## Bronisław Ceranka

Poznań University of Life Sciences, Faculty of Agronomy and Bioengineering,
Department of Mathematical and Statistical Methods, bronicer@up.poznan.pl


## Małgorzata Graczyk

Poznań University of Life Sciences, Faculty of Agronomy and Bioengineering,
Department of Mathematical and Statistical Methods, magra@up.poznan.pl

# Highly D-efficient Weighing Design and Its Construction

**Abstract:** In this paper, some aspects of design optimality on the basis of spring balance weighing designs are considered. The properties of D-optimal and D-efficiency designs are studied. The necessary and sufficient conditions determining the mentioned designs and some new construction methods are introduced. The methods of determining designs that have the required properties are based on a set of incidence matrices of balanced incomplete block designs and group divisible designs.

**Keywords:** balanced incomplete block design, efficient design, group divisible design, optimal design, spring balance weighing design

**JEL:** C02, C18, C90

# 1. Introduction

Amidst the complexity of the problems dealt with while conducting any research, the manner of experiment planning plays a significant role. In order to construct an experimental plan, we have to consider the information on the experimental material, the environmental conditions, and the optimal statistical properties. It is necessary to maintain the balance between these expectations to conduct the most effective experiment. Both the increase of effectiveness and the reduction of error variance depend on making appropriate assumptions connected with experimental errors and the design. Many reasons influence the value of the variance: some inaccuracies in the measuring of the examined characteristic or non-homogeneity of experimental designs. The above-mentioned attributes result in significant variability of observations. It can impact deduction on the grounds of obtaining measurements. Hence, we need to plan and carry out an experiment in such a way so as to minimise the influence of these elements with regard to the aim and manner of performing the experiment. An important part of research is such planning of experiments so as to obtain the best estimators of unknown measurements of objects. Optimal designs permit to obtain the value of unknown parameters with minimal variance. The design which is not optimal usually requires a larger number of measurements for assessment of the same parameters, frequently with a lesser degree of appraisal precision. Practically, optimal designs allow the reduction of experimental costs. Optimality depends on a statistical model and is connected with the variance matrix of estimators. The selection of an appropriate model is related to experimental conditions, whereas the choice of the optimality criterion depends upon a person performing the experiment. In many experiments, it is expected that the mean variance of the comparisons should be as small as possible. For that reason, D-optimal designs are taken into consideration. These are designs in which the product of the variances attains the smallest value.

In the presented paper, the issues concerning the above-mentioned problems are considered from the point of view of spring balance weighing designs. Any spring balance weighing design is defined as a design in which we determine unknown measurements of $p$ objects in $n$ measurement operations according to the model $\mathbf{y} = \mathbf{Xw} + \mathbf{e}$, where:
1) $\mathbf{y}$ is a $n \times 1$ random vector of the recorded results of measurements,
2) $\mathbf{X} = (x_{ij}) \in \mathbf{\Phi}_{n \times p}(0, 1)$, $\mathbf{\Phi}_{n \times p}(0, 1)$ denotes the class of matrices with elements $x_{ij} = 1$ or $0$, $i = 1, 2, \ldots, n, j = 1, 2, \ldots, p$,
3) $\mathbf{w}$ is a $p \times 1$ vector of unknown measurements of objects,
4) $\mathbf{e}$ is an $n \times 1$ random vector of errors, $\mathrm{E}(\mathbf{e}) = 0_n$ and $\mathrm{E}(\mathbf{ee'}) = \sigma^2\mathbf{G}$, $\mathbf{G}$ is known as a positive definite matrix.

The possibility of using the proposed methodology of measuring economic phenomena is presented in Banerjee (1975) and Ceranka and Graczyk (2014c).

In order to determine unknown measurements of objects, we use normal equations $\mathbf{X'G^{-1}X\hat{w}} = \mathbf{X'G^{-1}y}$, where $\hat{\mathbf{w}}$ is the vector estimated by the least squares method. If X is of full column rank, then the least squares estimator of w is given by $\hat{\mathbf{w}} = (\mathbf{X'G^{-1}X})^{-1}\,\mathbf{X'G^{-1}y}$ and the covariance matrix of $\hat{\mathbf{w}}$ is equal to $V(\hat{\mathbf{w}}) = = \sigma^2\,(\mathbf{X'G^{-1}X})^{-1}$.

