УДК 519.4

ТОЧНЫЕ *D*-ОПТИМАЛЬНЫЕ ПЛАНЫ ЭКСПЕРИМЕНТОВ ДЛЯ ЛИНЕЙНОЙ МНОЖЕСТВЕННОЙ РЕГРЕССИИ С НЕРАВНОТОЧНЫМИ НАБЛЮДЕНИЯМИ

В. П. КИРЛИЦА[1)]

[1)]*Белорусский государственный университет, пр. Независимости, 4, 220030, г. Минск, Беларусь*

Исследована проблема построения точных *D*-оптимальных планов экспериментов для линейной множественной регрессии в случае, когда дисперсия ошибок наблюдений зависит от точки наблюдения. Выделен класс функций, описывающих изменение дисперсии неравноточных наблюдений, для которых точные *D*-оптимальные планы экспериментов совпадают с точными *D*-оптимальными планами для равноточных наблюдений. Построены насыщенные *D*-оптимальные планы экспериментов для линейной множественной регрессии с тремя факторами для неравноточных наблюдений. Показано, что таких насыщенных оптимальных планов можно строить бесконечное несчетное множество.

Ключевые слова: точные *D*-оптимальные планы экспериментов; линейная множественная регрессия; равноточные наблюдения; неравноточные наблюдения; насыщенные оптимальные планы.

Образец цитирования:
Кирлица В. П. Точные *D*-оптимальные планы экспериментов для линейной множественной регрессии с неравноточными наблюдениями // Журн. Белорус. гос. ун-та. Математика. Информатика. 2017. № 3. С. 53–59.

For citation:
Kirlitsa V. P. Exact *D*-optimal designs of experiments for linear multiple regression with heteroscedastic observations. *J. Belarus. State Univ. Math. Inform.* 2017. No. 3. P. 53–59 (in Russ.).

Автор:
Валерий Петрович Кирлица – кандидат физико-математических наук, доцент; доцент кафедры математического моделирования и анализа данных факультета прикладной математики и информатики.

Author:
Valery P. Kirlitsa, PhD (physics and mathematics), docent; associate professor at the department of mathematical analysis and date analysis, faculty of applied mathematics and computer science.
kirlitsa@bsu.by

EXACT *D*-OPTIMAL DESIGNS OF EXPERIMENTS FOR LINEAR MULTIPLE REGRESSION WITH HETEROSCEDASTIC OBSERVATIONS

V. P. KIRLITSA[a]

[a]*Belarusian State University, 4 Niezaliežnasci Avenue, Minsk 220030, Belarus*

In article the problem of construction of exact *D*-optimal designs of experiments for linear multiple regression in a case when variance of errors of observations depend on a point in which is made is investigated. The class functions describing change of variance of heteroscedastic observations is constructed for which exact *D*-optimal designs coincides with exact *D*-optimal designs for homoscedastic observations. The saturated *D*-optimal designs of experiments for linear multiple regression with three factors for heteroscedastic observations are constructed. It is demonstrated that is possible to construct infinite incalculable set of such saturated optimal designs.

Key words: exact *D*-optimal designs of experiments; linear multiple regression; homoscedastic; heteroscedastic observations; saturated optimal designs.

Рассмотрим линейную модель множественной регрессии

$$y_j = \theta_1 x_{j1} + \ldots + \theta_m x_{jm} + \varepsilon\left(x^{(j)}\right),\ \ j = \overline{1,\, n},\ \ n \ge m, \tag{1}$$

где y_j – наблюдаемые переменные; $x^{(j)} = \left(x_{j1},\, \ldots,\, x_{jm}\right)$ – m-векторы контролируемых переменных, компоненты которых принадлежат единичному m-мерному кубу: $\left|x_i\right| \le 1,\ i = \overline{1,\, m}$; $\theta_1,\, \ldots,\, \theta_m$ – неизвестные параметры; $\varepsilon\left(x^{(j)}\right)$ – некоррелированные случайные ошибки наблюдений с нулевыми математическими ожиданиями и дисперсиями, зависящими от точки наблюдения $x^{(j)}$, удовлетворяющие неравенствам

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) \ge a_0 + a_1 x_{j1} + \ldots + a_m x_{jm} > 0, \tag{2}$$

для каждой реализации $x^{(j)},\ \ j = \overline{1,\, n}$. Функция $d\left(x^{(j)}\right)$ в (2) должна быть такой, что в вершинах единичного m-мерного куба неравенство (2) обращается в равенство. Константы $a_0,\, a_1,\, \ldots,\, a_m$ в (2) такие, что $a_0 > 0,\, \left|a_1\right| + \ldots + \left|a_m\right| < a_0$.

