



Małgorzata Graczyk 

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

Bronisław Ceranka 

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

A Highly D-efficient Spring Balance Weighing Design for an Even Number of Objects

Abstract: The problem of determining unknown measurements of objects in the model of spring balance weighing designs is presented. These designs are considered under the assumption that experimental errors are uncorrelated and that they have the same variances. The relations between the parameters of weighing designs are deliberated from the point of view of optimality criteria. In the paper, designs in which the product of the variances of estimators is possibly the smallest one, i.e. D-optimal designs, are studied. A highly D-efficient design in classes in which a D-optimal design does not exist are determined. The necessary and sufficient conditions under which a highly efficient design exists and methods of its construction, along with relevant examples, are introduced.

Keywords: highly D-efficient design, spring balance weighing design

JEL: C02, C18, C90

1. Introduction

The problems presented in this paper refer to the issues related to spring balance weighing designs. According to the standard works on this subject, 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 $\mathbf{y} = \mathbf{X}\mathbf{w} + \mathbf{e}$, where \mathbf{y} is a $n \times 1$ random vector of the recorded results of measurements, $\mathbf{X} = (x_{ij}) \in \Phi_{n \times p}(0, 1)$. $\Phi_{n \times p}(0, 1)$ denotes the class of matrices with elements $x_{ij} = 1$ or 0 , $i = 1, 2, \dots, n, j = 1, 2, \dots, p$, and \mathbf{w} is a $p \times 1$ vector of unknown measurements of objects. For an $n \times 1$ random vector of errors \mathbf{e} , we shall make the standing assumptions on the maps under consideration: $E(\mathbf{e}) = \mathbf{0}_n$ and $E(\mathbf{e}\mathbf{e}') = \sigma^2\mathbf{G}$, \mathbf{G} is known as a positive definite matrix. A more complete theory on applications of the spring balance weighing design may be obtained from Raghavarao (1971: 319–321), Banerjee (1975: 33–48), Shah and Sinha (1989: 1–15). Our purpose is to determine unknown measurements of p objects in n measurement operations. Statistically speaking, we determine the vector $\hat{\mathbf{w}} = (\hat{w}_1, \hat{w}_2, \dots, \hat{w}_p)'$ by the least squares method with the use of normal equations $\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\hat{\mathbf{w}} = \mathbf{X}'\mathbf{G}^{-1}\mathbf{y}$. In the case when \mathbf{X} is of full column rank, the least squares estimator of \mathbf{w} is given by $\hat{\mathbf{w}} = (\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{G}^{-1}\mathbf{y}$ and the covariance matrix of $\hat{\mathbf{w}}$ is equal to $V(\hat{\mathbf{w}}) = \sigma^2(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}$. Moreover, here we are concerned with determining an estimator of \mathbf{w} having the product of its variances of its components as small as possible. Subsequently, the D-optimal criterion as the function of the matrix $(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}$ is considered. The standard approach of determining a D-optimal design can be easily found (see Jacroux, Notz, 1983: 213–230; Jacroux, Wong, Masaro, 1983: 213–240; Shah, Sinha, 1989: 1–15; Masaro, Wong, 2008a: 1392–1400; 2008b: 4093–4101; Ceranka, Graczyk, 2014: 12–14; 2018: 476–481; 2019). From these papers, we get that the design $\mathbf{X}_d \in \Phi_{n \times p}(0, 1)$ is D-optimal if $\det(\mathbf{X}_d'\mathbf{G}^{-1}\mathbf{X}_d)^{-1} = \min\{\det(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X})^{-1}: \mathbf{X} \in \Phi_{n \times p}(0, 1)\}$. The crucial fact is that not in any class $\Phi_{n \times p}(0, 1)$ we are able to determine a D-optimal spring balance weighing design.

An attempt has been made here to expand the theory of optimal designs. The aim of this research is to develop the results concerning new methods of determining optimal designs in classes in which they have not been determined in the literature so far. From now on we make the assumption that $\mathbf{G} = \mathbf{I}_n$.

2. Highly D-efficient design

Let us consider the class of spring balance weighing designs $\Phi_{n \times p}(0, 1)$. Some investigations of a regular D-optimal design for even p are dealt with by Neubauer, Watkins and Zeitlin (1997: 2–8). They have provided the conditions under which any spring balance weighing design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ is optimal.

Theorem 1. Let p be even. Any non-singular spring balance weighing design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ is regular D-optimal if and only if

- a) $\mathbf{X}'\mathbf{X} = t(\mathbf{I}_p + \mathbf{1}_p \mathbf{1}'_p)$
- b) each row of \mathbf{X} contains $0.5p$ or $0.5p + 1$ ones.

