УДК 519.4

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*

Исследуется проблема построения непрерывных (число наблюдений не фиксируется) и точных (с фиксированным числом наблюдений) *D*-оптимальных планов экспериментов для линейной множественной регрессии в случае, когда дисперсия ошибок зависит от точки, в которой проводится наблюдение. Определен класс функций, описывающих изменение дисперсии неравноточных наблюдений, для которых можно построить непрерывные и точные *D*-оптимальные планы экспериментов. Для линейной множественной регрессии с тремя факторами построены непрерывные *D*-оптимальные планы экспериментов с четырьмя различными типами неравноточных наблюдений. Для каждого из этих типов выделен свой собственный класс функций, описывающих изменение дисперсии наблюдений.

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

Образец цитирования:
Кирлица ВП. *D*-оптимальные планы экспериментов для линейной множественной регрессии с неравноточными наблюдениями. *Журнал Белорусского государственного университета. Математика. Информатика.* 2020;2:59–67.
https://doi.org/10.33581/2520-6508-2020-2-59-67

For citation:
Kirlitsa VP. *D*-optimal experimental designs for linear multiple regression under heteroscedastic observations. *Journal of the Belarusian State University. Mathematics and Informatics.* 2020;2:59–67. Russian.
https://doi.org/10.33581/2520-6508-2020-2-59-67

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 modeling and data analysis, faculty of applied mathematics and computer science.
kirlitsa@bsu.by

D-OPTIMAL EXPERIMENTAL DESIGNS FOR LINEAR MULTIPLE REGRESSION UNDER HETEROSCEDASTIC OBSERVATIONS

V. P. KIRLITSA[a]

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

The problem of construction of «continuous» (number of observations is not fixed) and «exact» (number of observations is fixed) *D*-optimal experimental designs for linear multiple regression in the case when variance of errors of observations depends on regressor value is studied in this paper. Families of functions that determine heteroscedastic observations are found for which it is possible to construct «continuous» and «exact» *D*-optimal experimental designs. «Continuous» *D*-optimal experimental designs under four different types of heteroscedasticity are constructed for linear multiple regression with three regressors.

Keywords: *D*-optimal experimental designs; linear multiple regression; heteroscedastic observations.

Рассмотрим линейную модель множественной регрессии

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

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 ($\left|x_i\right| \leq 1, \ i = \overline{1, m}$); $\theta_1, \ldots, \theta_m$ – неизвестные параметры; $\varepsilon\left(x^{(j)}\right)$ – некоррелированные случайные ошибки наблюдений с нулевыми математическими ожиданиями и дисперсиями, зависящими от точки наблюдения $x^{(j)}$:

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) > 0, \ j = \overline{1, n}, \tag{2}$$

где $d\left(x_1, \ldots, x_m\right)$ – некоторая непрерывная функция. Далее в статье на функцию $d(x)$ будут наложены определенные ограничения снизу, позволяющие строить *D*-оптимальные планы экспериментов для модели наблюдений (1), (2).

Для равноточных наблюдений ($d(x) = d = \text{const}$) проблема построения точных *D*-оптимальных планов экспериментов для модели наблюдений (1) довольно полно исследована [1]. В статье [2] построены точные *D*-оптимальные планы экспериментов для линейной модели парной регрессии с неравноточными наблюдениями. В работе [3] исследовалась проблема построения точных *D*-оптимальных планов экспериментов для модели наблюдений (1) при линейном изменении дисперсии наблюдений. В статье автора [4] результаты построения точных *D*-оптимальных планов экспериментов для модели наблюдений (1), полученные в работе [3], были обобщены на более широкий класс дисперсий неравноточных наблюдений:

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) \geq a_0 + a_1 x_{j1} + \ldots + a_m x_{jm} > 0, \ a_0 > 0, \left|a_1\right| + \ldots + \left|a_m\right| < a_0, \tag{3}$$

для каждой точки наблюдения $x^{(j)}$. Функция $d\left(x_1, \ldots, x_m\right)$ в (3) должна быть такой, чтобы в вершинах единичного m-мерного куба ($\left|x_j\right| \leq 1, \ j = \overline{1, m}$) неравенство (3) обращалось в равенство. Как отмечалось в статье [4], класс функций $d\left(x_1, \ldots, x_m\right)$, описываемых неравенством (3), обширен. К нему принадлежат постоянные функции (равноточные наблюдения) $d\left(x_1, \ldots, x_m\right) = a_0 > 0, \ a_1 = 0, \ldots, a_m = 0$, с линейным изменением $d\left(x_1, \ldots, x_m\right) = a_0 + a_1 x_{j1} + \ldots + a_m x_{jm} > 0, \ a_0 > 0, \ \left|a_1\right| + \ldots + \left|a_m\right| < a_0$, вогнутые функции $d\left(x_1, \ldots, x_m\right) = -\sum\limits_{j=1}^{m} k_j \left(x_j - \dfrac{a_j}{2k_j}\right)^2 + \sum\limits_{j=1}^{m} \dfrac{a_j^2 + 4k_j^2}{4k_j} + a_0, \ k_j > 0, \ j = \overline{1, m}.$

В статье [4] процесс построения точных *D*-оптимальных планов экспериментов для модели наблюдений (1) основывался на следующих теоремах.