Класс функций $d(x)$, описываемых неравенством (2), довольно обширен. Этому классу принадлежат постоянные функции $d(x) = a_0,\ a_0 > 0$ (равноточные наблюдения); функции с линейным изменением – $d(x) = a_0 + a_1 x_1 + \ldots + a_m x_m$; вогнутые функции, удовлетворяющие (2).

Для равноточных наблюдений проблема построения точных *D*-оптимальных планов экспериментов довольно полно исследована [1]. Для линейной модели парной регрессии с неравноточными наблюдениями в [2] построены точные *D*-оптимальные планы экспериментов. В [3] исследовалась проблема построения таких планов при линейном изменении дисперсии наблюдений

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = a_0 + a_1 x_{j1} + \ldots + a_m x_{jm}.$$

В настоящей работе результаты, полученные в [3], распространены на более широкий класс изменения дисперсии наблюдений.

Теорема 1. *Существует точный D-оптимальный план* ε^0 *экспериментов для модели наблюдений (1), (2), все точки спектра этого плана* ε^0 *лежат в вершинах единичного m-мерного куба,* $\left|x_i\right| \le 1,\ i = \overline{1,\, m}$.

Д о к а з а т е л ь с т в о. Пусть ε – произвольный, невырожденный план экспериментов. Покажем, что от плана ε можно перейти к плану ε_1, у которого все контролируемые переменные $x_{ij} = \pm 1,\ i = \overline{1,\, n}$, $j = \overline{1,\, m}$, и для которого $\left|M(\varepsilon)\right| \le \left|M(\varepsilon_1)\right|$, т. е. определитель информационной матрицы $M(\varepsilon_1)$ больше определителя информационной матрицы $M(\varepsilon)$ или равен ему. Информационная матрица плана экспериментов ε может быть представлена следующим образом:

$$M(\varepsilon) = \sum_{j=1}^{n} \frac{x^{(j)}\left(x^{(j)}\right)'}{d\left(x^{(j)}\right)} = \frac{x^{(1)}\left(x^{(1)}\right)'}{d\left(x^{(1)}\right)} + \sum_{j=2}^{n} \frac{x^{(j)}\left(x^{(j)}\right)'}{d\left(x^{(j)}\right)} = AA' + B,$$

где

$$A = \frac{x^{(1)}}{\sqrt{d\left(x^{(1)}\right)}};\ B = \sum_{j=2}^{n} \frac{x^{(j)}\left(x^{(j)}\right)'}{d\left(x^{(j)}\right)}.$$

Тогда в силу известной формулы [4, с. 302]

$$\left|M(\varepsilon)\right| = \left|AA' + B\right| = |B|\left(1 + \frac{\left(x^{(1)}\right)' B^{-1} x^{(1)}}{d\left(x^{(1)}\right)}\right) = |B| + \frac{\left(x^{(1)}\right)' B_1 x^{(1)}}{d\left(x^{(1)}\right)},$$

где B_1 – матрица, элементы которой являются алгебраическими дополнениями элементов матрицы B. Поскольку $d\left(x^{(1)}\right) \geq a_0 + a_1 x_{11} + \ldots + a_m x_{1m}$, то верна оценка сверху для определителя информационной матрицы:

$$\left|M(\varepsilon)\right| = |B| + \frac{\left(x^{(1)}\right)' B_1 x^{(1)}}{d\left(x^{(1)}\right)} \leq |B| + \frac{\left(x^{(1)}\right)' B_1 x^{(1)}}{a_0 + a_1 x_{11} + \ldots + a_m x_{1m}}. \qquad (3)$$

Покажем, что максимальное значение определителя информационной матрицы плана ε при выборе точки спектра $x^{(1)}$ в единичном m-мерном кубе достигается тогда, когда эта точка лежит в одной из вершин этого куба, т. е. когда все $x_{1j} = \pm 1,\ j = \overline{1, m}.$ Для того чтобы убедиться в этом, покажем, что максимальное значение правой части неравенства (3) по $x_{1j} \in [-1, 1],\ j = \overline{1, m},$ также достигается, когда $x^{(1)}$ – одна из вершин единичного m-мерного куба. А поскольку $d\left(x^{(1)}\right) = a_0 + a_1 x_{11} + \ldots + a_m x_{1m}$ в вершинах единичного m-мерного куба, то неравенство (3) в этих вершинах обращается в равенство.

