PRACE NAUKOWE AKADEMII EKONOMICZNEJ WE WROCŁAWIU
Nr 1162 — 2007
Applications of Mathematics and Statistics in Economics

Łukasz Kuźmiński

Wrocław University of Economics, Poland

COST-OPTIMAL EXPERIMENTAL DESIGN – OUTLINE OF PROBLEM AND SOME SOLUTIONS APPROACHES

1. Introduction

In the classic theory of optimal designs methods of seeking of formal (discrete) and real (exact) optimal designs are known. These methods are based on the known criterions such as A-, C-, D-, E- G- and V-optimality. The theory of optimum experimental design allows to obtain designs that have the greatest possible statistic precision of the parameters estimate of the linear regression function

$$Y(\mathbf{x}) = h(\mathbf{x},\boldsymbol{\theta}) + \varepsilon(\mathbf{x}), \tag{1}$$

where $h(\mathbf{x},\boldsymbol{\theta})$ is the so called response surface the form

$$h(\mathbf{x},\boldsymbol{\theta}) = \theta_1 f_1(\mathbf{x}) + \ldots + \theta_r f_r(\mathbf{x}). \tag{2}$$

The total cost of experiment depends on the same factors as the precision of the experimental design. These factors are: number of observations and the allocation of observations in the experimental area.

In the theory of cost-optimal designs two cases are differentiated:

1) cost of experiment depends only on the number of observations, the complete cost of experiment grows proportionally to the number of observations,

2) cost of experiment is additionally dependent on allocation of points in the experimental area in which the observations are done.

In the first case the cost-optimal design is the same as the optimal design with the additional condition

$$nK \leqslant k_0, \tag{3}$$

where n – number of observations, K – cost of one observation, k_0 – total cost outlays.

In the second case this problem is more difficult, because the cost of one observation is not the same in each point of the experimental area. The total cost of experiment P_n, in this case, is the sum of costs $k(\mathbf{x}_i)$ ($i = 1, \ldots, n$) of the individual observations

$$K(P_n) = \sum_{i=1}^{n} k(\mathbf{x}_i). \tag{4}$$

In this article only the second case is researched.

2. The various approaches for the cost-optimal designs construction

The cost-optimal designs can be obtained according to three various approaches. In two of these approaches the real cost-optimal design can be find whereas in one of these approaches only the formal cost-optimal design can be established.

Approach 1

We are looking for $P^*_{n_0,\psi_0} \in \mathcal{P}_0$, where $\mathcal{P}_0 = \bigcup_{n=n_{min}}^{n=n_{max}} P_n$ such that total cost $K(P^*_{n_0,\psi_0}) = \inf_{P_n \in \mathcal{P}_0} \{K(P_n) | \psi(P_n) \leqslant \psi_0\}$, where ψ_0 is the assumed admissible value of the function ψ which defines the precision in the sense of the criterion of ψ-optimality. In other words, we are looking for the design $P^*_{n_0,\psi_0}$ that has, at lowest cost, the efficiency at least e_{ψ_0}, where $e_{\psi_0} = \dfrac{\psi(P^*_{n,\psi})}{\psi_0}$, and $P^*_{n,\psi}$ is the ψ-optimal design in $\mathcal{P}_0$.

Approach 2

In this case, we are looking for the design $P^*_{n_0,\psi} \in \mathcal{P}_0$, where $\mathcal{P}_0 = \bigcup_{n=n_{min}}^{n=n_{max}} P_n$ such that $\psi(P^*_{n_0,\psi}) = \inf_{P_n \in \mathcal{P}_0} \{\psi(P_n) | K(P_n) \leqslant k_0\}$. By this $\psi(P_n)$ is the value of optimality criterion function ψ for designs $P_n \in \mathcal{P}_0$. This approach lets seek a design of experiment at the highest efficiency e_ψ that fulfils the limit of costs given in the form $\sum_{i=1}^{n} k(x_i) \leqslant k_0$.

Approach 3

This approach lets seek the formal cost-optimal design. We assume, as before, that the potential observations are of different costs, say $k(\mathbf{x}_i)$, $(i=1,\ldots,n)$. If the total cost has to be lower or equal to a prescribed amount k_0, we have the restriction $\sum_{i=1}^{n} k(\mathbf{x}_i) \leqslant k_0$.