Теорема 1. *Существует точный D-оптимальный план экспериментов* ε^0 *для модели наблюдений (1), (3), все точки спектра которого лежат в вершинах единичного m-мерного куба.*

Как следствие, из теоремы 1 вытекает теорема 2.

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

$$D\left\{\varepsilon\left(x^{(j)}\right)\right\} = d\left(x^{(j)}\right) \geq a_0,\ a_0 > 0,\ j = \overline{1, n}, \tag{4}$$

для функций $d\left(x_1, \ldots, x_m\right)$ *таких, что неравенство (4) обращается в равенство в вершинах единичного m-мерного куба, точные D-оптимальные планы экспериментов остаются такими же, как и для равноточных наблюдений.*

Примерами функций, удовлетворяющих (4), могут служить функции $d(x) = a_0$; $d(x) = -kx_1^2 - \ldots$ $\ldots - kx_m^2 + a_0 + mk$, $k > 0$; $d(x) = a_0 + m - |x_1| - \ldots - |x_m|$, $a_0 > 0$, и ряд других.

Теорема 1 позволяет эффективно строить оптимальные планы экспериментов, используя ЭВМ. Например, для частного случая модели наблюдений (1)

$$y_j = \theta_0 + \theta_1 x_{j1} + \theta_2 x_{j2} + \varepsilon\left(x^{(j)}\right),\ j = \overline{1, n},\ n \geq 3, \tag{5}$$

с дисперсиями наблюдений $d\left(x_1, x_2\right)$, удовлетворяющими неравенству

$$d\left(x_1, x_2\right) \geq 8 + 4x_1 - 3x_2, \tag{6}$$

построим точные D-оптимальные планы экспериментов. Функции $d\left(x_1, x_2\right)$, удовлетворяющие (6), должны быть такими, чтобы в вершинах единичного квадрата $x^{(1)} = (1, 1)$, $x^{(2)} = (-1, 1)$, $x^{(3)} = (-1, -1)$, $x^{(4)} = (1, -1)$ неравенство (6) обращалось в равенство. Дисперсии наблюдений в этих вершинах должны быть равны $d_1 = d(1, 1) = 9$, $d_2 = d(-1, 1) = 1$, $d_3 = d(-1, -1) = 7$, $d_4 = d(1, -1) = 15$. Такими функциями могут быть, например, $d\left(x_1, x_2\right) = 8 + 4x_1 - 3x_2$; $d\left(x_1, x_2\right) = -\left(x_1 - 2\right)^2 - \left(x_2 + \frac{3}{2}\right)^2 + \frac{65}{4}$; $d\left(x_1, x_2\right) = -3\left(x_1 - \frac{2}{3}\right)^2 - \left(x_2 + \frac{3}{2}\right)^2 + \frac{187}{12}$ и ряд других. Точный D-оптимальный план экспериментов для модели (5), (6) с n наблюдениями согласно теореме 1 имеет следующую структуру:

$$\varepsilon_n^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ m_1, & m_2, & m_3, & m_4 \end{Bmatrix},$$

где $0 \leq m_j \leq n - 1$, $j = 1, 2, 3, 4$, $m_1 + m_2 + m_3 + m_4 = n$. Информационная матрица плана экспериментов ε_n^0 равна

$$M\left(\varepsilon_n^0\right) = \frac{m_1}{d_1}\begin{pmatrix}1\\1\\1\end{pmatrix}(1, 1, 1) + \frac{m_2}{d_2}\begin{pmatrix}1\\-1\\1\end{pmatrix}(1, -1, 1) + \frac{m_3}{d_3}\begin{pmatrix}1\\-1\\-1\end{pmatrix}(1, -1, -1) + \frac{m_4}{d_4}\begin{pmatrix}1\\1\\-1\end{pmatrix}(1, 1, -1) = \begin{pmatrix} a & b & c \\ b & a & e \\ c & e & a \end{pmatrix},$$

