УДК 519.4

В. П. КИРЛИЦА

О СТРУКТУРЕ ТОЧНЫХ *A*-ОПТИМАЛЬНЫХ ПЛАНОВ ЭКСПЕРИМЕНТОВ ДЛЯ ЛИНИИ РЕГРЕССИИ С КВАДРАТИЧНЫМ ИЗМЕНЕНИЕМ ДИСПЕРСИИ НАБЛЮДЕНИЙ

Исследована структура точных A-оптимальных планов экспериментов для линии регрессии с контролируемой переменной $x \in [-1, 1]$ и квадратичным изменением дисперсии наблюдений $a_2x^2 + a_1x + a_0 = a_0\left(k_2x^2 + k_1x + 1\right)$, $a_0 > 0$, $k_2 = \frac{a_2}{a_0}$, $k_1 = \frac{a_1}{a_0}$. Показано, что точки спектра таких планов должны находиться на концах интервала $[-1, 1]$. Число наблюдений в этих точках зависит от коэффициентов k_1, k_2. Получена формула, определяющая число наблюдений в точке -1 в точном A-оптимальном плане экспериментов. Доказано, что для дисперсии наблюдений с $k_1 = 0$ и $-1 < k_2 \le 1$ точные A-оптимальные планы экспериментов для линии регрессии являются одновременно и D-оптимальными.

Ключевые слова: линия регрессии; неравноточные наблюдения; D- и A-оптимальные планы экспериментов; структура точных A-оптимальных планов; квадратичное изменение дисперсии.

The structure of exact A-optimal designs of experiment for a line of regress with a controllable variable $x \in [-1, 1]$ and square-law change of variance observations $a_2x^2 + a_1x + a_0 = a_0\left(k_2x^2 + k_1x + 1\right)$, $a_0 > 0$, $k_2 = \frac{a_2}{a_0}$, $k_1 = \frac{a_1}{a_0}$, is investigated. It is shown that points of a spectrum of such designs should be on the interval ends of $[-1, 1]$. The number of observations in these points depends on factors k_1, k_2. Formula defining number of observations in a point -1 is resulted for exact A-optimal designs. It is proved that for variance of observation with $k_1 = 0$ and $-1 < k_2 \le 1$ exact A-optimal designs are simultaneously D-optimal.

Key words: line of regress; heteroscedastic observations; D- and A-optimal designs of experiments; structure of exact A-optimal designs; square-law change of variance.

Рассмотрим линейную модель неравноточных наблюдений

$$y_i = \theta_0 + \theta_1 x_i + \varepsilon\left(x_i\right), i = 1, 2, \ldots, n, n \ge 2, \tag{1}$$

где y_i – наблюдаемые значения; θ_0, θ_1 – неизвестные параметры; x_i – контролируемые переменные из интервала $[-1, 1]$; $\varepsilon(x_i)$ – случайные некоррелированные ошибки наблюдений с нулевым средним и дисперсией

$$D\left\{\varepsilon\left(x_i\right)\right\} = a_2x_i^2 + a_1x_i + a_0 = a_0\left(k_2x_i^2 + k_1x_i + 1\right) > 0, x_i \in \left[-1, 1\right], \tag{2}$$

где $a_0 > 0$; $k_2 = \frac{a_2}{a_0}, k_1 = \frac{a_1}{a_0}$. Кроме того, k_1, k_2 должны удовлетворять одному из условий:

$$1) \left|k_1\right| < 2\sqrt{k_2}, k_2 > 0;\ 2)\ k_1 = \pm 2\sqrt{k_2}, 0 < k_2 < 1;\ 3) \left|k_1\right| < 1 + k_2, -1 < k_2 \le 0;\ 4) \left|k_1\right| > 2\sqrt{k_2}, 0 < k_2 < 1.$$

Выполнение одного из условий 1) – 4) гарантирует выполнение неравенства (2).