The regression equation has the form (1) and the response surface is given by (2). We connect the cost of observation in every given allocation point in the experimental area by modifying the variance $Vh(\mathbf{x})$ in the given point $\mathbf{x}$ to $\sqrt{k(\mathbf{x})}Vh(\mathbf{x})$. After this modification the total cost of experiment depends on the value of observations variance in this point. This leads to the so called heteroscedastic model for the regression function (1) where the covariance matrix $\Sigma(P_n)$ of the observation vector following from $P_n = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$ has now the form

$$\Sigma(P_n) = \mathrm{diag}((\sqrt{k(\mathbf{x}_1)}, \ldots, \sqrt{k(\mathbf{x}_n)})^T)\sigma^2. \tag{5}$$

It can be shown [3] that a formal cost-optimal design $\xi_\psi = \{\mathbf{x}_j, p_j\}_{j=1}^{s}$ for estimation the parameters of (2) can be obtained from the formal ψ-optimal design $\tilde{\xi}_\psi = \{\mathbf{x}_j, q_j\}_{j=1}^{s}$ for the homoscedastic model with the response surface

$$\tilde{h}(\mathbf{x},\boldsymbol{\theta}) = \theta_1 \frac{f_1(\mathbf{x})}{\sqrt{k(\mathbf{x})}} + \ldots + \theta_r \frac{f_r(\mathbf{x})}{\sqrt{k(\mathbf{x})}} \tag{6}$$

and variance $\sigma^2(\mathbf{x}) = \mathrm{const}$, where

$$p_j = \frac{q_j}{K(\mathbf{x}_j)\sum_{l=1}^{s}\frac{q_l}{K(\mathbf{x}_l)}}. \tag{7}$$

3. The D-optimum criterion

In the further part of this article the three approaches are illustrated on numerical examples. In these examples the D-optimum criterion is used. We define this criterion for the class of real designs and the class of formal designs.

Definition 1

The real design $P_{n_0,D} \in \mathcal{P}_0$ is D-optimum design in the class $\mathcal{P}_0$, if

$$\frac{1}{n_0^r}\det M^{-1}(P_{n_0,D}) = \inf_{P_n\in\mathcal{P}_0} \frac{1}{n^r}\det M^{-1}(P_n), \tag{8}$$

where $M(P_n)$ is the matrix design given in the form

$$\mathbf{M}(P_n)\frac{1}{n}\mathbf{F}_{\mathbf{x}}^{\mathrm{T}}\mathbf{F}_{\mathbf{x}}, \tag{9}$$

and $\mathbf{F_X} = [\mathbf{f}(\mathbf{x}^{(1)})...\mathbf{f}(\mathbf{x}^{(n)})]^{\mathrm{T}}$.

From definition 1 follows that the D-optimum design minimizes the generalized variance of the OLS estimators of $\theta_1, ..., \theta_r$.

Definition 2

The formal design $\xi_D \in \Xi_0$ is D-optimal formal design in class Ξ_0, if

$$\det M^{-1}(\xi_D) = \inf_{\xi\in\Xi_0} \det M^{-1}(\xi), \tag{10}$$

where Ξ_0 is some given class of the formal designs such that $\mathcal{P}_0 \in \Xi_0$ and $\det M(\xi) > 0$. The members of the class Ξ_0 are designs independent explicitly on n.

4. The numerical examples of an application of the showing approaches

Example 1

In this example the procedure of seeking a cost-optimal design which satisfies some precision condition related to D-optimum criterion in according to the approach 1 is shown. Let the linear regression function of two variables x_1 and x_2 be given in the form

$$Y(\mathbf{x};\boldsymbol{\theta}) = \theta_1 x_1 + \theta_2 x_2 + \varepsilon(x), \tag{11}$$

whereas the cost function of the design P_n is given by the equation

$$K(P_n) = \sum_{i=1}^{n} x_{1i}^2 \sum_{i=1}^{n} x_{2i}^2. \tag{12}$$

We assume that the experimental area D_1 is the unit square, i.e.