где

$$a = \frac{m_1}{d_1} + \frac{m_2}{d_2} + \frac{m_3}{d_3} + \frac{m_4}{d_4},\ b = \frac{m_1}{d_1} - \frac{m_2}{d_2} - \frac{m_3}{d_3} + \frac{m_4}{d_4},$$

$$c = \frac{m_1}{d_1} + \frac{m_2}{d_2} - \frac{m_3}{d_3} - \frac{m_4}{d_4},\ e = \frac{m_1}{d_1} - \frac{m_2}{d_2} + \frac{m_3}{d_3} - \frac{m_4}{d_4}.$$

Определитель информационной матрицы равен

$$\left|M\left(\varepsilon_n^0\right)\right| = f\left(m_1, m_2, m_3\right) = a^3 + 2bce - a\left(b^2 + c^2 + e^2\right),$$

где $0 \leq m_j \leq n - 1$, $j = 1, 2, 3$, $1 \leq m_1 + m_2 + m_3 \leq n$. Вычисляя функцию $f\left(m_1, m_2, m_3\right)$ и фиксируя m_1, m_2, m_3, при которых эта функция принимает наибольшее значение, строим точные D-оптимальные планы экспериментов. Для $n = 5$, 6 оптимальные планы таковы:

$$\varepsilon_5^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ 1, & 2, & 1, & 1 \end{Bmatrix}, \ \left|M\left(\varepsilon_5^0\right)\right| = 1{,}066\,667;$$

$$\varepsilon_6^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ 1, & 2, & 2, & 1 \end{Bmatrix}, \ \left|M\left(\varepsilon_6^0\right)\right| = 1{,}896\,296.$$

Для модели (1) с неравноточными наблюдениями можно построить непрерывные D-оптимальные планы экспериментов в случае, когда дисперсия ошибок наблюдений $d\left(x_1, \ldots, x_m\right)$ удовлетворяет неравенству

$$d\left(x_1, \ldots, x_m\right) \geq \frac{\sigma^2}{m}\left(x_1^2 + \ldots + x_m^2\right), \ \sigma \neq 0, \ |x_1| + \ldots + |x_m| \neq 0. \tag{7}$$

В вершинах единичного m-мерного куба неравенство (7) должно обращаться в равенство. Неравенству (7) удовлетворяют, например, функции $d\left(x_1, \ldots, x_m\right) = \frac{\sigma^2}{m}\left(x_1^2 + \ldots + x_m^2\right)$; $d\left(x_1, \ldots, x_m\right) = \frac{\sigma^2}{m}\left(2m - x_1^2 - \ldots - x_m^2\right)$; $d\left(x_1, \ldots, x_m\right) = \frac{\sigma^2}{m}\left(2m - |x_1| - \ldots - |x_m|\right)$; $d\left(x_1, \ldots, x_m\right) = \sigma^2$ (равноточные наблюдения) и ряд других.

Обозначим через $X = \left(x_{i,\,j}\right)$, $i = \overline{1, n}$, $j = \overline{1, m}$, $n \geq m$, матрицу плана экспериментов для модели наблюдений (1), в которой i-я строка – координаты i-й точки $x^{(i)}$, предназначенной для проведения наблюдения. В статье [4] доказана следующая теорема.

Теорема 3. *Для модели наблюдений (1), (7) матрица плана экспериментов X со взаимно ортогональными столбцами и элементами, равными* ± 1, *определяет непрерывный D-оптимальный план экспериментов*

$$\varepsilon^0 = \left\{x^{(i)}, \ p_i = \frac{1}{n}, \ i = \overline{1, n}\right\}, \tag{8}$$

где p_i – веса наблюдений в точках $x^{(i)}$.

Примером непрерывных D-оптимальных планов (8) для модели наблюдений (1), (7), как известно, являются планы полных факторных экспериментов, в которых точки спектра сосредоточены во всех 2^m вершинах единичного m-мерного куба с равными весами $\frac{1}{2^m}$. При больших значениях m они содержат чрезмерно много точек. Процесс подбора матрицы X плана эксперимента, удовлетворяющей теореме 3, но с меньшим, чем 2^m, числом точек, связан с построением матриц Адамара. Так, например, для модели наблюдений (1), (7) с четырьмя факторами D-оптимальный план может содержать четыре точки: $x^{(1)} = (1, 1, 1, 1)$, $x^{(2)} = (1, -1, 1, -1)$, $x^{(3)} = (1, 1, -1, -1)$, $x^{(4)} = (1, -1, -1, 1)$, а не 16 точек, как в полных факторных экспериментах.