Действительно, если $k_2 = 0$, то квадратичная зависимость в (2) вырождается в линейную. В этом случае (2) будет выполняться, если одновременно будут выполняться неравенства $1 - k_1 > 0$, $1 + k_1 > 0$, т. е. если $\left|k_1\right| < 1$. Это означает выполнимость ограничения 3).

При $k_2 < 0$ ветви параболы $y = k_2x^2 + k_1x + 1$ направлены вниз и неравенство (2) будет выполняться, если будут выполняться неравенства $k_2 - k_1 + 1 > 0$, $k_2 + k_1 + 1 > 0$, т. е. если $\left|k_1\right| < 1 + k_2$, $-1 < k_2 < 0$. А это и есть условие ограничений 3).

Для $k_2 > 0$ выполнимость неравенства (2) зависит от знака дискриминанта квадратичного трехчлена $k_1x^2 + k_1x + 1$. Если $D = k_1^2 - 4k_2 < 0$ и $k_2 > 0$, то неравенство (2) соблюдается при выполнении условий 1). Если $D = 0$ и $k_2 > 0$, то корни уравнения $k_2x^2 + k_1x + 1 = 0$ совпадают и равны $\frac{-k_1}{2k_2}$. Положительность дисперсии (2) на интервале $[-1, 1]$ будет гарантирована, если выполняется неравенство $\left|\frac{-k_1}{2k_2}\right| > 1$, т. е. если $\left|k_1\right| > 2k_2$. Если $D = 0$, то $k_1 = \pm 2\sqrt{k_2}$ и неравенство $\left|k_1\right| > 2k_2$ сводится к неравенству $\sqrt{k_2} > k_2$. Последнее неравенство выполняется, если $0 < k_2 < 1$. Итак, при выполнении условий 2) неравенство (2) выполняется. Наконец, если $D > 0$ и $k_2 > 0$, то в этом случае ветви параболы $y = k_2x^2 + k_1x + 1$ направлены вверх, она пересекает ось x в двух точках, которые одновременно либо отрицательные, либо положительные. Для выполнения неравенства (2) надо потребовать, чтобы эти точки лежали вне интервала $[-1, 1]$. Это выполняется, если $0 < k_2 < 1$, т. е. если выполняется ограничение 4).

Цель данной статьи – определить структуру точных A-оптимальных планов экспериментов [1] для модели наблюдений (1), (2). Среднее значение дисперсий оценок неизвестных параметров регрессионной модели наблюдений (1), полученных по A-оптимальным планам экспериментов, минимально в классе линейных несмещенных оценок [1].

В частном случае, когда $k_2 = 0$, квадратичное изменение дисперсии наблюдений (2) вырождается в линейное. Структура точных A-оптимальных планов экспериментов для этого частного случая исследована автором статьи в [2]. В данной статье рассматривается более широкий класс изменения дисперсии наблюдений (2), который включает в себя линейное изменение $(k_2 = 0)$ и равноточные наблюдения $(k_1 = k_2 = 0)$.

Квадратичное изменение дисперсии наблюдений

Докажем следующую теорему.

Теорема. *При выполнении одного из условий* 1) – 4) *структура точных A-оптимальных планов экспериментов для модели наблюдений* (1), (2) *имеет следующий вид:*

$$\varepsilon_n^0 = \begin{Bmatrix} -1; & 1 \\ n_1 - t; & n - n_1 + t \end{Bmatrix}, t = 0, \pm 1, \pm 2, \ldots; \tag{3}$$

$$\eta_{t-1} \le k_1 \le \eta_t, \eta_t = \frac{n(1+k_2)(n - 2n_1 + 2t + 1)}{2n_1(n_1 - 2t - n - 1) + n(2t + n + 1) + 2t(t+1)}, \tag{4}$$

где $n_1 = \left[\frac{n}{2}\right]$ – целая часть $\frac{n}{2}$ и $n_1 - t \in \{1, 2, \ldots, n-1\}$.