According to the conditions given above, in any class $\Phi_{n \times p}(0, 1)$, a regular D-optimal design does not exist. In such a case, a highly D-efficient design is considered. For details, we refer the reader to Bulutoglu, Ryan (2009: 17–20). Ceranka and

Graczyk (2018: 477–478) gave the definition of D-efficiency as $D_{eff} = \left(\frac{\det(\mathbf{X}'\mathbf{X})}{\det(\mathbf{Y}'\mathbf{Y})} \right)^{\frac{1}{p}}$,

where \mathbf{Y} is the matrix of the regular D-optimal spring balance weighing design given in Theorem 1. We indicate the highly D-efficient design when $D_{eff} \geq 0.95$.

Therefore, the aim of the study presented here is to provide a highly D-efficient design in the class $\Phi_{n \times p}(0, 1)$. The advantage of using the highly D-efficient design lies in the fact that the product of the variances of obtained estimators is second best compared to the regular D-optimal design. Simultaneously, we determine highly D-efficient designs in classes in which a regular D-optimal design does not exist.

Ceranka and Graczyk (2018: 478–479) gave the following Theorem.

Theorem 2. Let p be even. In any non-singular spring balance weighing design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ having $0.5p$ ones in each row $\det(\mathbf{X}'\mathbf{X}) \leq (p-1) \left(\frac{np}{4(p-1)} \right)^p$.

An upper bound is attained if and only if $\mathbf{X}'\mathbf{X} = \frac{n}{4(p-1)}(p\mathbf{I}_p + (p-2)\mathbf{1}_p \mathbf{1}'_p)$, where $0.25np(p-1)^{-1}$ and $0.25n(p-2)(p-1)^{-1}$ are integers.

The design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ of the form given in Theorem 2 is considered as highly D-efficient.

In classes in which regular a D-optimal spring balance weighing design does not exist, highly D-efficient designs are determined. In the presented paper, we study the designs with uncorrelated experimental errors according to the design matrix in the form

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_1 \\ \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \\ \mathbf{x}_4 \end{bmatrix}, \quad (1)$$

where \mathbf{X}_1 is the matrix of the highly D-efficient spring balance weighing design $\Phi_{(n-4) \times p}(0, 1)$ and \mathbf{x}_h , $h = 1, 2, 3, 4$, are $p \times 1$ vectors of 0 and 1. The problem is to formulate the relations between $\mathbf{X}_1 \in \Phi_{(n-4) \times p}(0, 1)$ and vectors \mathbf{x}_h and to determine the replacing of one of the elements in \mathbf{x}_h . In order to determine the highly D-efficient design $\mathbf{X} \in \Phi_{n \times p}(0, 1)$, we have to give the upper bound of $\det(\mathbf{X}'\mathbf{X})$. Thus, for $\mathbf{X} \in \Phi_{n \times p}(0, 1)$ in the form (1) we calculate

$$\begin{aligned} \det(\mathbf{X}'\mathbf{X}) &= \det(\mathbf{X}'_1\mathbf{X}_1 + [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4][\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4]') = \\ &= \det(\mathbf{X}'_1\mathbf{X}_1) \det(\mathbf{I}_4 + [\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4](\mathbf{X}'_1\mathbf{X}_1)^{-1}[\mathbf{x}_1 \ \mathbf{x}_2 \ \mathbf{x}_3 \ \mathbf{x}_4]'), \end{aligned}$$

see in Harville (1997: 416, 419). We assume that the vectors \mathbf{x}_h are of the form

$$\begin{cases} \mathbf{x}'_h \mathbf{1}_p = t_h & 0 < t_h \leq p \\ \mathbf{x}'_h \mathbf{x}_{h'} = \mathbf{x}'_{h'} \mathbf{x}_h = u_{hh'}, & 0 \leq u_{hh'} \leq \min(t_h, t_{h'}) \\ h, h' = 1, 2, 3, 4, h \neq h' \end{cases}$$

For $\mathbf{X}_1 \in \Phi_{(n-4) \times p}(0, 1)$, we have $\det(\mathbf{X}'_1\mathbf{X}_1) = (p-1) \left(\frac{(n-4)p}{4(p-1)} \right)^p$

and

$$\mathbf{X}'_1\mathbf{X}_1 = \frac{n-4}{4(p-1)} (p\mathbf{I}_p + (p-2)\mathbf{1}_p\mathbf{1}'_p)$$

and here $0.25(n-4)p(p-1)^{-1}$ and $0.25(n-4)(p-2)(p-1)^{-1}$ are integers. Moreover,

$$\mathbf{X}'_1\mathbf{X}_1^{-1} = \frac{4(p-1)}{n-4} \left(p\mathbf{I}_p - \frac{p-2}{4(p-1)} \mathbf{1}_p\mathbf{1}'_p \right).$$