Результаты теоремы 3 можно использовать для построения точных D-оптимальных планов экспериментов с числом наблюдений N, кратным n, т. е. $N = ns$, $s \geq 1$. Для модели наблюдений (1), (7) с четырьмя факторами точки спектра $x^{(1)} = (1, 1, 1, 1)$, $x^{(2)} = (1, -1, 1, -1)$, $x^{(3)} = (1, 1, -1, -1)$, $x^{(4)} = (1, -1, -1, 1)$ точного D-оптимального плана экспериментов будут такими же, как и для непрерывного плана, но в каждой точке спектра надо проводить по s наблюдений. Определитель информационной матрицы такого плана равен $(4s)^4$.

Для частного случая модели наблюдений (1), (2)

$$y_i = \theta_1 x_{i1} + \theta_2 x_{i2} + \theta_3 x_{i3} + \varepsilon\left(x^{(i)}\right), \ i = \overline{1, n}, \ n \geq 3, \tag{9}$$

можно получить некоторые новые результаты относительно построения D-оптимальных планов экспериментов с неравноточными наблюдениями по сравнению с публикациями [4; 5].

Для модели (9) с неравноточными наблюдениями можно сконструировать четыре типа непрерывных D-оптимальных планов экспериментов. Конкретный тип планов экспериментов, как будет показано далее, определяется видом функции, описывающей изменение дисперсии наблюдений.

Введем следующие обозначения. Пронумеруем вершины единичного куба $\left|x_j\right| \leq 1$, $j = \overline{1, 3}$:

$$
\begin{gathered}
x^{(1)}=(1,1,1),\ x^{(2)}=(1,-1,1),\ x^{(3)}=(1,-1,-1),\ x^{(4)}=(1,1,-1),\\
x^{(5)}=(-1,1,1),\ x^{(6)}=(-1,-1,1),\ x^{(7)}=(-1,-1,-1),\ x^{(8)}=(-1,1,-1).
\end{gathered}
\tag{10}
$$

Пусть $d_j=d\left(x^{(j)}\right)$, $j\geq 1$, – дисперсии наблюдений в вершинах единичного куба, и эти значения положительны.

В процессе построения непрерывных D-оптимальных планов для модели (9) с неравноточными наблюдениями покажем, что для любого из четырех типов планов экспериментов на каждой из шести граней единичного куба можно надлежащим образом выбрать три различные вершины куба, определяющие точки спектра оптимальных планов. Определители информационных матриц оптимальных планов принимают значения $\dfrac{16}{27d_id_jd_k}$, где i, j, k – номера таких вершин. Эти определители должны быть равны между собой. Следовательно, произведение дисперсий $d_id_jd_k$ должно быть неизменным, т. е. $d_id_jd_k=F$, где F – фиксированное значение. Таким образом, для каждого из четырех типов непрерывных D-оптимальных планов экспериментов для модели (9) с неравноточными наблюдениями можно построить шесть планов, определители информационных матриц которых равны $\dfrac{16}{27F}$, где F может быть выбрано произвольным положительным числом.

Теорема 4. *Для модели (9) с некоррелированными ошибками наблюдений, имеющими средние значения* 0 *и дисперсии* $d(x_1,x_2,x_3)$, *можно построить шесть непрерывных D-оптимальных планов экспериментов для каждого из четырех различных типов дисперсий наблюдений* $d(x_1,x_2,x_3)$ *с заданным значением определителей информационных матриц, равным* $\dfrac{16}{27F}$.

Для дисперсий наблюдений $d(x_1,x_2,x_3)$ *первого типа, удовлетворяющих неравенству*

$$
d(x_1,x_2,x_3)\geq\frac{1}{4}\Big((d_1+d_3)x_1^2+(d_1+d_2)x_2^2+(d_2+d_3)x_3^2+2d_1x_1x_2-2d_3x_1x_3-2d_2x_2x_3\Big),
\tag{11}
$$

в котором равенство выполняется в точках $x^{(1)}, x^{(2)}, x^{(3)}$ *и* $|x_1|+|x_2|+|x_3|\neq 0$, *произведение* $d_1d_2d_3=F$, *оптимальные планы имеют вид*

$$
\begin{gathered}
\varepsilon_1^0=\begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_2^0=\begin{Bmatrix} x^{(5)}, & x^{(7)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_3^0=\begin{Bmatrix} x^{(1)}, & x^{(5)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\\
\varepsilon_4^0=\begin{Bmatrix} x^{(2)}, & x^{(3)}, & x^{(7)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_5^0=\begin{Bmatrix} x^{(3)}, & x^{(7)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_6^0=\begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(5)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix}.
\end{gathered}
\tag{12}
$$

Для дисперсий наблюдений $d(x_1,x_2,x_3)$ *второго типа, удовлетворяющих неравенству*

$$
d(x_1,x_2,x_3)\geq\frac{1}{4}\Big((d_2+d_4)x_1^2+(d_3+d_4)x_2^2+(d_2+d_3)x_3^2+2d_4x_1x_2+2d_2x_1x_3+2d_3x_2x_3\Big),
\tag{13}
$$

