Журнал Белорусского государственного университета. Математика. Информатика. 2020;3:80–85
Journal of the Belarusian State University. Mathematics and Informatics. 2020;3:80–85

УДК 519.4

D-ОПТИМАЛЬНЫЕ ПЛАНЫ ЭКСПЕРИМЕНТОВ ДЛЯ ТРИГОНОМЕТРИЧЕСКОЙ РЕГРЕССИИ НА ОТРЕЗКЕ 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) *D*-оптимальных планов экспериментов для тригонометрической 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 *D*-оптимальные планы экспериментов. Для тригонометрической регрессии с тремя факторами построены непрерывные *D*-оптимальные планы экспериментов с различными типами неравноточных наблюдений. Для каждого из этих типов выделен свой собственный класс функций, описывающих изменение дисперсии наблюдений.

Ключевые слова: непрерывные *D*-оптимальные планы экспериментов; тригонометрическая регрессия; равноточные наблюдения; неравноточные наблюдения.

D-OPTIMAL DESIGNS OF EXPERIMENTS FOR TRIGONOMETRIC REGRESSION ON INTERVAL WITH HETEROSCEDASTIC OBSERVATIONS

V. P. KIRLITSA[a]

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

In article the problem of construction continuous (number of observations is not fixed) *D*-optimal designs of experiments for trigonometric regression in a case when variance of errors of observations depend on a point in which is made is investigated. Class of functions which describe change variance of heteroscedastic observations is defined for which it is possible construct continuous *D*-optimal designs of experiments. For trigonometric regression with three factors it is constructed continuous *D*-optimal designs of experiments with different types heteroscedastic observations. For each of these types the own class of functions describing change variance of observations is defined.

Keywords: continuous *D*-optimal designs of experiments; trigonometric regression; homoscedastic observations; heteroscedastic observations.

Образец цитирования:
Кирлица ВП. *D*-оптимальные планы экспериментов для тригонометрической регрессии на отрезке с неравноточными наблюдениями. *Журнал Белорусского государственного университета. Математика. Информатика.* 2020;3:80–85.
https://doi.org/10.33581/2520-6508-2020-3-80-85

For citation:
Kirlitsa VP. *D*-optimal designs of experiments for trigonometric regression on interval with heteroscedastic observations. *Journal of the Belarusian State University. Mathematics and Informatics.* 2020;3:80–85. Russian.
https://doi.org/10.33581/2520-6508-2020-3-80-85

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

Рассмотрим модель наблюдений

$$y_j = \theta_1 + \sum_{s=1}^{k} \left(\theta_{2s} \cos\left(sx_j\right) + \theta_{2s+1} \sin\left(sx_j\right) \right) + \varepsilon\left(x_j\right),\ j \geq 2k+1, \tag{1}$$

где y_j – наблюдаемые переменные; x_j – контролируемые переменные, принадлежащие интервалу $[0, 2\pi)$; θ_1, θ_{2s}, θ_{2s+1} – неизвестные параметры, подлежащие оцениванию; $\varepsilon\left(x_j\right)$ – некоррелированные ошибки наблюдений со средним значением, равным нулю, и дисперсией, зависящей от точки наблюдения:

$$D\left\{\varepsilon\left(x_j\right)\right\} = d\left(x_j\right) > 0,\ j \geq 2k+1. \tag{2}$$

Функция $d(x)$ – некоторая положительная функция.

В монографиях [1; 2] для модели равноточных наблюдений ($d(x) = \sigma^2$, $\sigma^2 > 0$) доказано, что план экспериментов

$$\varepsilon_n^0 = \begin{Bmatrix} x_1^0, \ldots, x_n^0 \\ \dfrac{1}{n}, \ldots, \dfrac{1}{n} \end{Bmatrix} \tag{3}$$

является непрерывным D-оптимальным. В (3) точки спектра плана

$$x_i^0 = \frac{2\pi(i-1)}{n},\ i = \overline{1, n},\ n \geq 2k+1. \tag{4}$$

Для частного случая $n = 2k + 2$ оптимальный план (3) впервые был построен в публикации [3]. В статье [4] для модели тригонометрической регрессии (1) с равноточными наблюдениями, в которой объясняющая переменная изменяется в интервале $[-a, a]$, $0 < a \leq \pi$, рассматривается построение оптимальных планов экспериментов для оценки неизвестных параметров модели. В этой статье отмечено, что структура оптимального плана экспериментов существенно зависит от размера интервала планирования. Также в ней приведена обширная библиография работ, посвященных исследованию тригонометрических моделей наблюдений.

В данной статье покажем, что план (3) остается D-оптимальным для определенного класса неравноточных наблюдений.

