Colloquium Biometricum 43
2013, 37–45

A NOTE ON D-OPTIMAL CHEMICAL BALANCE WEIGHING DESIGNS AND THEIR APPLICATIONS

Krystyna Katulska, Łukasz Smaga

Faculty of Mathematics and Computer Science
Adam Mickiewicz University of Poznań
Umultowska 87, 61-614 Poznań, Poland
e-mails: krakat@amu.edu.pl, ls@amu.edu.pl

Summary

In this paper, we consider the chemical balance weighing designs for estimation of individual unknown weights of objects using D-optimality criterion. We suppose that the random errors are equally non-negative correlated and they have equal variances. The upper bound of the determinant of the information matrix of weighing designs is proved. Some sufficient conditions for this upper bound to be attained are given. The construction of D-optimal designs, which satisfy this upper bound, is also presented. Some application of optimal weighing designs in bioengineering experiment given in the literature is also described.

Keywords and phrases: chemical balance weighing design, correlated errors, D-optimality criterion

Classification AMS 2010: 62K05, 05B20

1. Introduction

Let us assume that an $n\times 1$ vector of observations $\mathbf{y}=[y_1, y_2, \ldots, y_n]'$ follow the linear model $\mathbf{y}=\mathbf{Xw}+\mathbf{e}$, where $\mathbf{w}=[w_1, w_2, \ldots, w_p]'$ is a vector of unknown weights of p objects, $\mathbf{X}=[x_{ij}]$ is the $n\times p$ design matrix,

$\mathbf{e} = [e_1, e_2, \ldots, e_n]'$ is the vector of random errors. In the chemical balance weighing design, $x_{ij} = 1$ $(x_{ij} = -1)$ if the jth object is placed on the right (left) pan during the ith weighing operation. We assume that $\mathrm{E}(\mathbf{e}) = [0, 0, \ldots, 0]'$ is $n \times 1$ null vector and $\mathrm{Var}(\mathbf{e}) = \sigma^2 \mathbf{G}$, where $\sigma > 0$ is the known parameter and the matrix $\mathbf{G}$ is the known positive definite matrix. The weighing design is identified with its design matrix.

Among all the weighing designs we would like to choose the best one for a specific criterion. We consider D-optimality criterion. The design $\hat{\mathbf{X}}$ is D-optimal in the class of the designs $C \subseteq M_{n \times p}(\{-1,1\})$, where $M_{n \times p}(\{-1,1\})$ is the set of all $n \times p$ matrices with elements 1 or -1, if $\det(\hat{\mathbf{X}}'\mathbf{G}^{-1}\hat{\mathbf{X}}) = \max\{\det(\hat{\mathbf{X}}'\mathbf{G}^{-1}\hat{\mathbf{X}}) : \mathbf{X} \in C\}$ The matrix $\hat{\mathbf{X}}'\mathbf{G}^{-1}\hat{\mathbf{X}}$ is called the information matrix of the weighing design $\mathbf{X}$.

Many results about D-optimal weighing designs are known in the literature (see, for example, Galil and Kiefer 1980, Jacroux et al. 1983) when $\mathbf{G}$ is the identity matrix, i.e. random errors are uncorrelated and have equal variances. The case where random errors form the first order autoregressive process was considered in Katulska and Smaga (2012, 2013), Li and Yang (2005), Yeh and Lo Huang (2005). In Masaro and Wong (2008), D-optimal weighing designs in some subclasses of the class of weighing designs for p objects assuming that the random errors are equally correlated and they have equal variances are considered.

In this paper, we present the results about D-optimal chemical balance weighing designs assuming that the random errors are equally non-negative correlated and they have equal variances. Some construction of the D-optimal weighing designs is also given. At the end, we present the application of optimal weighing designs in bioengineering experiment given in the literature.