в котором равенство выполняется в точках $x^{(2)}, x^{(3)}, x^{(4)}$ *и* $|x_1|+|x_2|+|x_3|\neq 0$, *произведение* $d_2d_3d_4=F$, *оптимальные планы имеют вид*

$$
\begin{gathered}
\varepsilon_7^0=\begin{Bmatrix} x^{(2)}, & x^{(3)}, & x^{(4)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_8^0=\begin{Bmatrix} x^{(5)}, & x^{(6)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_9^0=\begin{Bmatrix} x^{(4)}, & x^{(5)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\\
\varepsilon_{10}^0=\begin{Bmatrix} x^{(2)}, & x^{(3)}, & x^{(6)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_{11}^0=\begin{Bmatrix} x^{(3)}, & x^{(4)}, & x^{(8)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix},\
\varepsilon_{12}^0=\begin{Bmatrix} x^{(2)}, & x^{(5)}, & x^{(6)}\\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3}\end{Bmatrix}.
\end{gathered}
\tag{14}
$$

Для дисперсий наблюдений $d(x_1, x_2, x_3)$ третьего типа, удовлетворяющих неравенству

$$d(x_1, x_2, x_3) \geq \frac{1}{4}\left((d_1 + d_3)x_1^2 + (d_3 + d_4)x_2^2 + (d_1 + d_4)x_3^2 - 2d_3x_1x_2 + 2d_1x_1x_3 - 2d_4x_2x_3\right), \tag{15}$$

в котором равенство выполняется в точках $x^{(1)}, x^{(3)}, x^{(4)}$ *и* $|x_1| + |x_2| + |x_3| \neq 0$, *произведение* $d_1d_3d_4 = F$, *оптимальные планы имеют вид*

$$\begin{gathered}\varepsilon_{13}^0 = \begin{Bmatrix} x^{(1)}, & x^{(3)}, & x^{(4)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{14}^0 = \begin{Bmatrix} x^{(5)}, & x^{(6)}, & x^{(7)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{15}^0 = \begin{Bmatrix} x^{(1)}, & x^{(4)}, & x^{(5)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \\ \varepsilon_{16}^0 = \begin{Bmatrix} x^{(3)}, & x^{(6)}, & x^{(7)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{17}^0 = \begin{Bmatrix} x^{(3)}, & x^{(4)}, & x^{(7)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{18}^0 = \begin{Bmatrix} x^{(1)}, & x^{(5)}, & x^{(6)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}. \end{gathered} \tag{16}$$

Для дисперсий наблюдений $d(x_1, x_2, x_3)$ четвертого типа, удовлетворяющих неравенству

$$d(x_1, x_2, x_3) \geq \frac{1}{4}\left((d_2 + d_4)x_1^2 + (d_1 + d_2)x_2^2 + (d_1 + d_4)x_3^2 - 2d_2x_1x_2 - 2d_4x_1x_3 + 2d_1x_2x_3\right), \tag{17}$$

в котором равенство выполняется в точках $x^{(1)}, x^{(2)}, x^{(4)}$ *и* $|x_1| + |x_2| + |x_3| \neq 0$, *произведение* $d_1d_2d_4 = F$, *оптимальные планы имеют вид*

$$\begin{gathered}\varepsilon_{19}^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(4)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{20}^0 = \begin{Bmatrix} x^{(6)}, & x^{(7)}, & x^{(8)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{21}^0 = \begin{Bmatrix} x^{(1)}, & x^{(4)}, & x^{(8)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \\ \varepsilon_{22}^0 = \begin{Bmatrix} x^{(2)}, & x^{(6)}, & x^{(7)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{23}^0 = \begin{Bmatrix} x^{(4)}, & x^{(7)}, & x^{(8)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}, \ \varepsilon_{24}^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(6)} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}. \end{gathered} \tag{18}$$

В планах экспериментов (12), (14), (16), (18) точки спектров – это точки (10).