The problems considered in such research are related to determining the estimator of the vector of unknown measurements of objects having certain properties in the class. Besides, the optimality criteria are the functions of the matrix $(\mathbf{X'G^{-1}X})^{-1}$. Therefore, the ways of determining and assessing have to be adapted to the classes in which we look for solutions. Hence, the obtained solutions have to belong to the set {0, 1}.

The problems of optimality of spring balance weighing designs under different assumptions concerning the measurement error are discussed in the literature, see Masaro and Wong (2008a; 2008b), Ceranka and Graczyk (2014a; 2014b; 2014c). Unfortunately, the properties of experimental designs and the relations between their parameters are the reasons why we are not able to determine the optimal design in any class $\Phi_{n \times p}(0, 1)$. In such a situation, we indicate the design for which the obtained assessment is as close to the best design as possible.

The aim of the study presented here is to provide the best design in the class $\mathbf{\Phi}_{n \times p}(0, 1)$, in which a regular D-optimal design does not exist. We define the best design as the design of the largest possible efficiency factor. We indicate such a design and provide construction methods.

## 2. Efficiency of the design

Let us consider the class of spring balance weighing designs $\mathbf{\Phi}_{n \times p}(0, 1)$. The design $\mathbf{X}_d \in \Phi_{n \times p}(0, 1)$ is D-optimal if $\det(\mathbf{X}_d'\mathbf{G}^{-1}\mathbf{X}_d) = \max\{\det\mathbf{M} : \mathbf{X} \in \mathbf{\Phi}\}$, where $\mathbf{M} = \mathbf{X'G^{-1}X}$ is called the information matrix for the design $\mathbf{X}$. The issues concerning the determination of D-optimal designs were presented in the literature; see Raghavarao (1971), Banerjee (1975), Jacroux et al. (1983), Shah and Sinha (1989).

From now on, we have been working under assumption that the number of errors is even and the variance matrix of errors $\sigma^2\mathbf{G}$ is given by $\mathbf{G} = \mathbf{I}_n$, i.e. the errors are uncorrelated and they have the same variances. Neubauer et al. (1997) proved the following Theorem.

**Theorem 1.** If $p$ is even, then any non-singular spring balance weighing design $\mathbf{X} \in \mathbf{\Phi}_{n \times p}(0, 1)$ is regular D-optimal if and only if $\mathbf{X'X} = \dfrac{(p+2)n}{4(p+1)}\left(\mathbf{I}_p + \mathbf{1}_p\mathbf{1}_p'\right)$.

The condition given above implies that in any class $\mathbf{\Phi}_{n\times p}(0, 1)$, a regular D-optimal design does not exist. In such cases, a highly D-efficiency design is considered. We recall the definition given by Bulutoglu and Ryan (2009).

**Definition 1.** D-efficiency of the design $X \in \Phi_{n\times p}(0, 1)$ is given by the formula

$$\mathrm{D}_{\mathrm{eff}}(\mathbf{X}) = \left( \frac{\det(\mathbf{X'X})}{\max_{\mathbf{Y}\in\mathbf{\Phi}_{n\times p}(0,1)}\det(\mathbf{Y'Y})} \right)^{\frac{1}{p}}.$$

Usually, $\max_{\mathbf{Y}\in\mathbf{\Phi}_{n\times p}(0,1)}\det(\mathbf{Y'Y})$ is not known. To solve this problem, we consider the upper bound of $\det(\mathbf{Y'Y})$ and consequently

$$\mathrm{D}_{\mathrm{eff}}(\mathbf{X}) = \left( \frac{\det(\mathbf{X'X})}{\det(\mathbf{Y'Y})} \right)^{\frac{1}{p}}. \tag{1}$$

With the use of equation (1), we can determine a highly D-efficient design when $\mathrm{D}_{\mathrm{eff}}(\mathbf{X}) \geq 0.9$ or a regular D-optimal design when $\mathrm{D}_{\mathrm{eff}}(\mathbf{X}) = 1$.