2. D-optimal chemical balance weighing designs

In this section, we present some results about the D-optimal weighing designs when errors are equally non-negative correlated and they have equal variances. These assumptions imply that the covariance matrix of errors can be written in the form $\sigma^2 \mathbf{G}$, where

$$\mathbf{G} = g[(1-\rho)\mathbf{I}_n + \rho \mathbf{1}_n \mathbf{1}_n'], \qquad (2.1)$$

$g > 0$ and $\rho \in [0, 1)$. The inverse matrix of $\mathbf{G}$ is given by the following formula

$$\mathbf{G}^{-1} = \frac{1}{g(1-\rho)}\left[\mathbf{I}_n - \frac{\rho}{1+(n-1)\rho}\mathbf{1}_n\mathbf{1}_n'\right]. \tag{2.2}$$

The following lemma describes elements of the information matrix of weighing designs with the covariance matrix of errors $\sigma^2\mathbf{G}$, where the matrix $\mathbf{G}$ is given by (2.1).

Lemma 2.1. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p] \in \mathrm{M}_{n\times p}(\{-1, 1\})$ and the matrix $\mathbf{G}$ be given by (2.1). If $i, j = 1, 2, \ldots, p,$ then

$$\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)_{ij} = \mathbf{x}_i'\mathbf{G}^{-1}\mathbf{x}_j = \frac{1}{g(1-\rho)}\left[\mathbf{x}_i'\mathbf{x}_j - \frac{(\mathbf{x}_i'\mathbf{1}_n)(\mathbf{x}_j'\mathbf{1}_n)'\rho}{1+(n-1)\rho}\right].$$

Now, we prove the inequality which gives the upper bound for determinant of the information matrix of the weighing design in $\mathbf{M}_{n\times p}(\{-1, 1\})$.

Theorem 2.1. If $g > 0, \rho \in [0, 1)$ and $\mathbf{X} \in \mathrm{M}_{n\times p}(\{-1, 1\})$, then

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le \left(\frac{n}{g(1-\rho)}\right)^p, \tag{2.3}$$

where $\mathbf{G}$ is given by the formula (2.1).

Proof. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p] \in \mathrm{M}_{n\times p}(\{-1, 1\})$ and $\mathbf{G}$ be given by (2.1). By the Hadamard's inequality it follows that the determinant of the information matrix is smaller or equal to the product of diagonal elements of this matrix, it is

$$\det\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right) \le \prod_{k=1}^{p}\left(\mathbf{X}'\mathbf{G}^{-1}\mathbf{X}\right)_{kk} = \prod_{k=1}^{p}\mathbf{x}_k'\mathbf{G}^{-1}\mathbf{x}_k. \tag{2.4 }$$

By the assumption $\rho \in [0, 1)$, we conclude that $\rho/(1+(n-1)\rho) \geq 0$. Hence and by Lemma 2.1, for all $k = 1, 2, \ldots, p$ we obtain the following inequality

$$\mathbf{x}_k' \mathbf{G}^{-1} \mathbf{x}_k = \frac{1}{g(1-\rho)} \left[\mathbf{x}_k' \mathbf{x}_k - \frac{(\mathbf{x}_k' \mathbf{1}_n)^2 \rho}{1+(n-1)\rho} \right] \leq \frac{\mathbf{x}_k' \mathbf{x}_k}{g(1-\rho)}. \tag{2.5}$$

It is easy to see that $\mathbf{x}_k' \mathbf{x}_k = n$, since $\mathbf{X} \in \mathrm{M}_{n \times p}(\{-1, 1\})$. Therefore, from inequalities (2.4) and (2.5) we have the inequality (2.3).