Теорема 1. *Для модели наблюдений (1), (2) план (3) остается непрерывным D-оптимальным для неравноточных наблюдений, дисперсии которых $d(x)$ удовлетворяют неравенству*

$$d(x) \geq \sigma^2,\ \sigma \neq 0, \tag{5}$$

где равенство выполняется в точках (4) спектра плана (3).

Доказательство. Теорема будет доказана, если будет установлено, что для плана (3) справедливы условия теоремы эквивалентности Кифера – Вольфовица [1] для D-оптимальных планов:

$$\frac{1}{d(x)} f'(x) M^{-1}\left(\varepsilon_n^0\right) f(x) \leq 2k+1,\ x \in [0, 2\pi), \tag{6}$$

где $f(x) = (1, \cos x, \sin x, \ldots, \cos kx, \sin kx)$ – вектор базисных функций; $M\left(\varepsilon_n^0\right)$ – информационная матрица плана (3). В неравенстве (6) равенство должно выполняться в точках (4) спектра плана (3).

Вначале докажем, что информационная матрица плана (3) при ограничении (5) имеет диагональный вид:

$$M\left(\varepsilon_n^0\right) = \sigma^{-2} \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & \frac{1}{2} & 0 & \ldots & 0 \\ 0 & 0 & \frac{1}{2} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \frac{1}{2} \end{pmatrix}. \tag{7}$$

В работах [1; 2] приводится доказательство утверждения (7) для равноточных наблюдений. Однако оно фрагментарно и в монографии [1, с. 123] дано для четных значений n. Представим полное доказательство (7) для любых n.

Будем использовать формулы Эйлера и формулу для суммы геометрической прогрессии. Пусть m и l – натуральные числа, принимающие значения от 1 до k, причем $m \neq l$. Диагональные элементы информационной матрицы, начиная со второго, вычислим, используя формулы Эйлера.

$$\sum_{j=1}^{n} \frac{1}{n} \cos^2 \frac{2m(j-1)\pi}{n} = \frac{1}{2n} \sum_{j=1}^{n} \left(1 + \cos \frac{4m(j-1)\pi}{n}\right) = \frac{1}{2} + \frac{1}{4n} \sum_{j=1}^{n} \left(e^{i\frac{4m(j-1)\pi}{n}} + e^{-i\frac{4m(j-1)\pi}{n}}\right) =$$

$$= \frac{1}{2} + \frac{1}{4n}\left(\frac{1-e^{i4m\pi}}{1-e^{i\frac{4m\pi}{n}}} + \frac{1-e^{-i4m\pi}}{1-e^{-i\frac{4m\pi}{n}}}\right) = \frac{1}{2} + \frac{1}{4n}\left(\frac{\left(1-e^{i4m\pi}\right)\left(1-e^{-i\frac{4m\pi}{n}}\right) + \left(1-e^{-i4m\pi}\right)\left(1-e^{i\frac{4m\pi}{n}}\right)}{2-2\cos\frac{4m\pi}{n}}\right) =$$

$$= \frac{1}{2} + \frac{1}{4n} \frac{2 - 2\cos 4m\pi - 2\cos\frac{4m\pi}{n} + 2\cos 4m\pi\left(1 - \frac{1}{n}\right)}{2 - 2\cos\frac{4m\pi}{n}} = \frac{1}{2},$$

$$\sum_{j=1}^{n} \frac{1}{n} \sin^2 \frac{2m(j-1)\pi}{n} = \frac{1}{2n} \sum_{j=1}^{n} \left(1 - \cos \frac{4m(j-1)\pi}{n}\right) = \frac{1}{2} - \frac{1}{2n} \sum_{j=1}^{n} \cos \frac{4m(j-1)\pi}{n} = \frac{1}{2}.$$

Недиагональные элементы равны нулю. Используя формулы Эйлера, получаем

$$\sum_{j=1}^{n} \frac{1}{n} \cos \frac{2m(j-1)\pi}{n} = \frac{1}{2n}\left(\frac{1-e^{i2m\pi}}{1-e^{i\frac{2m\pi}{n}}} + \frac{1-e^{-i2m\pi}}{1-e^{-i\frac{2m\pi}{n}}}\right) = \frac{1}{2n} \frac{-2\cos\frac{2m\pi}{n} + 2\cos\frac{2m\pi}{n}}{2 - 2\cos\frac{2m\pi}{n}} = 0,$$