$$D_1 = \{\mathbf{x} = (x_1, x_2) : -1 \leqslant x_j \leqslant 1,\ x_j \neq 0\ \ i = 1,2\}. \tag{13}$$

The information matrix of the design P_n for the equation (11) has the form

$$\mathbf{M}(P_n) = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_{1i}^2 & \frac{1}{n}\sum_{i=1}^{n} x_{1i}x_{2i} \\ \frac{1}{n}\sum_{i=1}^{n} x_{1i}x_{2i} & \frac{1}{n}\sum_{i=1}^{n} x_{1i}^2 \end{bmatrix} \tag{14}$$

and

$$\det \mathbf{M}(P_n) = \frac{1}{n^2}\sum_{i=1}^{n} x_{1i}^2 \sum_{i=1}^{n} x_{2i}^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_{1i}x_{2i}\right)^2. \tag{15}$$

We are looking for the design $P^*_{n_0,\psi_0}$ for which

$$\det \mathbf{M}^{-1}(P^*_{n_0,\psi_0}) \leqslant 2 \;\text{ that is }\; \det \mathbf{M}(P^*_{n_0,\psi_0}) \geqslant 0.5 \tag{16}$$

and

$$K(P^*_{n,\psi_0}) = \inf_{P_{n,\psi} \in \mathcal{P}_0} \left\{ K(P_n) \middle| \psi(P_n) \leqslant \psi_0 \right\}, \tag{17}$$

where $\psi_0 = 2$ and $\det \mathbf{M}^{-1}(P_n)$ is the D-optimum criterion (see Definition 1).

From (15) and (16) we obtain

$$\frac{1}{n^2}\sum_{i=1}^{n} x_{1i}^2 \sum_{i=1}^{n} x_{2i}^2 - \left(\frac{1}{n}\sum_{i=1}^{n} x_{1i}x_{2i}\right)^2 \geqslant 0.5\,. \tag{18}$$

Because of (12) we get

$$K(P_n) \geq n^2 0.5 + \left(\sum_{i=1}^{n} x_{1i}x_{2i}\right)^2 \;\text{ for every } P_n \in \mathcal{P}_0 \text{ satisfying (16).} \tag{19}$$

From (19) it follows, that the minimum cost $K(P_n)$ of design, which fulfils the condition (16) of the precision of P_n, is obtained for $n = 2$ observations which fulfils the equation

$$x_{11}x_{21} + x_{12}x_{22} = 0\,. \tag{20}$$

In this case $K(P_n) = K(P_2) = 2$ and from (12) follows that

$$(x_{11}^2 + x_{12}^2)(x_{21}^2 + x_{22}^2) = 2\,. \tag{21}$$

The solution of (20) and (21) is given for example by

$$x_{11} = 0.644 \quad x_{21} = 1$$
$$x_{12} = 1 \quad x_{22} = -0.644.$$

The resulting design is therefore as follows:

$$P^*_{2.\psi_0} = \{(0.644\,,1) \quad (1\,,-0.644)\}. \tag{22}$$

The design $P^*_{2.\psi_0}$ is the cost-optimal design which fulfils the condition (17).

Example 2

In this example the procedure of seeking the D-cost-optimal design according to the approach 2 is shown. The linear regression function has the form

$$Y(x;\boldsymbol{\theta}) = \theta_1 + \theta_2 x^2 + \varepsilon(x)\,. \tag{23}$$

The cost function of the design P_n is given by the equation

$$K(P_n) = \sum_{i=1}^{n}(x_i + 2). \tag{24}$$

We assume that the experimental area $D_1 = [-1,1]$.

In the first step of this procedure we are seeking for a design $P_1^* = \{x\}$ that fulfils the following conditions

$$\det \mathbf{M}(P_1^*) = \sup_{P_1 \in \mathcal{P}_1} \det \mathbf{M}(P_1) \;\text{ and }\; K(P_1^*) \leqslant k_0, \tag{25}$$

where $\mathcal{P}_1$ denotes a class of experimental designs for $n = 1$ observations and k_0 denotes the total limit-cost of experiment. In this example $k_0 = 4.9$.