The proof is complete. ■

Definition 2.1. *A*ny chemical balance weighing design $\mathbf{X} \in \mathrm{M}_{n \times p}(\{-1, 1\})$ with covariance matrix of errors $\sigma^2 \mathbf{G}$, where $\mathbf{G}$ is given by (2.1), $\rho \in [0, 1)$, is said to be D*-optimal if it satisfies the equality in (2.3), that is

$$\det(\mathbf{X}' \mathbf{G}^{-1} \mathbf{X}) = \left(\frac{n}{g(1-\rho)} \right)^p. \tag{2.6}$$

By Theorem 2.1 it follows that D*-optimal weighing design is D-optimal. In general, the opposite implication is not true, because there are D-optimal designs which are not D*-optimal for some design parameters. The sufficient conditions under which the chemical balance weighing design is D*-optimal are given in the following theorem.

Theorem 2.2. Assume that $\mathbf{X} \in \mathrm{M}_{n \times p}(\{-1, 1\})$ is the weighing design with the covariance matrix of errors $\sigma^2 \mathbf{G}$, where $\mathbf{G}$ is given by (2.1) with $g > 0$, then the design $\mathbf{X}$ is D*-optimal if

1. $\rho = 0$ and $\mathbf{X}'\mathbf{X} = n\mathbf{I}_p$.
2. $\rho \in (0, 1)$, $\mathbf{X}'\mathbf{X} = n\mathbf{I}_p$ and $\mathbf{X}'\mathbf{1}_n = \mathbf{0}_p$.

Proof. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_p] \in \mathrm{M}_{n \times p}(\{-1, 1\})$ be the design with the covariance matrix of errors $\sigma^2 \mathbf{G}$, where the matrix $\mathbf{G}$ is given by (2.1).

1. If $\mathbf{X}'\mathbf{X} = n\mathbf{I}_p$, then $\mathbf{x}_i' \mathbf{x}_j = 0$ for $i \neq j, i, j = 1, 2, \ldots, p$. Hence and by Lemma 2.1 it follows that if $\rho = 0$, then the information matrix of the design $\mathbf{X}$ is equal to the matrix $(n/g)\mathbf{I}_p$, so its determinant satisfies the equality (2.6).

2. If $\rho \in (0,1)$ and $\mathbf{x}_k' \mathbf{1}_n = 0 \; (k = 1, 2, \ldots, p)$, then by Lemma 2.1 we deduce that the information matrix $\mathbf{X}'\mathbf{G}^{-1}\mathbf{X} = n/(g(1-\rho))\mathbf{I}_p$ and therefore the equality (2.6) is also satisfied.
Thus, the design $\mathbf{X}$ is D*-optimal in both cases. This completes the proof. ■

Masaro and Wong (2008) proved the following theorem.

Theorem 2.3. (Masaro and Wong, 2008). Let $\sigma^2\mathbf{G}$, where the matrix $\mathbf{G}$ is given by (2.1), be the covariance matrix of random errors. Moreover, we assume that
$n \equiv 0 \pmod 4$, $n \geq p+1$, $\mathbf{Z} \in D_0 = \{\mathbf{X} \in \mathrm{M}_{n\times p}(\{-1, 1\}) : \mathbf{X}'\mathbf{X} = n\mathbf{I}_p\}$ and $\mathbf{Z}'\mathbf{1}_n = \mathbf{0}_p$. Then the design $\mathbf{Z}$ is D-optimal in the subclass D_0 for all $\rho > 0$.

So, Theorems 2.1 and 2.2 expand the result of Theorem 2.3. In Masaro and Wong (2008), some construction of design $\mathbf{Z}$, which satisfies the assumptions of Theorem 2.3, is also given. We present and use this construction in the proof of Theorem 3.1.

3. Construction of D*-optimal weighing designs

In this section, we present a construction of D*-optimal weighing designs by means of Hadamard matrices. Lots of information about Hadamard matrices can be found in Hedayat and Wallis (1978). We give the construction in the proof of the following theorem.

Theorem 3.1. Let $\rho = 0$, $n \geq p$ or $\rho \in (0,1)$, $n \geq p+1$. The existence of a Hadamard matrix $\mathbf{H}_n$ of order n implies the existence of the D*-optimal design matrix $\mathbf{X} \in \mathrm{M}_{n\times p}(\{-1, 1\})$.