Будем увеличивать значение правой части неравенства (3), последовательно варьируя значения $x_{11}, \ldots, x_{1m}$ в интервале $[-1, 1]$. Сначала определим верхнее значение правой части (3) при изменении $x_{11} = x$ от -1 до 1. Поскольку $|B|$ не зависит от $x^{(1)}$, то для этого достаточно вычислить максимум по $x \in [-1, 1]$ функции

$$f_{11}(x) = \frac{\left(x^{(1)}\right)' B_1 x^{(1)}}{a_0 + a_1 x + a_2 x_{12} + \ldots + a_m x_{1m}} = \frac{\alpha_{11} x^2 + \beta_{11} x + \gamma_{11}}{a_1 x + c_1},$$

где $\alpha_{11} = \left|\sum_{j=2}^{n} \frac{z^{(j)}\left(z^{(j)}\right)'}{d\left(x^{(j)}\right)}\right| \geq 0,\ z^{(j)} = \left(x_{j2}, \ldots, x_{jm}\right)';\ c_1 = a_0 + a_2 x_{12} + \ldots + a_m x_{1m};\ \beta_{11}, \gamma_{11}$ – некоторые константы, не зависящие от x. Покажем, что максимум $f_{11}(x)$ достигается при $x = -1$ либо $x = 1$.

Если $\alpha_{11} = 0$, то

$$\frac{df_{11}(x)}{dx} = \frac{c_1 \beta_{11} - a_1 \gamma_{11}}{\left(a_1 x + c_1\right)^2}. \qquad (4)$$

В случае когда числитель в (4) равен нулю, функция $f_{11}(x)$ остается постоянной на интервале $[-1, 1]$. Можно считать, что ее максимальное значение достигается при $x = \pm 1$. Если числитель в (4) отличен от нуля, то производная положительна либо отрицательна на всем интервале $[-1, 1]$. В этом случае максимум $f_{11}(x)$ будет достигаться при $x = -1$ либо $x = 1$.

Когда $\alpha_{11} \neq 0$, то это означает, что $\alpha_{11} > 0$. Производная функции $f_{11}(x)$ равна следующему выражению:

$$\frac{df_{11}(x)}{dx} = \frac{\alpha_{11}a_1x^2 + 2\alpha_{11}c_1x + c_1\beta_{11} - a_1\gamma_{11}}{(a_1x + c_1)^2}. \tag{5}$$

Пусть $D = 4\alpha_{11}\left(\alpha_{11}c_1^2 - a_1\left(c_1\beta_{11} - a_1\gamma_1\right)\right)$ – дискриминант числителя в (5). Если $D \leq 0$, то производная в (5) не меняет своего знака на интервале $[-1, 1]$, т. е. $f_{11}(x)$ либо строго возрастает, либо убывает на $[-1, 1]$. Если $D > 0$, то функция $f_{11}(x)$ является выпуклой. Действительно,

$$\frac{d^2 f_{11}(x)}{dx^2} = \frac{D}{2\alpha_{11}(a_1x + c_1)^2} > 0.$$

Итак, в любом случае функция $f_{11}(x)$ достигает своего максимального значения при $x_{11}^0 = -1$ либо $x_{11}^0 = 1$.

Далее, будем максимизировать правую часть неравенства (3) по $x_{12} = x$ при условии, что $x_{11} = x_{11}^0$. Аналогичным образом проведем процесс максимизации функции $f_{12}(x)$:

$$f_{12}(x) = \frac{\alpha_{12}x^2 + \beta_{12}x + \gamma_{12}}{a_2x + c_2},$$

где

$$\alpha_{12} = \left|\sum_{j=2}^{n} \frac{z^{(j)}\left(z^{(j)}\right)}{d\left(x^{(j)}\right)}\right| \geq 0, \; z^{(j)} = \left(x_{j1}, x_{j3}, \ldots, x_{jm}\right)', \; c_2 = a_0 + a_1x_{11}^0 + a_3x_{13} + \ldots + a_mx_{1m}.$$