Because

$$\mathbf{M}(P_1) = \begin{bmatrix} 1 & x^2 \\ x^2 & x^4 \end{bmatrix} \;\text{ and }\; \det \mathbf{M}(P_1) = 0, \tag{26}$$

we obtain that $\det \mathbf{M}(P_1)$ is independent from $x \in [-1,1]$. Additionally (comp. (24)) $K(P_1) = x + 2$ and $\min_{x \in [-1,1]} K(P_1) = \min_{x \in [-1,1]} (x+2) = 1$ for $x = -1$. From this we obtain immediately that $P_1^* = \{-1\}$.

Because

$$K(P_1^*) = 1 \leqslant k_0,$$

we are starting the second step of the procedure. In this step we are seeking for a design $P_2^* = \{-1,\, x\}$ that fulfils the similar conditions as the design in the first step, namely

$$\det \mathbf{M}(P_2^*) = \sup_{P_2 \in \mathcal{P}_2} \det \mathbf{M}(P_2) \;\text{ and }\; K(P_2^*) \leqslant k_0\,, \tag{27}$$

where $\mathcal{P}_2$ denotes a class of designs for $n = 2$ observations with the first observation in point −1.

For such designs we have

$$\mathbf{M}(P_2) = \begin{pmatrix} 1 & \frac{1+x^2}{2} \\ \frac{1+x^2}{2} & \frac{1+x^4}{2} \end{pmatrix} \text{ and } \det \mathbf{M}(P_2) = \frac{x^4}{4} - \frac{x^2}{2} + \frac{1}{4} = \frac{(x^2-1)^2}{4}. \quad (28)$$

Hence $\det \mathbf{M}(P_2^*) = \sup\limits_{P_2 \in \mathcal{P}_2} \frac{(x^2-1)^2}{4} = \frac{1}{4}$ for $x = 0$ and $P_2^* = \{-1, 0\}$. Therefore $K(P_2^*) = 3 \leqslant k_0$ and we are going to the next step.

In this step we are seeking the design $P_3^* = \{-1, 0, x\}$ that fulfils the conditions from p

$$\det \mathbf{M}(P_3^*) = \sup_{P_3 \in \mathcal{P}_3} \det \mathbf{M}(P_3) \text{ and } K(P_3^*) \leqslant k_0. \quad (29)$$

For designs $P_3 \in \mathcal{P}_3$ we have

$$\mathbf{M}(P_3) = \begin{pmatrix} 1 & \frac{1+x^2}{3} \\ \frac{1+x^2}{3} & \frac{1+x^4}{3} \end{pmatrix} \text{ and } \det \mathbf{M}(P_3) = \frac{2}{9}x^4 - \frac{2}{9}x^2 + \frac{2}{9}. \quad (30)$$

Hence $\det \mathbf{M}(P_3^*) = \sup\limits_{P_3 \in \mathcal{P}_3} \frac{2}{9}(x^4 - x^2 + 1) = \frac{2}{9}$ i.e.

$$\text{for } {}_1P_3 = \{-1,-1,0\}, \text{ for } {}_2P_3 = \{-1,0,0\} \text{ and for } {}_3P_3 = \{-1,0,1\}.$$

Because $K({}_1P_3) = 4$, $K({}_2P_3) = 5 > k_0$ and $K({}_3P_3) = 6$ so

$$P_3^* = \{-1,-1,0\}.$$

In this place we have to stop this procedure because the cost function for every $P_4 \in \mathcal{P}_4$ must follow the inequality $K(P_4) \geqslant 5 > k_0 = 4.9$.

Let us notice that the D-effectiveness P_3^* is rather poor, because from the formula

$$e_D(P_n) = \exp[2 - \sup_{x \in [-1;1]} \mathbf{f}^T(\mathbf{x})\mathbf{M}^{-1}(P_n)\mathbf{f}(\mathbf{x})] \quad (31)$$

it is easy to obtain for P_3^* that $e_D(P_3^*) = 0.368$.

Example 3

In this example the procedure of assigning D-optimal design is shown, for the regression equation

$$Y(x;\boldsymbol{\theta}) = \theta_1 + \theta_2 x + \varepsilon(x)\,, \tag{32}$$

and $V(\varepsilon_i) = \sigma_i^2 = (x_i^2 + 1)\sigma^2$, for $i = 1, ..., n$. It means that the regression model is heteroscedastic. This case is described in the approach 3.