Based on the results given by Ceranka and Graczyk (2016), we can formulate $\mathrm{D}_{\mathrm{eff}}(\mathbf{X})$ as the function dependent on the number of objects $p$ in the form

$$\mathrm{D}_{\mathrm{eff}}(\mathbf{X}) = \frac{p(p+1)}{(p-1)(p+2)}\left( \frac{p-1}{p+1} \right)^{\frac{1}{p}}$$ and the condition determining a highly D-efficiency design. Thus, we obtain the following Theorem.

**Theorem 2.** Any non-singular spring balance weighing design is highly D-efficient if and only if

$$\mathbf{X'X} = \frac{p}{4(p-1)}\mathbf{I}_p + \frac{n(p-2)}{4(p-1)}\mathbf{1}_p\mathbf{1}_p'. \tag{2}$$

Now, the problem is how to construct the design matrix. We suggest forming this matrix on the basis of a set of incidence matrices of balanced incomplete block designs and group divisible designs. In such a construction $p = v$ and $n = b$, where $v$ and $b$ are the parameters of block designs. It is worth noting that some issues related to a highly D-efficient design are given in Ceranka and Graczyk (2017).

# 3. Construction of highly D-efficiency designs

Let us consider any spring balance weighing design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ in the form

$$\mathbf{X} = \mathbf{N'}, \tag{3}$$

where $\mathbf{N}$ is the incidence matrix of the balanced incomplete block design with the parameters $v$, $b$, $r$, $k$, $\lambda$. For detailed properties of balanced incomplete block designs, we refer the reader to Raghavarao and Padgett (2005).

**Theorem 3.** Let $v$ be an even number and $\mathbf{N}$ the incidence matrix of the balanced incomplete block design with the parameters $v$, $b$, $r$, $k$, $\lambda$. If the conditions
(i)      $b = 2r$
(ii)     $4\lambda (v - 1) = b(v - 2)$
are fulfilled simultaneously, then $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ is a highly D-efficient spring balance weighing design.

    **Proof**. Let us consider $\mathbf{X} \in \Phi_{n \times p}(0, 1)$. Based on equation (2), we obtain

$$\mathbf{X'X} = \frac{bv}{4(v-1)} \mathbf{I}_p + \frac{b(v-2)}{4(v-1)} \mathbf{1}_p \mathbf{1}_p'. \tag{4}$$

    On the other hand, $\mathbf{X'X} = \mathbf{NN'} = (r - \lambda)\mathbf{I}_p + \lambda \mathbf{l}_p \mathbf{l}_p'$. Thus, condition (4) is satisfied if and only if (i) and (ii) hold.

    **Corollary 1.** Any $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ in the form $\mathbf{X} = \mathbf{N'}$, where $\mathbf{N}$ is the incidence matrix of the balanced incomplete block design with the parameters $v = 2t$, $b = 2t(2t - 1)$, $r = t(2t - 1)$, $k = t$, $\lambda = t(t - 1)$, $t = 2, 3, \ldots$, is a highly D-efficient spring balance weighing design.

    **Corollary 2.** Any $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ in the form $\mathbf{X} = \mathbf{N'}$, where $\mathbf{N}$ is the incidence matrix of the balanced incomplete block design with the parameters $v = 2t$, $b = \binom{2t}{t}$, $r = \binom{2t-1}{t-1}$, $k = t$, $\lambda = \binom{2(t-1)}{t-2}$, $t = 2, 3, \ldots$, is a highly D-efficient spring balance weighing design.