Рассуждая аналогично, получаем, что максимальное значение функции $f_{12}(x)$ достигается при $x_{12}^0 = -1$ либо $x_{12}^0 = 1$. Очевидно, что $f_{11}\left(x_{11}^0\right) \leq f_{12}\left(x_{12}^0\right)$. Продолжим вычислять максимум правой части (3) по x_{13} при условии, что $x_{11} = x_{11}^0$, $x_{12} = x_{12}^0$. На финальной стадии будем максимизировать правую часть (3) по x_{1m} при условии, что $x_{1j} = x_{1j}^0$, $j = \overline{1, m-1}$. В итоге получим цепочку неравенств: $f_{11}\left(x_{11}^0\right) \leq \ldots \leq f_{1m}\left(x_{1m}^0\right)$. В функции $f_{1m}\left(x_{1m}^0\right)$ все значения $x_{1j}^0 = \pm 1$, $j = \overline{1, m}$, и эти значения определяют некоторую вершину $x_0^{(1)}$ единичного m-мерного куба. Следовательно, согласно (2) в этой вершине неравенство (3) обратится в равенство. Это означает, что определитель информационной матрицы плана экспериментов достигнет в этой вершине максимального значения $f_{1m}\left(x_{1m}^0\right)$ при вариации точки $x^{(1)}$ в m-мерном кубе.

Продолжим, как и ранее, максимизировать $|M(\varepsilon)|$ по точке $x^{(2)}$ спектра плана ε при условии, что $x^{(1)} = x_0^{(1)}$. В результате получим, что максимум $|M(\varepsilon)|$ будет достигаться в некоторой вершине $x_0^{(2)}$ единичного куба. При таком процессе максимизации определителя информационной матрицы плана экспериментов его значение увеличится, возможно, не в строгом смысле. Завершим этот процесс максимизацией определителя по точке спектра плана $x^{(n)}$ при условии, что вычисленные ранее точки $x_0^{(j)}$, $j = \overline{1, n-1}$, находятся в некоторых вершинах единичного куба. В итоге приходим к заключению, что максимальное значение определителя информационной матрицы плана экспериментов достигается в случае, когда все точки спектра плана находятся в вершинах единичного куба. Для построения точного D-оптимального плана экспериментов остается определить, в каких именно вершинах единичного куба надо провести n наблюдений. Процесс доказательства теоремы завершен.

Теорема 1 позволяет строить точные D-оптимальные планы экспериментов для модели наблюдений (1), (2). Например, рассмотрим модель наблюдений

$$y_j = \theta_0 + \theta_1 x_{1j} + \theta_2 x_{2j} + \varepsilon\left(x^{(j)}\right), \ j = \overline{1, k}, \ k \geq 3, \tag{6}$$

для которой дисперсии ошибок наблюдений определяются функцией $d\left(x_1, x_2\right)$, удовлетворяющей неравенству

$$d\left(x_1, x_2\right) \geq 10 - x_1 + 4x_2, \tag{7}$$

которое обращается в равенство в вершинах единичного квадрата: $\left|x_1\right| \leq 1, \left|x_2\right| \leq 1.$

Неравенство (7) определяет некоторый класс функций $d\left(x_1, x_2\right)$. В данном случае такими функциями могут быть:

$$d_1\left(x_1, x_2\right) = 10 - x_1 + 4x_2, \quad d_2\left(x_1, x_2\right) = -\left(x_1 + 0{,}5\right)^2 - \left(x_2 - 2\right)^2 + 16{,}25,$$

$$d_3\left(x_1, x_2\right) = -2\left(x_1 + \frac{1}{4}\right)^2 - 3\left(x_2 - \frac{2}{3}\right)^2 + \frac{395}{24}$$

и ряд других. Функция $d_2\left(x_1, x_2\right)$ – круговой параболоид с вершиной в точке $x_1 = -0{,}5$, $x_2 = 2$, функция $d_3\left(x_1, x_2\right)$ – эллиптический параболоид с вершиной в точке $x_1 = \frac{-1}{4}$, $x_2 = \frac{2}{3}$. Эти параболоиды лежат выше плоскости $d_1\left(x_1, x_2\right) = 10 - x_1 + 4x_2$ на единичном квадрате, за исключением вершин этого квадрата.