We assume the experimental area is the interval $D_1 = [-1\,, 1]$. From the assumption about $V(\varepsilon_i)$ it follows that the cost function has the form

$$k(x) = x^2 + 1\,. \tag{33}$$

We transform equation (11) into

$$\tilde{Y}(x;\boldsymbol{\theta}) = \frac{\theta_1}{\sqrt{x^2+1}} + \frac{\theta_2 x}{\sqrt{x^2+1}} + \tilde{\varepsilon}(x) \text{ where } \tilde{\varepsilon}(x) = \frac{\varepsilon(x)}{\sqrt{x^2+1}}, \tag{34}$$

therefore $V(\varepsilon_i) = \sigma^2$ for $i = 1, ..., n$. The regression model (34) is homoscedastic.

We are starting the iterative procedure from the zero step in which we assign some formal design on D_1, e.g. the design

$$\tilde{\xi}_0 = \begin{Bmatrix} 0 & 0{,}5 \\ \dfrac{2}{3} & \dfrac{1}{3} \end{Bmatrix}, \tag{35}$$

where $x_1 = 0,\ x_2 = 0.5$.

We will stop the procedure if in the k-th step the design $\tilde{\xi}_k$ attains the D-effectiveness at least 0.95, i.e. $e_D(\xi_k) \geq 0.95$.

The D-effectiveness can be calculated using the formula

$$e_D = \exp[r - \sup_{\mathbf{x}\in D_1} \mathbf{f}^T(\mathbf{x})\mathbf{M}^{-1}(\xi)\mathbf{f}(\mathbf{x})]\,, \tag{36}$$

where r is the number of parameters in equation (6), $\mathbf{M}^{-1}(\xi)$ denotes the inverse information matrix of design ξ, $\mathbf{f}(\mathbf{x}) = [f_1(\mathbf{x})...f_r(\mathbf{x})]$ (comp. [7, p. 100]).

In our example we obtain

$$e_D = \exp[2 - \sup_{x\in[-1;1]} \begin{bmatrix} \dfrac{1}{\sqrt{x^2+1}} & \dfrac{x}{\sqrt{x^2+1}} \end{bmatrix} \begin{bmatrix} \dfrac{2}{3} + \dfrac{1}{3}\cdot\dfrac{1}{1.25} & \dfrac{1}{3}\cdot\dfrac{0.5}{1.25} \\ \dfrac{1}{3}\cdot\dfrac{0.5}{1.25} & \dfrac{1}{3}\cdot\dfrac{0.25}{1.25} \end{bmatrix}^{-1} \begin{bmatrix} \dfrac{1}{\sqrt{x^2+1}} \\ \dfrac{x}{\sqrt{x^2+1}} \end{bmatrix}] =$$

$$= \exp[2 - \sup_{x\in[-1;1]} \left[\frac{1}{\sqrt{x^2+1}}, \frac{x}{\sqrt{x^2+1}} \right] \begin{bmatrix} 1.49 & -2.985 \\ -2.985 & 21.12 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{x^2+1}} \\ \frac{x}{\sqrt{x^2+1}} \end{bmatrix}] =$$

$$= \exp[2 - \sup_{x\in[-1;1]} \left[\frac{1}{x^2+1}(21.12x^2 - 5.97x + 1.49) \right]] = 4.6\cdot 10^{-6} < 0.95. \qquad (37)$$

Because $e_D < 0.95$, we are going to the first step of the iterative procedure. This iterative procedure is described in [7, p 107]. In this step we are building a family

$$\tilde{\Xi}(1) = \left\{ \tilde{\xi} \in \tilde{\Xi}_0 : S(\tilde{\xi}\,) \subseteq S(\tilde{\xi}_0) \cup \{x\} : \ x \in [-1,1] \right\} \qquad (38)$$

of formal designs on D_1, where $S(\tilde{\xi}\,)$ is the spectrum of design $\tilde{\xi}$.