    Based on the incidence matrices of two group divisible designs with the same association scheme, we construct a highly D-efficient spring balance weighing design. For this purpose, we consider the design $\mathbf{X}$ in the form

$$\mathbf{X} = \begin{bmatrix} \mathbf{N}_1' \\ \mathbf{N}_2' \end{bmatrix}, \tag{5}$$

where $\mathbf{N}_u$, $u = 1, 2$, is the incidence matrix of the group divisible design with the same association scheme with the parameters $v$, $b_u$, $r_u$, $k_u$, $\lambda_{1u}$, $\lambda_{2u}$, $n = b_1 + b_2$. Furthermore, let the condition

$$\lambda_{11} + \lambda_{12} = \lambda_{21} + \lambda_{22} \tag{6}$$

be satisfied.

**Theorem 4.** Let $v$ be an even number and let $\mathbf{N}_1$ and $\mathbf{N}_2$ be the incidence matrices of two group divisible designs with the same association scheme with the parameters $v$, $b_u$, $r_u$, $k_u$, $\lambda_{1u}$, $\lambda_{2u}$, $u = 1, 2$, for which condition (6) is satisfied. If the conditions
(i) $b_1 + b_2 = 2(r_1 + r_2)$,
(ii) $4\lambda(v - 1) = (v - 2)(b_1 + b_2)$
are fulfilled simultaneously, then $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ is a highly D-efficient spring balance weighing design.

**Proof**. For the design matrix $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ given by formula (5) we have

$$\mathbf{X}'\mathbf{X} = \mathbf{N}_1\mathbf{N}_1' + \mathbf{N}_2\mathbf{N}_2' = \frac{(b_1 + b_2)v}{4(v-1)}\mathbf{I}_v + \frac{(b_1 + b_2)(v - 2)}{4(v-1)}\mathbf{1}_v\mathbf{1}_v'. \tag{7}$$

On the other hand, $\mathbf{N}_1\mathbf{N}_1' + \mathbf{N}_2\mathbf{N}_2' = (r_1 + r_2 - \lambda)\mathbf{I}_v + \lambda\mathbf{1}_v\mathbf{1}_v'$. Under the above assumptions, (7) is satisfied if and only if $\lambda = \dfrac{(v - 2)(b_1 + b_2)}{4(v-1)}$, hence (ii). Taking into consideration the equation $\dfrac{(b_1 + b_2)v}{4(v-1)} = r_1 + r_2 - \lambda$, we obtain (i). Hence, we obtain our result.

Now, we present series of parameters of group divisible designs based on Clatworthy (1973).

**Theorem 5.** Let $\mathbf{N}_u$, $u = 1, 2$, be the incidence matrices of the group divisible design with the same association scheme with the parameters $v = 2k$ and $b_u = 2r_u$. If

1) $k = 2$ and

1.1) $r_1 = 3t + 1$, $\lambda_{11} = t + 1$, $\lambda_{21} = t$ and $r_2 = 3s + 2$, $\lambda_{12} = s$, $\lambda_{22} = s + 1$, $t = 1, 2, 3$, $s = 0, 1, 2,$

1.2) $r_1 = 3t + 2$, $\lambda_{11} = t + 2$, $\lambda_{21} = t$ and $r_2 = 3s + 4$, $\lambda_{12} = s$, $\lambda_{22} = s + 2$, $t = 1, 2$, $s = 0, 1, 2,$

1.3) $r_1 = t + 3$, $\lambda_{11} = t + 1$, $\lambda_{21} = 1$ and $r_2 = 2t$, $\lambda_{12} = 0$, $\lambda_{22} = t$, $t = 1, 2, ..., 5,$

1.4) $r_1 = 8$, $\lambda_{11} = 0$, $\lambda_{21} = 4$ and $r_2 = 3s + 4$, $\lambda_{12} = s + 4$, $\lambda_{22} = s$, $s = 1, 2$,

1.5) $r_1 = 9$, $\lambda_{11} = 5$, $\lambda_{21} = 2$ and $r_2 = 3(s + 2)$, $\lambda_{12} = s$, $\lambda_{22} = s + 3$, $s = 0, 1$,

2) $k = 3$ and

2.1) $r_1 = 2t$, $\lambda_{11} = 0$, $\lambda_{21} = t$ and $r_2 = 3t$, $\lambda_{12} = 2t$, $\lambda_{22} = t$, $t = 1, 2, 3$,