Доказательство. Опишем вначале процесс построения непрерывных D-оптимальных планов, соответствующих дисперсиям наблюдений $d(x_1, x_2, x_3)$ первого типа. Для D-оптимального плана ε_1^0 по теореме эквивалентности Кифера – Вольфовица [6] должно выполняться неравенство

$$\frac{1}{d(x_1, x_2, x_3)}(x_1, x_2, x_3)M^{-1}\left(\varepsilon_1^0\right)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq 3, \ |x_1| \leq 1, \ |x_2| \leq 1, \ |x_3| \leq 1, \tag{19}$$

где $M\left(\varepsilon_1^0\right)$ – информационная матрица плана экспериментов. В точках спектра плана ε_1^0 неравенство (19) должно обращаться в равенство (как необходимое условие). Исходя из условий теоремы Кифера – Вольфовица, построим класс функций $d(x_1, x_2, x_3)$, определяющих поведение дисперсии ошибок наблюдений для плана ε_1^0. Информационная матрица плана ε_1^0 равна

$$M\left(\varepsilon_1^0\right) = \frac{1}{3}\left(\frac{1}{d_1}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}(1, 1, 1) + \frac{1}{d_2}\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}(1, -1, 1) + \frac{1}{d_3}\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}(1, -1, -1)\right) = \frac{1}{3}\begin{pmatrix} a & b & c \\ b & a & e \\ c & e & a \end{pmatrix},$$

где

$$a = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3}, \ b = \frac{1}{d_1} - \frac{1}{d_2} - \frac{1}{d_3}, \ c = \frac{1}{d_1} + \frac{1}{d_2} - \frac{1}{d_3}, \ e = \frac{1}{d_1} - \frac{1}{d_2} + \frac{1}{d_3}. \tag{20}$$

Определитель матрицы $M\left(\varepsilon_1^0\right)$ равен $\dfrac{a^3+2bce-a\left(b^2+c^2+e^2\right)}{27}$. Матрица, обратная к матрице $M\left(\varepsilon_1^0\right)$, имеет вид

$$M^{-1}\left(\varepsilon_1^0\right)=\frac{3}{a^3+2bce-a\left(b^2+c^2+e^2\right)}\begin{pmatrix} a^2-e^2, & ce-ab, & be-ac \\ ce-ab, & a^2-c^2, & bc-ae \\ be-ac, & bc-ae, & a^2-b^2 \end{pmatrix}. \quad (21)$$

Разрешая неравенство (19) относительно $d\left(x_1, x_2, x_3\right)$ с учетом (21), получим класс функций $d\left(x_1, x_2, x_3\right)$, определяющих изменение дисперсии наблюдений для плана ε_1^0:

$$d\left(x_1, x_2, x_3\right) \geq f\left(x_1, x_2, x_3\right), \quad (22)$$

где

$$f(x)=\frac{\left(a^2-e^2\right)x_1^2+\left(a^2-c^2\right)x_2^2+\left(a^2-b^2\right)x_3^2+2(ce-ab)x_1x_2+2(be-ac)x_1x_3+2(bc-ae)x_2x_3}{a^3+2bce-a\left(b^2+c^2+e^2\right)}.$$

Если теперь в функции $f\left(x_1, x_2, x_3\right)$ вернуться к исходным обозначениям (20), то неравенство (22) обратится в неравенство (11). Необходимое условие оптимальности плана ε_1^0 также выполняется, так как в точках $x^{(1)}$, $x^{(2)}$, $x^{(3)}$ спектра плана неравенство (11) обращается в равенство. Точки спектра плана ε_1^0 лежат на грани $x_1 = 1$ единичного куба. Неравенству (11) удовлетворяет множество функций $d\left(x_1, x_2, x_3\right)$, поскольку значения d_1, d_2, d_3 различные, но такие, что $d_1d_2d_3 = F$. Определитель информационной матрицы плана ε_1^0 равен $\dfrac{16}{27d_1d_2d_3}$. Оптимальность плана ε_1^0 доказана.

Оптимальность остальных планов экспериментов ε_2^0–ε_6^0, соответствующих дисперсиям (11), докажем следующим образом. В вершинах единичного куба, принадлежащих грани $x_1 = -1$, значения дисперсий наблюдений, соответствующих (11), удовлетворяют соотношениям

$$d_5=d\left(x^{(5)}\right)\geq d_3,\ d_6=d\left(x^{(6)}\right)\geq d_1+d_2+d_3,\ d_7=d\left(x^{(7)}\right)\geq d_1,\ d_8=d\left(x^{(8)}\right)\geq d_2. \quad (23)$$