Consequently,

$$\det(\mathbf{X}'\mathbf{X}) \leq (p-1) \left(\frac{(n-4)p}{4(p-1)} \right)^p \det(\mathbf{T}),$$

where

$$\mathbf{T} = \mathbf{I}_4 + \frac{4(p-1)}{(n-4)p} [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3 \quad \mathbf{x}_4] \left(\mathbf{I}_p - \frac{p-2}{p(p-1)} \mathbf{1}_p \mathbf{1}'_p \right) [\mathbf{x}_1 \quad \mathbf{x}_2 \quad \mathbf{x}_3 \quad \mathbf{x}_4]'$$

The matrix \mathbf{T} has the diagonal elements equal to

$$1 + \frac{4(p-1)}{(n-4)p} \left(t_h - \frac{p-2}{p(p-1)} t_h^2 \right)$$

and off-diagonal of the form

$$\frac{4(p-1)}{(n-4)p} \left(u_{hh'} + \frac{p-2}{p(p-1)} t_h t_{h'} \right), \quad h, h' = 1, 2, 3, 4, h \neq h'$$

Next,

$$\det(\mathbf{T}) = A + 2B - D + \left(\frac{4(p-1)}{(n-4)p} \right)^4 (C - 2E),$$

where

$$A = \prod_{h=1}^4 \left(1 + \frac{4(p-1)}{(n-4)p} \left(t_h - \frac{p-2}{p(p-1)} t_h^2 \right) \right),$$

$$B = \sum_{h=1}^4 \left(1 + \frac{4(p-1)}{(n-4)p} \left(t_h - \frac{p-2}{p(p-1)} t_h^2 \right) \right) \prod_{s,z} \frac{4(p-1)}{(n-4)p} \left(u_{sz} + \frac{p-2}{p(p-1)} t_s t_z \right),$$

$$s, z = 1, 2, 3, 4, s \neq z \neq h, s < z,$$

$$C = \sum_{h,s,z} \left(u_{1h} - \frac{p-2}{p(p-1)} t_1 t_h \right)^2 \left(u_{sz} - \frac{p-2}{p(p-1)} t_s t_z \right)^2,$$

$$h, s, z = 2, 3, 4, s \neq z \neq h, s < z,$$

$$D = \sum_{h,h',s,z} \left(1 + \frac{4(p-1)}{(n-4)p} \left(t_h - \frac{p-2}{p(p-1)} t_h^2 \right) \right) \left(1 + \frac{4(p-1)}{(n-4)p} \left(t_{h'} - \frac{p-2}{p(p-1)} t_{h'}^2 \right) \right) \frac{4(p-1)}{(n-4)p} \left(u_{sz} + \frac{p-2}{p(p-1)} t_s t_z \right),$$

$$h = 1, 2, 3, h' = 2, 3, 4, s, z = 1, 2, 3, 4, s \neq z \neq h \neq h', s < z, h < h',$$

$$E = \left(u_{12} - \frac{p-2}{p(p-1)} t_1 t_2 \right) \left(u_{34} - \frac{p-2}{p(p-1)} t_3 t_4 \right) \left(u_{13} - \frac{p-2}{p(p-1)} t_1 t_3 \right) \\ \left(\left(u_{24} - \frac{p-2}{p(p-1)} t_2 t_4 \right) + \left(u_{14} - \frac{p-2}{p(p-1)} t_1 t_4 \right) \left(u_{23} - \frac{p-2}{p(p-1)} t_2 t_3 \right) \right) + \\ \left(u_{13} - \frac{p-2}{p(p-1)} t_1 t_3 \right) \left(u_{14} - \frac{p-2}{p(p-1)} t_1 t_4 \right) \left(u_{23} - \frac{p-2}{p(p-1)} t_2 t_3 \right) \left(u_{24} - \frac{p-2}{p(p-1)} t_2 t_4 \right).$$