2.2) $r_1 = 2t + 5$, $\lambda_{11} = t + 1$, $\lambda_{21} = t + 2$ and $r_2 = 3t$, $\lambda_{12} = t + 1$, $\lambda_{22} = t$, $t = 1, 2$,

2.3) $r_1 = 6$, $\lambda_{11} = 4$, $\lambda_{21} = 2$ and $r_2 = 5s + 4$, $\lambda_{12} = 2s$, $\lambda_{22} = 2(s + 1)$, $s = 0, 1$,

2.4) $r_1 = 8$, $\lambda_{11} = 4$, $\lambda_{21} = 3$ and $r_2 = 5s + 2$, $\lambda_{12} = 2s$, $\lambda_{22} = 2s + 1$, $s = 0, 1$,

3) $k = 4$ and

3.1) $r_1 = 2(t + 1)$, $\lambda_{11} = 0$, $\lambda_{21} = t + 1$ and $r_2 = 2(6 - t)$, $\lambda_{12} = 6$, $\lambda_{22} = 5 - t$, $t = 1, 2, 3$,

3.2) $r_1 = 3t + 2$, $\lambda_{11} = t + 2$, $\lambda_{21} = t + 1$ and $r_2 = 3(4 - t)$, $\lambda_{12} = 5 - t$, $\lambda_{22} = 5 - t$, $t = 1, 2$,

4) $k = 4$ and $r_1 = 4t$, $\lambda_{11} = 0$, $\lambda_{21} = 2$ and $r_2 = 5t$, $\lambda_{12} = 4t$, $\lambda_{22} = 2t$, $t = 1, 2$,

5) $k = 2t + 1$ and $r_1 = 2t$, $\lambda_{11} = 0$, $\lambda_{21} = t$ and $r_2 = 2t + 1$, $\lambda_{12} = 2t$, $\lambda_{22} = t$, $t = 1, 2, 3, 4$,

6) $k = 2(t + 1)$ and $r_1 = 2t + 1$, $\lambda_{11} = 2t + 1$, $\lambda_{21} = t$ and $r_2 = 2(t + 1)$, $\lambda_{12} = 0$, $\lambda_{22} = t + 1$

$t = 1, 2, 3, 4$,

then any $\mathbf{X} \in \mathbf{\Phi}_{n \times p}(0, 1)$ in the form (5) is a highly D-efficient spring balance weighing design.

**Proof**. Clearly, the parameters given above satisfy conditions (i)–(ii) of Theorem 4.

# 4. Example

Here, we consider the experiment in which we determine unknown measurements of $p = 4$ objects using $n = 12$ measurements. We are interested in determining the design having the best statistical properties in the class $\mathbf{X} \in \mathbf{\Phi}_{12 \times 4}(0, 1)$. Firstly, let us consider the existence of a regular D-optimal design. Based on Theorem 1, a regular D-optimal spring balance weighing design in the class $\mathbf{\Phi}_{12 \times 4}(0, 1)$ does not exist, as $0.25n(p + 2)(p + 1)^{-1} = 3.6$ is not an integer. Thus, in this class, we determine a highly D-efficient design. Based on Theorem 5, subsection 1.1, for $t = 1$ and $s = 0$, we take the group divisible designs with the parameters $v = 4$, $b_1 = 8$, $r_1 = 4$, $k_1 = 2$, $\lambda_{11} = 2$, $\lambda_{21} = 1$ and $v = 4$, $b_4 = 8$, $r_2 = 2$, $k_2 = 2$, $\lambda_{12} = 0$, $\lambda_{22} = 1$ given by the incidence matrices

$$\mathbf{N}_1 = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \end{bmatrix} \quad \mathbf{N}_2 = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}.$$

and

Therefore, we form the highly D-efficient design $\mathbf{X} \in \mathbf{\Phi}_{12 \times 4}(0, 1)$ as

$$\mathbf{X} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

and $D_{eff}(\mathbf{X}) = 0.9779$.