Согласно (23) минимальное значение произведения трех различных дисперсий таково: $d_5d_7d_8 \geq$ $\geq F = d_3d_1d_2$. Следовательно, на грани $x_1 = -1$ можно построить непрерывный D-оптимальный план только в том случае, если $d_5 = d_3$, $d_7 = d_1$, $d_8 = d_2$. Тогда $d_5d_7d_8 = F$. Это достигается, если в точках $x^{(5)}$, $x^{(7)}$, $x^{(8)}$ плана ε_2^0 в (11) выполняется равенство. Но есть ли такие функции $d\left(x_1, x_2, x_3\right)$, удовлетворяющие (11), для которых в точках спектров планов ε_1^0, ε_2^0 неравенство (11) обращается в равенство? Ответ утвердительный. Примерами таких функций являются следующие:

$$d\left(x_1, x_2, x_3\right)=\frac{1}{4}\left(\left(d_1+d_3\right)x_1^2+\left(d_1+d_2\right)x_2^2+\left(d_2+d_3\right)x_3^2+2d_1x_1x_2-2d_3x_1x_3-2d_2x_2x_3\right);$$

$$d\left(x_1, x_2, x_3\right)=\frac{1}{4}\left(\left(d_1+d_3-\frac{k}{3}\right)x_1^2+\left(d_1+d_2-\frac{k}{3}\right)x_2^2+\left(d_2+d_3-\frac{k}{3}\right)x_3^2+B\right),$$

$$B=2d_1x_1x_2-2d_3x_1x_3-2d_2x_2x_3+k,\ k>3\max\left\{d_1+d_3,\ d_1+d_2,\ d_2+d_3\right\}.$$

Можно построить и ряд других подобных функций. Итак, дисперсиям вида (11) на грани $x_1 = -1$ единичного куба соответствует оптимальный план ε_2^0.

Аналогичным образом можно показать оптимальность планов ε_3^0–ε_6^0, соответствующих дисперсиям наблюдений первого типа, а также оптимальность остальных планов, соответствующих дисперсиям наблюдений второго, третьего и четвертого типов. Теорема 4 доказана.

Непрерывные D-оптимальные планы, построенные по теореме 4, можно использовать для построения точных D-оптимальных планов экспериментов с числом наблюдений n, кратным 3, т. е. $n = 3s$, $s \geq 1$. Так, для модели (9) неравноточных наблюдений с дисперсиями $d\left(x_1, x_2, x_3\right)=0{,}25\left(-x_1^2-x_2^2-\right.$ $\left.-\,x_3^2+2x_1x_2-2x_1x_3-2x_2x_3+9\right)$ D-оптимальный план с шестью наблюдениями равен

$$\varepsilon^0(6) = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)} \\ 2, & 2, & 2 \end{Bmatrix}, \ \left|M\left(\varepsilon^0(6)\right)\right| = 128.$$

Для модели (9) равноточных наблюдений с дисперсиями $\sigma^2 = 1$ D-оптимальный план с шестью наблюдениями (согласно работе [1]) имеет вид

$$\varepsilon^0(6) = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ 2, & 2, & 1, & 1 \end{Bmatrix}, \ \left|M\left(\varepsilon^0(6)\right)\right| = 192.$$

В точках $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, $x^{(4)}$ единичного куба дисперсии наблюдений для обоих планов равны 1, однако определители информационных матриц этих планов разные. Это вызвано тем, что дисперсии равноточных наблюдений не принадлежат классам функций, удовлетворяющих неравенствам (11), (13), (15), (17).

Для модели (9) неравноточных наблюдений, имеющих дисперсии

$$d(x_1, x_2, x_3) \geq \frac{\sigma^2}{3}\left(x_1^2 + x_2^2 + x_3^2\right), \ \sigma^2 \neq 0, \ |x_1| + |x_2| + |x_3| \neq 0, \tag{24}$$

как частный случай теоремы 3 получаем следующую теорему.

Теорема 5. *Для модели (9) с некоррелированными ошибками наблюдений, имеющими средние значения* 0 *и дисперсии (24), можно построить шесть непрерывных D-оптимальных планов экспериментов:*

$$\varepsilon^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix}, \ \varepsilon^0 = \begin{Bmatrix} x^{(5)}, & x^{(6)}, & x^{(7)}, & x^{(8)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix}, \ \varepsilon^0 = \begin{Bmatrix} x^{(1)}, & x^{(4)}, & x^{(5)}, & x^{(8)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix},$$

$$\varepsilon^0 = \begin{Bmatrix} x^{(2)}, & x^{(3)}, & x^{(6)}, & x^{(7)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix}, \ \varepsilon^0 = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(5)}, & x^{(6)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix}, \ \varepsilon^0 = \begin{Bmatrix} x^{(3)}, & x^{(4)}, & x^{(7)}, & x^{(8)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix}.$$

Точки спектров каждого из этих планов лежат на одной и той же грани единичного куба.

Теорему 5 можно применять для построения точных D-оптимальных планов экспериментов с числом наблюдений n, кратным 4, т. е. $n = 4s$, $s \geq 1$. Точки спектров таких планов такие же, как и для непрерывных планов, но в каждой точке надо проводить по s наблюдений.

В теореме 5 во всех вершинах единичного куба дисперсии наблюдений принимают одно и то же значение.