$$\sum_{j=1}^{n} \frac{1}{n} \sin \frac{2m(j-1)\pi}{n} = \frac{1}{2ni}\left(\frac{1-e^{i2m\pi}}{1-e^{i\frac{2m\pi}{n}}} - \frac{1-e^{-i2m\pi}}{1-e^{-i\frac{2m\pi}{n}}}\right) = \frac{1}{2n} \frac{2\sin\frac{2m\pi}{n} + 2\sin 2m\pi\left(1 - \frac{1}{n}\right)}{2 - 2\cos\frac{2m\pi}{n}} = 0,$$

$$\sum_{j=1}^{n} \frac{1}{n} \sin \frac{2m(j-1)\pi}{n} \cos \frac{2l(j-1)\pi}{n} = \frac{1}{2n}\left(\sum_{j=1}^{n} \sin \frac{2(m+l)(j-1)\pi}{n} + \sum_{j=1}^{n} \sin \frac{2(m-l)(j-1)\pi}{n}\right) = 0.$$

В силу (5), (7) имеем

$$\frac{1}{d(x)} f'(x) M^{-1}\left(\varepsilon_n^0\right) f(x) = \frac{\sigma^2}{d(x)}\left(1 + 2\left(\sin^2 x + \cos^2 x + \ldots + \sin^2 kx + \cos^2 kx\right)\right) =$$

$$= \frac{\sigma^2}{d(x)}(1 + 2k) \leq 2k + 1, \tag{8}$$

причем в неравенстве (8) равенство достигается в точках (4) спектра плана (3). Условия теоремы эквивалентности Кифера – Вольфовица выполнены. Теорема 1 доказана.

Множество функций, удовлетворяющих неравенству (5), обширно. Ему соответствуют равноточные наблюдения с дисперсиями $d(x) = \sigma^2$ и, например, функции

$$d(x) = \sigma^2 + \beta \cdot \left|\sin(2nx)\right|, \ \beta > 0; \ d(x) = \sigma^2 + \beta \cdot \left((x) \operatorname{mod} \Delta\right), \ \beta > 0, \ \Delta = \frac{2\pi}{n};$$

$$d(x) = \sigma^2 \sqrt{\frac{\Delta^2}{4} - \left(x - s\Delta + \frac{\Delta}{2}\right)^2}, \ x \in \left[(s-1)\Delta, s\Delta\right], \ s = \overline{1, n}, \ \Delta = \frac{2\pi}{n}.$$

Можно предложить и ряд других функций.

Замечание. Доказанная теорема верна и для планов экспериментов (3), в которых точки спектра (4) сдвинуты на угол φ, т. е.

$$x_i^0 = \frac{2\pi(i-1)}{n} + \varphi,\ i = \overline{1, n},\ n \ge 2k+1.$$

Доказательство проводится по той же схеме с незначительными изменениями.

Особенность построенных планов экспериментов (3) с неравноточными наблюдениями состоит в том, что в точках (4) спектров этих планов дисперсии наблюдений принимают одни и те же значения, равные σ^2. Возникает вопрос, можно ли построить планы экспериментов для неравноточных наблюдений, для которых дисперсии наблюдений в точках спектра планов экспериментов принимали бы различные значения. Ответ утвердительный. Для частного случая модели наблюдений (1)

$$y_j = \theta_1 + \theta_2 \cos x_j + \theta_3 \sin x_j + \varepsilon(x_j),\ j \ge 3, \tag{9}$$

можно сконструировать так называемые насыщенные планы экспериментов с тремя наблюдениями, для которых в точках спектра $0, \frac{2\pi}{3}, \frac{4\pi}{3}$ дисперсии наблюдений будут принимать положительные значения d_1, d_2, d_3 соответственно.

Теорема 2. *Для модели наблюдений (9) с некоррелированными ошибками наблюдений, имеющими среднее значение, равное нулю, и дисперсии* $d(x)$*, удовлетворяющие неравенству*

$$\begin{aligned} d(x) \ge \frac{1}{9}\Big(&(4d_1 + d_2 + d_3)\cos^2 x + 3(d_2 + d_3)\sin^2 x + \sqrt{3}(d_3 - d_2)\sin 2x + \\ &+ 2(2d_1 - d_2 - d_3)\cos x + 2\sqrt{3}(d_2 - d_3)\sin x + d_1 + d_2 + d_3\Big), \end{aligned} \tag{10}$$

в котором равенство достигается в точках $0, \frac{2\pi}{3}, \frac{4\pi}{3}$, *насыщенный* D*-оптимальный план экспериментов имеет вид*

$$\varepsilon_3^0 = \begin{Bmatrix} 0, & \frac{2\pi}{3}, & \frac{4\pi}{3} \\ \frac{1}{3}, & \frac{1}{3}, & \frac{1}{3} \end{Bmatrix}. \tag{11}$$

Доказательство. Для D-оптимального плана ε_3^0 по теореме эквивалентности Кифера – Вольфовица [1] должно выполняться неравенство

$$\frac{1}{d(x)}(1, \cos x, \sin x) M^{-1}(\varepsilon_3^0) \begin{pmatrix} 1 \\ \cos x \\ \sin x \end{pmatrix} \le 3,\ x \in [0, 2\pi), \tag{12}$$

причем (12) должно обращаться в равенство в точках спектра плана (11). Проверим справедливость этого утверждения.