Для модели (6) с тремя равноточными наблюдениями ($k = 3$) можно построить четыре точных D-оптимальных плана экспериментов [5; 6]:

$$\begin{gathered}
\varepsilon_1 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)} \\ 1, & 1, & 1 \end{Bmatrix}, \ \varepsilon_2 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(4)} \\ 1, & 1, & 1 \end{Bmatrix}, \\
\varepsilon_3 = \begin{Bmatrix} x^{(1)}, & x^{(3)}, & x^{(4)} \\ 1, & 1, & 1 \end{Bmatrix}, \ \varepsilon_4 = \begin{Bmatrix} x^{(2)}, & x^{(3)}, & x^{(4)} \\ 1, & 1, & 1 \end{Bmatrix},
\end{gathered} \tag{8}$$

где $x^{(1)} = (1, 1)$; $x^{(2)} = (-1, 1)$; $x^{(3)} = (-1, -1)$; $x^{(4)} = (1, -1)$ – вершины единичного квадрата. Планы (8) – это так называемые насыщенные оптимальные планы экспериментов, у которых число наблюдений совпадает с числом неизвестных параметров. Определители информационных матриц этих планов равны 16 при условии, что дисперсия каждого наблюдения равна 1.

Для модели наблюдений (6) с тремя неравноточными наблюдениями (7) точки спектра насыщенного D-оптимального плана в силу теоремы 1 должны располагаться в вершинах единичного квадрата. Всего имеется 20 различных способов провести три наблюдения в четырех вершинах квадрата. Вычисления показывают, что для неравноточных наблюдений (7) четыре наиболее высоких значения определителя информационной матрицы планов экспериментов достигаются для планов (8). Эти значения следующие:

$$\left|M\left(\varepsilon_1\right)\right| = 0{,}011\,722, \quad \left|M\left(\varepsilon_2\right)\right| = 0{,}016\,41, \quad \left|M\left(\varepsilon_3\right)\right| = 0{,}035\,165, \quad \left|M\left(\varepsilon_4\right)\right| = 0{,}030\,476.$$

Наибольшее значение определителя информационной матрицы соответствует плану ε_3. Это и есть D-оптимальный насыщенный план экспериментов для модели наблюдений (6), (7). Важно отметить, что план ε_3 находится среди планов (8) для равноточных наблюдений.

Проведем аналогичные расчеты для модели (6) с четырьмя наблюдениями ($k = 4$). Для модели равноточных наблюдений точный D-оптимальный план экспериментов единствен [1]:

$$\varepsilon = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ 1, & 1, & 1, & 1 \end{Bmatrix}, \tag{9}$$

определитель информационной матрицы этого плана равен 64 при условии, что дисперсии всех наблюдений равны 1. Для неравноточных наблюдений (7) при $k = 4$ в силу теоремы 1 четыре наблюдения нужно провести в вершинах единичного квадрата. Всего есть 35 способов проведения этих наблюдений. Вычисления показывают, что из этих способов только для плана (9) определитель информационной матрицы достигает максимального значения, равного 0,093 773. Таким образом, план (9) – D-оптимальный план для неравноточных наблюдений вида (7).

Для модели наблюдений (6) с пятью равноточными наблюдениями D-оптимальные планы экспериментов, согласно [1], имеют вид

$$\begin{gathered}\varepsilon_1=\begin{Bmatrix}x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)}\\ 2, & 1, & 1, & 1\end{Bmatrix},\ \varepsilon_2=\begin{Bmatrix}x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)}\\ 1, & 2, & 1, & 1\end{Bmatrix},\\ \varepsilon_3=\begin{Bmatrix}x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)}\\ 1, & 1, & 2, & 1\end{Bmatrix},\ \varepsilon_4=\begin{Bmatrix}x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)}\\ 1, & 1, & 1, & 2\end{Bmatrix}.\end{gathered} \tag{10}$$

Определители информационных матриц планов (10) равны 0,896 для случая, когда дисперсии всех наблюдений равны 5. Для неравноточных наблюдений (7) при $k = 5$, согласно теореме 1, пять наблюдений нужно провести в четырех вершинах единичного квадрата для построения D-оптимального плана экспериментов. Всего имеется 56 таких реализаций. Вычисления показывают, что четыре наибольших определителя информационных матриц из этих случаев достигаются на четырех планах (10). Эти значения таковы:

$$|M(\varepsilon_1)| = 0{,}157\,07,\quad |M(\varepsilon_2)| = 0{,}152\,381,\quad |M(\varepsilon_3)| = 0{,}171136,\quad |M(\varepsilon_4)| = 0{,}175\,824.$$

Отсюда заключаем, что план ε_4 – D-оптимальный для неравноточных наблюдений. Опять замечаем, что D-оптимальный план ε_3 находится среди оптимальных планов для равноточных наблюдений.