## 5. Discussion

As you can see from Theorem 1, it is not possible to construct a regular D-optimal spring balance weighing design in any class $\mathbf{\Phi}_{n \times p}(0, 1)$. For example, in the classes $\mathbf{\Phi}_{n \times p}(0, 1)$ for $p = 4$ objects and $n = 5, 6, 7, 8, 9$ measurements, a D-optimal design does not exist as for these numbers $0.25n(p + 2)(p + 1)^{-1}$ is not an integer. On the other hand, the properties of the design matrix $\mathbf{X}$ should be taken into consideration. Hence, we introduce highly D-efficient designs relevantly good from the point of view of optimality. Moreover, we present a new construction method of this type of design.

## References

Banerjee K.S. (1975), *Weighing Designs for Chemistry, Medicine, Economics, Operations Research, Statistics*, Marcel Dekker Inc., New York.
Bulutoglu D.A., Ryan K.J. (2009), *D-optimal and near D-optimal 2ᵏ fractional factorial designs of resolution V*, "Journal of Statistical Planning and Inference", vol. 139, pp. 16–22.

Ceranka B., Graczyk M. (2014a), *The problem of D-optimality in some experimental designs*, "International Journal of Mathematics and Computer Application Research", vol. 4, pp. 11–18.

Ceranka B., Graczyk M. (2014b), *Regular E-optimal spring balance weighing designs with correlated errors*, "Communication in Statistics – Theory and Methods", vol. 43, pp. 947–953.

Ceranka B., Graczyk M. (2014c), *On certain A-optimal biased spring balance weighing designs*, "Statistics in Transition new series", vol. 15(2), pp. 317–326.

Ceranka B., Graczyk M. (2016), *Recent developments in D-optimal spring balance weighing designs*, to appear.

Ceranka B., Graczyk M. (2017), *Highly D-efficient designs for even number of objects*, Revstat.

Clatworthy W.H. (1973), *Tables of Two-Associated-Class Partially Balanced Design*, NBS Applied Mathematics Series 63.

Jacroux M., Wong C.S., Masaro J.C. (1983), *On the optimality of chemical balance weighing design*, "Journal of Statistical Planning and Inference", vol. 8, pp. 213–240.

Masaro J., Wong C.S. (2008a), *Robustness of A-optimal designs*, "Linear Algebra and its Applications", vol. 429, pp. 1392–1408.

Masaro J., Wong C.S. (2008b), *D-optimal designs for correlated random errors*, "Journal of Statistical Planning and Inference", vol. 130, pp. 4093–4106.

Neubauer M.G., Watkins S., Zeitlin J. (1997), *Maximal j-simpplices in the real d-dimensional unit cube*, "Journal of Combinatorial Theory", Ser. A 80, pp. 1–12.

Raghavarao D. (1971), *Constructions and combinatorial problems in design of experiment*, John Wiley and Sons, New York.

Raghavarao D., Padgett L.V. (2005), *Block Designs, Analysis, Combinatorics and Applications*, Series of Applied Mathematics 17, Word Scientific Publishing Co. Pte. Ltd., Singapore.

Shah K.R., Sinha B.K. (1989), *Theory of Optimal Designs*, Springer-Verlag, Berlin.

**O pewnych aspektach efektywności w sprężynowych układach wagowych**

**Streszczenie:** W artykule zaprezentowane zostały problemy związane z optymalnością układu doświadczalnego z punktu widzenia sprężynowych układów wagowych. Przeanalizowano własności D-optymalnych i wysoce D-efektywnych układów. Podano warunki konieczne i dostateczne wyznaczające te układy. Ponadto zaprezentowano nową metodę konstrukcji tych układów, opartą na macierzach incydencji układów o grupach podzielnych oraz układów zrównoważonych o blokach niekompletnych.

**Słowa kluczowe:** sprężynowy układ wagowy, układ efektywny, układ o grupach podzielnych, układ zrównoważony o blokach niekompletnych

**JEL:** C02, C18, C90