Информационная матрица плана ε_3^0 равна

$$M(\varepsilon_3^0) = \frac{1}{3}\left(\frac{1}{d_1}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}(1, 1, 0) + \frac{1}{d_2}\begin{pmatrix} 1 \\ -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}\left(1, -\frac{1}{2}, \frac{\sqrt{3}}{2}\right) + \frac{1}{d_3}\begin{pmatrix} 1 \\ -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}\left(1, -\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)\right) = \frac{1}{3}\begin{pmatrix} a & b & c \\ b & k & e \\ c & e & m \end{pmatrix},$$

где

$$a = \frac{1}{d_1} + \frac{1}{d_2} + \frac{1}{d_3},\ b = \frac{1}{d_1} - \frac{1}{2d_2} - \frac{1}{2d_3},\ c = \frac{\sqrt{3}}{2d_2} - \frac{\sqrt{3}}{2d_3},\ e = -\frac{\sqrt{3}}{4d_2} + \frac{\sqrt{3}}{4d_3}, \tag{13}$$

$$k = \frac{1}{d_1} + \frac{1}{4d_2} + \frac{1}{4d_3}, \; m = \frac{3}{4d_2} + \frac{3}{4d_3}.$$

Матрица, обратная к матрице $M\left(\varepsilon_3^0\right)$, имеет вид

$$M^{-1}\left(\varepsilon_3^0\right) = \frac{3}{akm + 2bce - c^2k - e^2a - b^2m}\begin{pmatrix} km - e^2, & ce - bm, & be - kc \\ ce - bm, & am - c^2, & bc - ae \\ be - kc, & bc - ae, & ak - b^2 \end{pmatrix}. \tag{14}$$

Используя (13), преобразуем элементы матрицы (14).

$$km - e^2 = \frac{3\left(d_1 + d_2 + d_3\right)}{4d_1d_2d_3}, \; ce - bm = \frac{3\left(2d_1 - d_2 - d_3\right)}{4d_1d_2d_3}, \; be - kc = \frac{3\sqrt{3}\left(d_2 - d_3\right)}{4d_1d_2d_3},$$

$$am - c^2 = \frac{3\left(d_2 + 4d_1 + d_3\right)}{4d_1d_2d_3}, \; bc - ae = \frac{3\sqrt{3}\left(d_3 - d_2\right)}{4d_1d_2d_3}, \; ak - b^2 = \frac{9\left(d_3 + d_2\right)}{4d_1d_2d_3}, \tag{15}$$

$$akm + 2bce - c^2k - e^2a - b^2m = \frac{27}{4d_1d_2d_3}.$$

Разрешая неравенство (12) относительно функции $d(x)$ и учитывая соотношения (14), (15), получаем требуемое неравенство (10). Легко проверить, что неравенство (10) обращается в равенство в точках спектра плана (11). Условия теоремы эквивалентности Кифера – Вольфовица выполнены. Теорема 2 доказана.

Есть предположение, что точки спектра плана (11) можно сдвинуть на угол φ, т. е.

$$x_1^0 = \varphi, \; x_2^0 = \frac{2\pi}{3} + \varphi, \; x_3^0 = \frac{4\pi}{3} + \varphi, \tag{16}$$

при этом изменится вид функций $d(x)$ в неравенстве (10). Расчеты, проведенные для угла $\varphi = \frac{\pi}{6}$, подтверждают это. Пусть в соответствии с (16) $D_1 = d\left(x_1^0\right)$, $D_2 = d\left(x_2^0\right)$, $D_3 = d\left(x_3^0\right)$. Получим выражение для дисперсии $d(x)$ наблюдений, соответствующих оптимальному плану экспериментов с точками спектра (16), равными весами $\frac{1}{3}$ и углом $\varphi = \frac{\pi}{6}$:

$$d(x) \geq \frac{1}{9}\Big(3\left(D_1 + D_2\right)\cos^2 x + \left(D_1 + D_2 + 4D_3\right)\sin^2 x +$$

$$+ \sqrt{3}\left(D_1 - D_2\right)\sin 2x + 2\sqrt{3}\left(D_1 - D_2\right)\cos x + 2\left(D_1 + D_2 - 2D_3\right)\sin x + D_1 + D_2 + D_3\Big). \tag{17}$$

В точках (16) спектра плана с углом $\varphi = \frac{\pi}{6}$ неравенство (17) обращается в равенство. Было бы интересно установить закон преобразования значений d_1, d_2, d_3 в значения D_1, D_2, D_3 для произвольных углов φ.