Процесс построения оптимальных планов для неравноточных наблюдений для трех – пяти наблюдений приводит к следующей гипотезе: оптимальные планы для неравноточных наблюдений вида (2) – это некоторое подмножество оптимальных планов для равноточных наблюдений. Пока это только гипотеза. Если она подтвердится, то это значительно упростит процесс построения D-оптимальных планов для неравноточных наблюдений вида (2).

Следствием из теоремы 1 является следующая теорема.

Теорема 2. *Для модели наблюдений (1) с некоррелированными ошибками наблюдений, имеющими средние значения 0 и дисперсии*

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) \geq a_0,\ \ a_0 > 0,\ \ j = \overline{1,\ n}, \tag{11}$$

для функций $d(x)$ таких, что неравенство (11) обращается в равенство в вершинах единичного m-мерного куба: $|x_i| \leq 1,\ i = \overline{1,\ m},$ точные D-оптимальные планы ε^0 остаются такими же, как и для равноточных наблюдений.

Д о к а з а т е л ь с т в о. В силу теоремы 1 точки спектра D-оптимального плана ε^0 лежат в вершинах единичного m-мерного куба и в этих вершинах дисперсии наблюдений одинаковы и равны a_0. Таким образом, условия построения D-оптимальных планов для неравноточных наблюдений (11) совпадают с условиями построения таких же планов с равноточными наблюдениями. Это завершает доказательство теоремы.

Например, для модели наблюдений (1) с четырьмя факторами и функцией $d(x_1,\ \ldots,\ x_4) = 20 - 2|x_1| - |x_2| - 4|x_3| - 6|x_4|$, определяющей изменение дисперсии наблюдений, точные D-оптимальные планы совпадают с оптимальными планами для равноточных наблюдений с дисперсиями $a_0 = 7$ [1].

Используя теорему 2, построим насыщенные D-оптимальные планы экспериментов для модели наблюдений (6) с неравноточными наблюдениями на единичном квадрате. План экспериментов называется насыщенным, если в своем спектре он содержит число точек, равное числу неизвестных параметров. Покажем, что таких планов можно построить бесконечное, несчетное множество.

В [6] был предложен алгоритм построения насыщенных D-оптимальных планов для линейной множественной регрессии с тремя факторами для равноточных наблюдений. Согласно [6] можно построить 24 таких плана. В [5] удалось обобщить процесс построения подобных планов для равноточных наблюдений. В частности, в [5] было показано, что можно строить бесконечное, несчетное множество D-оптимальных планов для модели (1) с тремя факторами и равными дисперсиями наблюдений. Оказывается, что для модели неравноточных наблюдений (1), (2) существует, при определенном подборе функции $d(x_1, \ldots, x_m)$, бесконечное несчетное множество D-оптимальных планов. Проиллюстрируем это на примере модели наблюдений (6) с функцией $d(x_1, x_2)$, описывающей изменение дисперсии наблюдений на единичном квадрате:

$$d(x_1, x_2) = 2 - x_1,\ 0 \le x_1 \le 1,\ |x_2| \le x_1;\ d(x_1, x_2) = 2 - x_2,\ 0 \le x_2 \le 1,\ |x_1| \le x_2,$$

$$d(x_1, x_2) = 2 + x_1,\ -1 \le x_1 \le 0,\ |x_2| \le |x_1|;\ d(x_1, x_2) = 2 + x_2,\ -1 \le x_2 \le 0,\ |x_1| \le |x_2|. \quad (12)$$

Функция (12) описывает боковую поверхность правильной прямоугольной пирамиды с высотой, равной 1, основание которой квадрат: $x_3 = 1,\ |x_1| \le 1,\ |x_2| \le 1.$ Ясно, что на единичном квадрате $|x_1| \le 1,\ |x_2| \le 1$ функция $d(x_1, x_2) \ge a_0 = 1.$ Следовательно, в силу теоремы 2 насыщенные D-оптимальные планы для модели наблюдений (6) с дисперсиями наблюдений (12) совпадают с насыщенными планами для равноточных наблюдений, которые согласно [5] имеют вид

$$\varepsilon_1 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & \lambda x^{(3)} + (1-\lambda)x^{(4)} \\ 1, & 1, & 1 \end{Bmatrix},\ \varepsilon_2 = \begin{Bmatrix} x^{(2)}, & x^{(3)}, & \lambda x^{(1)} + (1-\lambda)x^{(4)} \\ 1, & 1, & 1 \end{Bmatrix},$$