Proof. Let $\mathbf{H}_n \in \mathrm{M}_{n\times n}(\{-1, 1\})$ be a Hadamard matrix of order n. Then $\mathbf{H}_n'\mathbf{H}_n = n\mathbf{I}_n$. So, if $\rho = 0$, then the weighing design composed from p columns of the matrix $\mathbf{H}_n$ is D*-optimal. Now, we assume that $\rho \in (0,1)$. One can negate every row of the matrix $\mathbf{H}_n$ whose first element is -1 and thus obtaining the matrix $\overline{\mathbf{H}}_n$ with the first column of ones, but still $\overline{\mathbf{H}}_n'\overline{\mathbf{H}}_n = n\mathbf{I}_n$.

Let $\overline{\mathbf{H}}_n = [\mathbf{1}_n, \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n-1}]$. The columns of $\overline{\mathbf{H}}_n$ are orthogonal and sum of elements in each column (except the first one) of this matrix is zero because the columns $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n-1}$ are orthogonal to the first column $\mathbf{1}_n$. Therefore, $\mathbf{X}$ composed from p columns of $\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{n-1}$ is the D*-optimal design matrix by Theorem 2.2. ■

We illustrate the above construction of D*-optimal weighing design by an example.

Example 3.1. The following matrix is a Hadamard matrix of size 12×12 ($-$ denotes -1 and $+$ represents 1)

$$\mathbf{H}_{12} = \begin{bmatrix} + & - & + & + & + & + & + & + & + & + & + & + \\ + & + & - & + & - & + & - & + & - & + & - & + \\ + & + & + & - & + & + & - & - & - & - & + & + \\ - & + & + & + & - & + & + & - & + & - & - & + \\ + & + & + & + & + & - & + & + & - & - & - & - \\ + & - & + & - & - & - & + & - & - & + & - & + \\ - & - & + & + & - & - & - & + & - & - & + & + \\ + & - & - & + & + & - & - & - & + & - & - & + \\ + & + & - & - & - & - & + & + & + & - & + & + \\ + & - & - & + & - & + & + & - & - & - & + & - \\ + & + & + & + & - & - & - & - & + & + & + & - \\ + & - & + & - & - & + & - & + & + & - & - & - \end{bmatrix}.$$

The first element of the fourth and the seventh row of the matrix $\mathbf{H}_{12}$ is equal to -1. Therefore, we negate these rows and we get the following matrix:

$$\overline{\mathbf{H}}_{12} = \begin{bmatrix} + & - & + & + & + & + & + & + & + & + & + & + \\ + & + & - & + & - & + & - & + & - & + & - & + \\ + & + & + & - & + & + & - & - & - & - & + & + \\ + & - & - & - & + & - & - & + & - & + & + & - \\ + & + & + & + & + & - & + & + & - & - & - & - \\ + & - & + & - & - & - & + & - & - & + & - & + \\ + & + & - & - & + & + & + & - & + & + & - & - \\ + & - & - & + & + & - & - & - & + & - & - & + \\ + & + & - & - & - & - & + & + & + & - & + & + \\ + & - & - & + & - & + & + & - & - & - & + & - \\ + & + & + & + & - & - & - & - & + & + & + & - \\ + & - & + & - & - & + & - & + & + & - & - & - \end{bmatrix}.$$

Therefore, we construct D*-optimal design from p columns of $\overline{\mathbf{H}}_{12}$ (from the second column to the last). For example, the design with the matrix

$$\mathbf{X}' = \begin{bmatrix} - & + & + & - & + & - & + & - & + & - & + & - \\ + & - & + & - & + & + & - & - & - & - & + & + \\ + & + & - & - & + & - & - & + & - & + & + & - \\ + & - & + & + & + & - & + & + & - & - & - & - \\ + & + & + & - & - & - & + & - & - & + & - & + \end{bmatrix}$$

is D*-optimal weighing design for $p = 5$ objects.