As we want to maximise $\det(\mathbf{T})$, we simultaneously determine the maximum values of

$$t_h - \frac{p-2}{p(p-1)} t_h^2 \tag{2}$$

and the minimum values of

$$\left(u_{hh'} - \frac{p-2}{p(p-1)} t_h t_{h'} \right)^2. \tag{3}$$

The maximum values of (2) are all attained if and only if $t_h = 0.5(p+2)$ and in that case

$$t_h - \frac{p-2}{p(p-1)} t_h^2 = \frac{p^3 + 8}{4p(p-1)}.$$

If $p = 0 \pmod 4$, then the minimum values of (3) are equal to

$$\frac{(p^2 + 8)^2}{16p^2 (p-1)^2}.$$

In this case

$$\det(\mathbf{T}) = \left(1 + \frac{p-1}{n-4}\right)^3 \left(1 + \frac{p^3 + 3p^2 + 32}{(n-4)p^2}\right)$$

and

$$\det(\mathbf{X}'\mathbf{X}) \leq (p-1) \left(\frac{(n-4)p}{4(p-1)}\right)^p \left(1 + \frac{p-1}{n-4}\right)^3 \left(1 + \frac{p^3 + 3p^2 + 32}{(n-4)p^2}\right). \quad (4)$$

The equality in (4) holds if and only if $t_h = 0.5(p + 2)$ and $u_{hh'} = 0.25(p + 2)$.

If $p + 2 = 0 \pmod 4$, then the minimum values of (3) are equal to

$$\frac{(p+2)^2 (p-4)^2}{16p^2 (p-1)^2}.$$

In this case

$$\det(\mathbf{T}) = \left(1 + \frac{(p-1)(p+2)}{(n-4)p}\right)^3 \left(1 + \frac{(p+2)(p^2 - 5p + 16)}{(n-4)p^2}\right)$$

and

$$\det(\mathbf{X}'\mathbf{X}) \leq (p-1) \left(\frac{(n-4)p}{4(p-1)}\right)^p \left(1 + \frac{(p-1)(p+2)}{(n-4)p}\right)^3 \left(1 + \frac{(p+2)(p^2 - 5p + 16)}{(n-4)p^2}\right). \quad (5)$$

The equality in (5) holds if and only if $t_h = 0.5(p + 2)$ and $u_{hh'} = 0.25(p + 4)$.

3. Example

Our examples demonstrate rather strikingly the manner of construction of the highly D-efficient design. The choice of these cases seems to be the best adapted to our theory. We want to determine unknown measurements of $p = 4$ objects using $n = 10$ measurements. We are interested in determining the design having the best

statistical properties in the class $\mathbf{X} \in \Phi_{10 \times 4}(0, 1)$. So, according to the above-presented theory, we take the design matrix of the highly D-efficient spring balance weighing design \mathbf{X}_1 in the class $\Phi_{6 \times 4}(0, 1)$ and we add four measurements where

$$\mathbf{X}_1 = \begin{pmatrix} 1100 \\ 1010 \\ 1001 \\ 0110 \\ 0101 \\ 0011 \end{pmatrix}.$$

Thus, the design matrix

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ 1110 \\ 1101 \\ 1011 \\ 0111 \end{pmatrix} \in \Phi_{10 \times 4}(0, 1)$$

is highly D-efficient.

We are interested in determining the design having the best statistical properties in the class $\mathbf{X} \in \Phi_{14 \times 6}(0, 1)$. So, according to the above-presented theory, we take the design matrix of the highly D-efficient spring balance weighing design \mathbf{X}_1 in the class $\Phi_{10 \times 6}(0, 1)$ and we add four measurements where

$$\mathbf{X}_1 = \begin{pmatrix} 110010 \\ 110001 \\ 101100 \\ 101001 \\ 100110 \\ 011100 \\ 011010 \\ 010101 \\ 001011 \\ 000111 \end{pmatrix}.$$

So, the design

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ 111000 \\ 110100 \\ 101100 \\ 011100 \end{pmatrix} \in \Phi_{14 \times 6}(0, 1)$$

is highly D-efficient.

The problem of estimation of unknown measurements of objects in the model of spring balance weighing design is presented. Of particular interest is a new construction method of the design matrix \mathbf{X} which allows us to determine the optimal design in the class of matrices $\Phi_{n \times p}(0, 1)$ in the cases not considered in the literature.

4. Conclusions

In the paper, the theory and practice of the spring balance weighing design is presented. It is not possible to determine the D-optimal spring balance weighing design in any class $\Phi_{n \times p}(0, 1)$. So, new construction methods of highly D-efficient designs in such classes are given. In the above-presented examples, the methods of determining highly D-efficient designs in classes in which D-optimal spring balance weighing designs have not been determined so far in the literature are introduced. It is worth noting that in the highly D-efficient spring balance weighing design we are able to determine unknown measurements of object with the minimal product of variances of their estimators.