$$\varepsilon_3 = \begin{Bmatrix} x^{(3)}, & x^{(4)}, & \lambda x^{(1)} + (1-\lambda)x^{(2)} \\ 1, & 1, & 1 \end{Bmatrix},\ \varepsilon_4 = \begin{Bmatrix} x^{(1)}, & x^{(4)}, & \lambda x^{(2)} + (1-\lambda)x^{(3)} \\ 1, & 1, & 1 \end{Bmatrix},\ \lambda \in [0, 1]. \quad (13)$$

Планы (13) образуют бесконечное, несчетное множество.

Теорему 1, очевидно, можно обобщить на случай, когда функция, определяющая поведение дисперсии наблюдений, изменяется от одного наблюдения к другому, т. е. зависит от номера, или порядка, наблюдений:

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) \ge a_{0,j} + a_{1j}x_{1j} + \ldots + a_{mj}x_{mj} > 0,\ \ j = \overline{1, n}.$$

Например, для модели с четырьмя наблюдениями (6) и функциями

$$d_j(x_1, x_2) = 10 - x_1 + 4x_2,\ \ j = 1, 2, 3,\ \ d_4(x_1, x_2) = 10 - x_1 + 8x_2$$

D-оптимальный план, как показывают вычисления, имеет следующую структуру. Первые три наблюдения должны проводиться в точках $x^{(1)} = (1, 1),\ x^{(2)} = (-1, 1),\ x^{(3)} = (-1, 1)$ в произвольном порядке, однако четвертое наблюдение надо провести в точке $x^{(4)} = (1, -1).$ Определитель информационной матрицы такого плана равен 0,421 978.

Библиографические ссылки

1. *Moyssiadis C., Kounias S.* Exact *D*-optimal observations 2^k designs of resolution 3, when $N \equiv 1$ or 2 mod 4 // Math. Operationsforch. U. Statist. Ser.: Statist. 1968. Vol. 14, № 3. P. 367–379.
2. *Кирлица В. П.* Точные *D*-оптимальные планы экспериментов для линейной модели парной регрессии // Вестн. БГУ. Сер. 1, Физика. Математика. Информатика. 2016. № 2. С. 116–122.
3. *Kirlitsa V. P.* Exact *D*-optimal designs of experiments for linear multiple model with linear variance of observations // Comput. data anal. model. Robust. comput. intensive methods. Minsk, 2004. Vol. 1. P. 165–167.
4. *Ермаков С. М., Жиглявский А. А.* Математическая теория оптимального эксперимента. М., 1987.
5. *Кирлица В. П.* О структуре насыщенных *D*-оптимальных планов экспериментов // Вестн. БГУ. Сер. 1, Физика. Математика. Информатика. 2013. № 3. С. 86–90.
6. *Williamson J.* Determinants whose elements are 0 and 1 // Am. Math. Monthly. 1946. Vol. 53. P. 427–434.

References

1. Moyssiadis C., Kounias S. Exact *D*-optimal observations 2^k designs of resolution 3, when $N \equiv 1$ or 2 mod 4. *Math. Operationsforch. U. Statist. Ser.: Statist.* 1968. Vol. 14, No. 3. P. 367–379.
2. Kirlitsa V. P. Exact *D*-optimal designs of experiments for linear pair regression. *Vestnik BGU. Ser. 1, Fiz. Mat. Inform.* 2016. No. 2. P. 116–122 (in Russ.).
3. Kirlitsa V. P. Exact *D*-optimal designs of experiments for linear multiple model with linear variance of observations. *Comput. data anal. model. Robust. comput. intensive methods.* Minsk, 2004. Vol. 1. P. 165–167.
4. Ermakov S. M., Zhiglavski A. A. [The mathematical theory of optimal design]. Moscow, 1987 (in Russ.).
5. Kirlitsa V. P. [About structure of *D*-optimal saturated designs of experiment]. *Vestnik BGU. Ser. 1, Fiz. Mat. Inform.* 2013. No. 3. P. 86–90 (in Russ.).
6. Williamson J. Determinants whose elements are 0 and 1. *Am. Math. Monthly.* 1946. Vol. 53. P. 427–434.

Статья поступила в редколлегию 03.05.2017.
Received by editorial board 03.05.2017.