Each design $\tilde{\xi} \in \tilde{\Xi}(1)$ has the form

$$\tilde{\xi} = \begin{Bmatrix} 0 & 0.5 & x \\ q_1 & q_2 & q_3 \end{Bmatrix}, \qquad (39)$$

where $q_1 + q_2 + q_3 = 1$. The information matrix design of every design $\tilde{\xi}$ has the form

$$\mathbf{M}(\tilde{\xi}) = \begin{bmatrix} q_1 + 0.8q_2 + q_3\dfrac{1}{x^2+1} & 0.4q_2 + q_3\dfrac{x}{x^2+1} \\ 0.4q_2 + q_3\dfrac{x}{x^2+1} & 0.20q_2 + q_3\dfrac{x^2}{x^2+1} \end{bmatrix}, \qquad (40)$$

whereas $\det \mathbf{M}(\tilde{\xi}) = q_3(q_1 - 0.8q_2)\dfrac{x^2}{x^2+1} + \dfrac{q_2q_3}{x^2+1}(0.2 - 0.8x) + 0.2q_1q_2$. The determinant of $\mathbf{M}(\tilde{\xi})$ attains maximum (the function $\psi_D(\mathbf{M}(\tilde{\xi})) = \det \mathbf{M}^{-1}(\tilde{\xi})$ attains minimum) for $q_1 = 0; q_2 = 0.5; q_3 = 0.5; x = -1$.

In the class $\tilde{\Xi}(1)$ the design

$$\tilde{\xi}_1 = \begin{Bmatrix} -1 & 0.5 \\ 0.5 & 0.5 \end{Bmatrix} \qquad (41)$$

should be chosen, for which

$$\mathbf{M}(\tilde{\xi}_1) = \begin{bmatrix} 0.65 & -0.05 \\ -0.05 & 0.35 \end{bmatrix} \qquad (42)$$

and

$$\mathbf{M}^{-1}(\tilde{\xi}_1)=\begin{bmatrix} 1.556 & 0.222 \\ 0.222 & 2.889 \end{bmatrix}. \tag{43}$$

Calculating the D-effectiveness from the equation (36), we obtain

$$e_D=\exp[2-\sup_{x\in[-1;1]}\left[\frac{1}{\sqrt{x^2+1}},\frac{x}{\sqrt{x^2+1}}\right]\begin{bmatrix} 1.556 & 0.222 \\ 0.222 & 2.889 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{x^2+1}} \\ \frac{x}{\sqrt{x^2+1}} \end{bmatrix}]=$$

$$=\exp[2-\sup_{x\in[-1;1]}\left[\frac{1}{x^2+1}(2.889x^2+0.444x+1.556)\right]]=0.64<0.95. \tag{44}$$

The D-effectiveness of the design $\tilde{\xi}_1$ is lower than 0.95, therefore we are starting with the second step of the iterative procedure. In this step we construct the set

$$\tilde{\Xi}(2)=\left\{\tilde{\xi}\in\tilde{\Xi}_0 : S(\tilde{\xi}\ )\subseteq S(\tilde{\xi}_1)\cup\{x\}:\ x\in[-1,1]\right\}. \tag{45}$$

Each design $\tilde{\xi}\in\tilde{\Xi}(2)$ has the form

$$\tilde{\xi}=\begin{Bmatrix} -1 & 0.5 & x \\ q_1 & q_2 & q_3 \end{Bmatrix}. \tag{46}$$

The information matrix design of the every design $\tilde{\xi}\in\tilde{\Xi}(2)$ has the form

$$\mathbf{M}(\tilde{\xi})=\begin{bmatrix} 0.5q_1+0.8q_2+q_3\frac{1}{x^2+1} & -0.5q_1+0.4q_2+q_3\frac{x}{x^2+1} \\ -0.5q_1+0.4q_2+q_3\frac{x}{x^2+1} & 0.5q_1+0.2q_2+q_3\frac{x^2}{x^2+1} \end{bmatrix}. \tag{47}$$

The determinant of $\mathbf{M}(\tilde{\xi})$ attains maximum (the function $\psi_D(\mathbf{M}(\tilde{\xi}))=\det\mathbf{M}^{-1}(\tilde{\xi})$ attains minimum) for $q_1=0.5; q_2=0; q_3=0.5; x=1$.