Теперь исследуем вопрос, можно ли для модели (9) с четырьмя наблюдениями построить оптимальный план

$$\varepsilon_4^0 = \begin{Bmatrix} 0, & \frac{\pi}{2}, & \pi, & \frac{3\pi}{2} \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{Bmatrix} \tag{18}$$

с неравноточными наблюдениями, для которого дисперсии наблюдений в точках спектра плана (18) будут различными.

Введем для плана (18) обозначения для дисперсий наблюдений: $d_1 = d(0°)$, $d_2 = d(90°)$, $d_3 = d(180°)$, $d_4 = d(270°)$. Проводя для плана (18) рассуждения, аналогичные тем, что были изложены при доказательстве теоремы 2, получим неравенство, которому должна удовлетворять дисперсия $d(x)$ наблюдений:

$$d(x) \geq \frac{1}{3(d_1 + d_2 + d_3 + d_4)}\Big((d_2 d_3 + 4d_1 d_3 + d_1 d_2 + d_3 d_4 + d_1 d_4)\cos^2 x +$$
$$+ (4d_2 d_4 + d_1 d_4 + d_1 d_2 + d_2 d_3 + d_3 d_4)\sin^2 x + (d_3 d_4 - d_1 d_4 - d_2 d_3 + d_1 d_2)\sin 2x +$$
$$+ 2(d_1 d_2 - d_2 d_3 + d_1 d_4 - d_3 d_4)\cos x + 2(d_1 d_2 - d_1 d_4 + d_2 d_3 - d_3 d_4)\sin x +$$
$$+ d_1 d_2 + d_2 d_3 + d_1 d_4 + d_3 d_4\Big). \qquad (19)$$

Условия выполнения в (19) равенства в точках 0°, 90°, 180°, 270° приводят к системе уравнений

$$4(d_2 + d_3 + d_4) = 3(d_1 + d_2 + d_3 + d_4),\ 4(d_1 + d_3 + d_4) = 3(d_1 + d_2 + d_3 + d_4),$$
$$4(d_1 + d_2 + d_4) = 3(d_1 + d_2 + d_3 + d_4),\ 4(d_1 + d_2 + d_3) = 3(d_1 + d_2 + d_3 + d_4). \qquad (20)$$

Решая систему уравнений (20), получаем $d_1 = d_2 = d_3 = d_4 = \sigma^2$, $\sigma \neq 0$. Таким образом, неравенство (19) сводится к ранее полученному неравенству (5). Следовательно, для модели (9) с четырьмя наблюдениями можно строить оптимальные планы экспериментов, дисперсии наблюдений которых удовлетворяют неравенству (5).

Библиографические ссылки

1. Ермаков СМ, Жиглявский АА. *Математическая теория оптимального эксперимента*. Москва: Наука; 1987. 320 с.
2. Федоров ВВ. *Теория оптимального эксперимента.* Москва: Наука; 1971. 312 с.
3. Hoel PG. Minimax designs in two dimensional regression. *Annals of Mathematical Statistics.* 1965;36(4):1097–1106. DOI: 10.1214/aoms/1177699984.
4. Dette H, Melas VB. Optimal designs for estimating individual coefficients in Fourier regression models. *Annals of Statistics.* 2003;31(5):1669–1692. DOI: 10.1214/aos/1065705122.

References

1. Ermakov SM, Zhiglyavskii AA. *Matematicheskaya teoriya optimal'nogo eksperimenta* [The mathematical theory of optimal design]. Moscow: Nauka; 1987. 320 p. Russian.
2. Fedorov VV. *Teoriya optimal'nogo eksperimenta* [Optimal experiment theory]. Moscow: Nauka; 1971. 312 p. Russian.
3. Hoel PG. Minimax designs in two dimensional regression. *Annals of Mathematical Statistics.* 1965;36(4):1097–1106. DOI: 10.1214/aoms/1177699984.
4. Dette H, Melas VB. Optimal designs for estimating individual coefficients in Fourier regression models. *Annals of Statistics.* 2003;31(5):1669–1692. DOI: 10.1214/aos/1065705122.

Статья поступила в редколлегию 29.06.2020.
Received by editorial board 29.06.2020.