Покажем, что в принципе нельзя обобщить теорему 5 на случай разных значений дисперсий наблюдений в вершинах единичного куба. Проиллюстрируем это, например, для плана ε с точками спектра $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, $x^{(4)}$ и дисперсиями в этих точках $d_1 = 2$, $d_2 = d_3 = d_4 = 1$. Информационная матрица плана

$$\varepsilon = \begin{Bmatrix} x^{(1)}, & x^{(2)}, & x^{(3)}, & x^{(4)} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix} \tag{25}$$

равна

$$M(\varepsilon) = \frac{1}{4}\begin{pmatrix} 3,5 & -0,5 & -0,5 \\ -0,5 & 3,5 & -0,5 \\ -0,5 & -0,5 & 3,5 \end{pmatrix}.$$

Обратная матрица

$$M^{-1}(\varepsilon) = 4\begin{pmatrix} 0,3 & 0,05 & 0,05 \\ 0,05 & 0,3 & 0,05 \\ 0,05 & 0,05 & 0,3 \end{pmatrix}.$$

Если план ε претендует на D-оптимальность, то согласно критерию оптимальности Кифера – Вольфовица [6] должно выполняться неравенство

$$\frac{1}{d(x_1, x_2, x_3)}(x_1, x_2, x_3)M^{-1}(\varepsilon)\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \leq 3. \qquad (26)$$

В (26) в точках спектра плана (25) должно выполняться равенство. Однако в точке $x^{(1)}$, в которой дисперсия равна 2,0, неравенство (26) не обращается в равенство, так как левая часть (26) в этой точке равна 2,4.

Библиографические ссылки

1. Moyssiadis C, Kounias S. Exact *D*-optimal *N* observations 2^k designs of resolution III, when $N \doteq 1$ or 2 mod 4. *Series Statistics*. 1983;14(3):367–379. DOI: 10.1080/02331888308801711.

2. Кирлица ВП. Точные *D*-оптимальные планы экспериментов для линейной модели парной регрессии. *Вестник БГУ. Серия 1. Физика. Математика. Информатика*. 2016;2:116–122.

3. Kirlitsa VP. Exact *D*-optimal designs of experiments for linear multiple model with linear variance of observations. In: *Computer data analysis and modeling. Robustness and computer intensive methods. Proceedings of the Seventh International conference*; *2004 September 6–10; Minsk, Belarus. Volume 1.* Minsk: Belarusian State University; 2004. p. 165–167.

4. Кирлица ВП. Точные *D*-оптимальные планы экспериментов для линейной множественной регрессии с неравноточными наблюдениями. *Журнал Белорусского государственного университета. Математика. Информатика*. 2017;3:53–59.

5. Кирлица ВП. Построение *D*-оптимальных планов экспериментов для линейной множественной регрессии с неравноточными наблюдениями. *Журнал Белорусского государственного университета. Математика. Информатика*. 2019;2:27–33.

6. Ермаков СМ, Жиглявский АА. *Математическая теория оптимального эксперимента*. Москва: Наука; 1987. 320 с.

References

1. Moyssiadis C, Kounias S. Exact *D*-optimal *N* observations 2^k designs of resolution III, when $N \doteq 1$ or 2 mod 4. *Series Statistics*. 1983;14(3):367–379. DOI: 10.1080/02331888308801711.

2. Kirlitsa VP. Exact *D*-optimal designs of experiments for linear model of pair regression. *Vestnik BGU. Seriya 1. Fizika. Matematika. Informatika.* 2016;2:116–122. Russian.

3. Kirlitsa VP. Exact *D*-optimal designs of experiments for linear multiple model with linear variance of observations. In: *Computer data analysis and modeling. Robustness and computer intensive methods. Proceedings of the Seventh International conference*; *2004 September 6–10; Minsk, Belarus. Volume 1.* Minsk: Belarusian State University; 2004. p. 165–167.

4. Kirlitsa VP. Exact *D*-optimal designs of experiments for linear multiple regression with heteroscedastic observations. *Journal of the Belarusian State University. Mathematics and Informatics.* 2017;3:53–59. Russian.

5. Kirlitsa VP. Construction *D*-optimal designs of experiments for linear multiple regression with heteroscedastic observations. *Journal of the Belarusian State University. Mathematics and Informatics.* 2019;2:27–33. Russian.

6. Ermakov SM, Zhiglyavskii AA. *Matematicheskaya teoriya optimal'nogo eksperimenta* [The mathematical theory of optimal design]. Moscow: Nauka; 1987. 320 p. Russian.

Статья поступила в редколлегию 12.02.2020.
Received by editorial board 12.02.2020.