In the class $\tilde{\Xi}(2)$ the design

$$\tilde{\xi}_2=\begin{Bmatrix} -1 & 1 \\ 0.5 & 0.5 \end{Bmatrix} \tag{48}$$

should be chosen for which

$$\mathbf{M}(\tilde{\xi}_2) = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix} \tag{49}$$

and

$$\mathbf{M}^{-1}(\tilde{\xi}_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}. \tag{50}$$

If we calculate the D-effectiveness from the equation (36), we will obtain

$$e_D = \exp[2 - \sup_{x\in[-1;1]} \left[\frac{1}{\sqrt{x^2+1}}, \frac{x}{\sqrt{x^2+1}} \right] \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{x^2+1}} \\ \frac{x}{\sqrt{x^2+1}} \end{bmatrix}] =$$

$$= \exp[2 - \sup_{x\in[-1;1]} \left[\frac{1}{x^2+1}(2x^2+2) \right]] = 1 \geqslant 0.95.$$

In this place we stop the iterative procedure, because the design $\tilde{\xi}_2$ attains the D-effectiveness higher than 0.95.

Additionally, the $\tilde{\xi}_2$ design is the D-optimum design for the equation (13), because the D-effective equals 1.

Finally, we are seeking for the D-cost-optimal design for the equation (11) by solving the equation (7). In our case

$$\begin{aligned} p_1 &= \frac{q_1}{K(x_1)\sum_{i=1}^{s}\frac{q_i}{K(x_i)}} = \frac{0.5}{2(\frac{0.5}{2}+\frac{0.5}{2})} = 0.5, \\ p_2 &= \frac{q_2}{K(x_2)\sum_{i=1}^{s}\frac{q_i}{K(x_i)}} = \frac{0.5}{2(\frac{0.5}{2}+\frac{0.5}{2})} = 0.5, \end{aligned} \tag{51}$$

therefore D-cost-optimal design has the form

$$\xi_2 = \begin{Bmatrix} -1 & 1 \\ 0.5 & 0.5 \end{Bmatrix}. \tag{52}$$

5. Summary

In this article the methods of seeking the formal and real cost-optimal design are described. Three approaches to the solution of this problem are shown. Each approach is shown on the numerical examples. Two approaches involve the real cost-optimal designs and one the formal cost-optimal design. In presenting examples the D-optimum criterion is used.

References

[1] Bandamer H., Nather W., *Theorie und Anwendung der optimalen Versuchsplanung II*, Akademie-Verlag, Berlin 1980.

[2] Chernoff H., "Locally Optimum Designs for Estimating Parameters", *The Annals of Mathematical Statistics* 1953, 24 (4), 586-602.

[3] Elfving G., "Optimum Allocation in Linear Regression Theory", *The Annals of Mathematical Statistic* 1952, 23 (2), 255-262.

[4] Jung W., "Kostenoptimale Versuchsplanung im Regressionsmodell", *Biometrica* 1974, 16, 31-47.

[5] Kiefer J., "Optimum Experimental Design", *Journal of the Royal Statistical Society* 1959, 21 (2), 272-319.

[6] Kiefer J., Wolfowitz J., "Optimum Design in Regression Problems", *The Annals of Mathematical Statistics* 1959, 30 (2), 271-294.

[7] Wawrzynek J., *Statistical Experimental Design in Regression Problem under Small Sample Conditions*, AE, Wrocław 1994.

KOSZTOWO-OPTYMALNE PLANOWANIE EKSPERYMENTÓW – ZARYS PROBLEMU I PEWNE PODEJŚCIA DO ROZWIĄZANIA

Streszczenie

W artykule opisano metody poszukiwania formalnego i rzeczywistego planu kosztowo-optymalnego. Zostały przedstawione trzy podejścia do rozwiązania problemu. Każde podejście jest pokazane za pomocą przykładów liczbowych. Dwa podejścia dotyczą rzeczywistego planu kosztowo-optymalnego, a jedno planu formalnego. W przedstawionych przykładach wykorzystano kryterium D-optymalności.

Słowa kluczowe: plan kosztowo-optymalny, kryterium D-optymalności, model heteroscedastyczny, funkcja kosztu.