4. Application

In this section, we present an example of the application of the optimal chemical balance weighing designs in the bioengineering experiment given in the literature. Namely, Gawande and Patkar (1999) described two experiments in which the effect of dextrin, peptone, yeast extract, ammonium dihydogen orthophosphate $(NH_4H_2PO_4)$, and magnesium sulfate $(MgSO_4 \cdot 7H_2O)$ on production of cyclodextrin glycosyltransferase from *Klebsiella pneumoniae pneumoniae* AS-22 was investigated in two experiments for 2^{5-1} fractional factorial designs. One of the designs and the observations are given in Table 1. 2^{5-1} fractional factorial designs were used for the optimization of medium composition for production of mentioned enzyme. The optimization medium resulted in 9-fold higher production as compared to that in the basal medium.

The matrices of the plan of factorial designs or fraction factorial designs are chosen from the set $\mathrm{M}_{n\times p}(\{-1, 1\})$, which is the set of the weighing designs. If we can not use the 2^k factorial design, then we choose some 2^{k-1} fractional factorial design. Of course, we would like to have a good fractional factorial design. To do this, we can take as a matrix of the plan of the 2^{k-1} fractional factorial design the design matrix of D-optimal weighing design. It is easy to calculate that Gawande and Patkar (1999) used the 2^{5-1} fractional factorial designs which matrices of the plan are D^*-optimal even when the errors are correlated with $\rho > 0$.

Table 1. Experimental results of one of fractional factorial designs from Gawande and Patkar (1999)

Flask no.	Dextrin	Peptone	Yeast extract	$NH_4H_2PO_4$	$MgSO_4 \cdot 7H_2O$	Activity (U/ml)
1	–	–	–	–	–	3.97
2	–	–	–	+	+	5.99
3	–	–	+	–	+	4.13
4	–	–	+	+	–	5.59
5	–	+	–	–	+	5.18
6	–	+	–	+	–	6.47
7	–	+	+	–	–	5.12
8	–	+	+	+	+	6.53
9	+	–	–	–	–	5.39
10	+	–	–	+	+	5.25
11	+	–	+	–	+	5.39
12	+	–	+	+	–	6.06
13	+	+	–	–	+	4.98
14	+	+	–	+	–	6.74
15	+	+	+	–	–	5.66
16	+	+	+	+	+	8.42

References

Galil Z., Kiefer J. (1980). D-optimum weighing designs. *Ann. Statist*. 8, 1293–1306.

Gawande B.N., Patkar A.Y. (1999). Application of factorial designs for optimization of cyclodextrin glycosyltransferase production from *Klebsiella pneumoniae pneumoniae* AS-22. *Biotechnology and Bioengineering* 64, 168–173.

Hedayat A., Wallis W.D. (1978). Hadamard matrices and their application. *Annals of Statistics* 6, 1184–1238.

Jacroux M., Wong C.S., Masaro J.C. (1983). On the optimality of chemical balance weighing design. *Journal of Statistical Planning and Inference* 8, 231–240.

Katulska K., Smaga Ł. (2012). D-optimal chemical balance weighing designs with $n \equiv 0 \pmod 4$ and 3 objects. *Communications in Statistics – Theory and Methods* 41, 2445–2455.

Katulska K., Smaga Ł. (2013). D-optimal chemical balance weighing designs with autoregressive errors. *Metrika* 76, 393–407.

Li C.H., Yang S.Y. (2005). On a conjecture in D-optimal designs with $n \equiv 0 \pmod 4$. *Linear Algebra and its Applications* 400, 279–290.

Masaro J.C., Wong C.S. (2008). D-optimal designs for correlated random vectors. *Journal of Statistical Planning and Inference* 138, 4093–4106.

Yeh H.G., Lo Huang M.N. (2005). On exact D-optimal designs with 2 two-level factors and *n* autocorrelated observations. *Metrika* 61, 261–275.