It is worth emphasising that spring balance weighing designs can be applied in all experiments in which the experimental factors are on two levels. Let us suppose that we study the real estate market and we are interested in the influence of the following factors: population density, the type of occupation, salary, and the location (each on two levels coded with 1 or 0). From the statistical point of view, we are interested in determining the influences of these factors using twenty different combinations. In the notation of weighing designs, we determine unknown measurements of $p = 4$ objects in $n = 10$ surveys, so we consider the class $\Phi_{10 \times 4}(0, 1)$. The scheme of determination of the measurements, i.e. the design matrix, is given in the above-presented example. Possible applications of the discussed designs should be searched wherever the measurement results can be written as a linear combination of unknown object measures with coefficients equal to 0 or 1. The examples of such applications are given in Beckman (1973: 561–565), Banerjee (1975), Ceranka and Katulska (1987: 98–108).

References

- Banerjee K. S. (1975), *Weighing Designs for Chemistry, Medicine, Economics, Operations Research, Statistics*, Marcel Dekker Inc., New York.
- Beckman R. J. (1973), *An applications of multivariate weighing designs*, "Communication in Statistics", no. 1(6), pp. 561–565.
- 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", no. 139, pp. 16–22.
- Ceranka B., Graczyk M. (2014), *The problem of D-optimality in some experimental designs*, "International Journal of Mathematics and Computer Application Research", no. 4, pp. 11–18.
- Ceranka B., Graczyk M. (2018), *Highly D-efficient designs for even number of objects*, "REVSTAT-Statistical Journal", no. 16, pp. 475–486.
- Ceranka B., Graczyk M. (2019), *Recent developments in D-optimal designs. Communication in Statistics – Theory and Methods*, Accepted to publication.
- Ceranka B., Katulska K. (1987), *Zastosowanie optymalnych sprężynowych układów wagowych*, "Siedemnaste Colloquium Metodologiczne z Agro-Biometrii", PAN, pp. 98–108.
- Harville D. A. (1997), *Matrix Algebra from a Statistician's Point of Perspective*, Springer-Verlag, New York.
- Jacroux M., Notz W. (1983), *On the optimality of spring balance weighing designs*, "The Annals of Statistics", no. 11(3), pp. 970–978.
- Jacroux M., Wong C. S., Masaro J. C. (1983), *On the optimality of chemical balance weighing design*, "Journal of Statistical Planning and Inference", no. 8, pp. 213–240.
- Masaro J., Wong C. S. (2008a), *Robustness of A-optimal designs*, "Linear Algebra and its Applications", no. 429, pp. 1392–1408.
- Masaro J., Wong C. S. (2008b), *D-optimal designs for correlated random errors*, "Journal of Statistical Planning and Inference", no. 130, pp. 4093–4106.
- Neubauer M. G., Watkins S., Zeitlin J. (1997), *Maximal j-simplices 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.
- Shah K. R., Sinha B. K. (1989), *Theory of Optimal Designs*, Springer-Verlag, Berlin.

Wysoco D-efektywny sprężynowy układ wagowy dla parzystej liczby obiektów

Streszczenie: W artykule zaprezentowano problemy związane z wyznaczaniem nieznanych miar obiektów w modelu sprężynowego układu wagowego. Układy te badano przy założeniu, że błędy pomiarów są nieskorelowane i mają równe wariancje. Relacje między parametrami układów wagowych rozważano z punktu widzenia kryteriów optymalności. Analizowano takie układy, w których iloczyn wariancji estymatorów jest możliwie najmniejszy, czyli układy D-optymalne. W klasach, w których nie istnieją układy D-optymalne, wyznaczono układy wysoco D-efektywne. Podano warunki konieczne i dostateczne, przy których spełnieniu układy wysoco efektywne istnieją, oraz ich przykładowe metody konstrukcji.

Słowa kluczowe: sprężynowy układ wagowy, układ wysoco D-efektywny

JEL: C02, C18, C90

 <p>OPEN ACCESS</p>	<p>© by the author, licensee Łódź University – Łódź University Press, Łódź, Poland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license CC-BY (http://creativecommons.org/licenses/by/3.0/)</p> <p>Received: 2019-01-11; verified: 2019-03-02. Accepted: 2019-08-29</p>
 <p>COPE Member since 2018 JM13714</p>	<p>This journal adheres to the COPE's Core Practices https://publicationethics.org/core-practices</p>