration of semi-Latin squares, *Discrete Maths.*, **167/168**, 73-84.
- [5] Bailey, R.A. and Monod, H. (2001): Efficient semi-Latin rectangles: Designs for plant disease experiments, *Scand. J. Statist.*, **28**, 257-270.

-
- [6] Bedford, D. and Whitaker, R.M. (2001): A new construction for efficient semi-Latin squares, *J. Statist. Plan. Inf.*, **98**, 287-292.
- [7] Chigbu, P.E. (2003): The “best” of the three optimal (4 x 4)/4 semi-Latin squares, *Sankhya: The Indian Journal of Statistics*, **65**(3), 641–648.
- [8] Darby, L.A. and Gilbert, N. (1958): The Trojan Square, *Euphytica*, **7**, 183-188.
- [9] Datta, A., Jaggi, S., Varghese, C. and Varghese, E. (2014): Structurally incomplete row-column designs with multiple units per cell. *Statistics and Applications*, **12**, (1&2): 71-79.
- [10] Datta, A., Jaggi, S., Varghese, C. and Varghese, E. (2015): Some series of row-column designs with multiple units per cell. *Calcutta Statistical Association Bulletin*, **67**, (265-266), 89-99.
- [11] Dharmalingam, M. (2002): Construction of partial triallel crosses based on Trojan square design, *J. App. Statist.*, **29**(5), 675-702.
- [12] Edmondson, R.N. (1998): Trojan square and incomplete Trojan square design for crop research, *J. Agric. Sci.* **131**, 135-142.
- [13] Edmondson, R.N. (2002): Generalized incomplete Trojan designs, *Biometrika*, **89**(4), 877-891.
- [14] Harshbarger, B. and Davis, L.L. (1952): Latinized rectangular lattices, *Biometrics*, **8**, 73-84.
- [14] Jaggi, Seema, Varghese, Cini, Varghese, Eldho and Sharma, V.K. (2010): Generalized incomplete Trojan-type designs, *Statistics and Probability Letters*, **80**, 706-710.
- [15] Jaggi, Seema, Varghese, Cini, and Varghese Eldho (2016): A series of generalized incomplete Trojan-type designs. *Journal of Combinatorics, Information and System Sciences: American Journal*, **40**(1-4), 53-60.
- [16] Preece, D.A. and Freeman, G.H. (1983): Semi-Latin squares and related designs, *J. R. Statist. Soc. B* **45**, 267-277.
- [17] Varghese, Cini and Jaggi, Seema (2011): Mating designs using generalized incomplete Trojan-type designs, *Journal of Statistics and Applications*, **6**(3-4), 85-93.
- [18] Williams, E.R. (1986): Row and column designs with contiguous replicates, *Australian J. Statist.*, **28**, 154-163.