Д о к а з а т е л ь с т в о. В статье [3] было показано, что структура точных A-оптимальных планов для модели наблюдений (1), (2) должна иметь вид

$$\varepsilon_n^0 = \begin{Bmatrix} -1; & 1 \\ s^0; & n - s^0 \end{Bmatrix},$$

где

$$s^0 = \operatorname{argmin} t_n(s; k_1, k_2); \tag{5}$$

$$t_n(s; k_1, k_2) = a_0 \frac{2k_1 s + n(k_2 - k_1 + 1)}{s(n-s)}, \tag{6}$$

минимум в (5) берется по $s \in \{1, 2, \ldots, n-1\}$. Однако задача дискретной оптимизации (5), (6) не позволяет в явном виде получить структуру точных A-оптимальных планов для модели наблюдений (1), (2).

Погрузим задачу дискретной оптимизации (5), (6) в задачу непрерывной оптимизации по s, т. е. будем полагать, что в функции (6) аргумент s меняется непрерывно на интервале $(0, n)$. Производная функции

$$f_n(s; k_1, k_2) = \frac{2k_1 s + n(k_2 - k_1 + 1)}{s(n-s)} \tag{7}$$

по s равна

$$\frac{\partial f_n(s; k_1, k_2)}{\partial s} = \frac{(1 + k_1 + k_2)s^2 - (1 + k_2 - k_1)(n-s)^2}{s^2(n-s)^2}. \tag{8}$$

Поскольку $1+k_1+k_2 > 0$ и $1+ k_1- k_2 > 0$, то производная (8) обращается в ноль, если

$$\sqrt{1 + k_1 + k_2} \cdot s = \sqrt{1 + k_2 - k_1} \cdot (n - s),$$

т. е. если

$$s_n^0(k_1, k_2) = \frac{\sqrt{1 + k_2 - k_1} \cdot n}{\sqrt{1 + k_2 + k_1} + \sqrt{1 + k_2 - k_1}} = \frac{\left(\sqrt{(1+k_2)^2 - k_1^2} + k_1 - k_2 + 1\right)n}{2k_1}. \tag{9}$$

Точка (9) – это точка минимума функции (7), так как левее этой точки производная (8) отрицательна, а правее – положительна. Действительно, неравенство

$$(1 + k_1 + k_2)s^2 < (1 + k_2 - k_1)(n - s)^2$$

имеет место, если $s < s_n^0(k_1, k_2)$, а неравенство

$$(1 + k_1 + k_2)s^2 > (1 + k_2 - k_1)(n - s)^2$$

выполняется, если $s > s_n^0(k_1, k_2)$.

Точка минимума (9) – это непрерывная функция аргумента k_1, так как при $k_1 = 0$ она не имеет разрыва и в силу правила Лопиталя $s_n^0\left(0,k_2\right)=\frac{n}{2}$.

Производная по k_1 точки минимума (9) отрицательна:

$$\frac{\partial s_n^0\left(k_1,k_2\right)}{\partial k_1}=\frac{\left(1+k_2\right)\left(\sqrt{\left(1+k_2\right)^2-k_1^2}-1-k_2\right)n}{2k_1^2\sqrt{\left(1+k_2\right)^2-k_1^2}}<0 \tag{10}$$

для всех k_1, k_2, удовлетворяющих одному из ограничений 1) – 4). Убедимся в этом для ограничений 1). Остальные случаи ограничений можно рассмотреть аналогично. Выполнение 1) означает, что дискриминант $D < 0$ и ветви параболы $y = k_2x^2 + k_1x + 1$ направлены вверх, т. е. в точках –1, 1 парабола принимает положительные значения, $k_2 - k_1 + 1 > 0$, $k_2 + k_1 + 1 > 0$. Значит, $\left(1+k_2\right)^2-k_1^2>0.$ По правилу Лопиталя

$$\frac{\partial s_n^0\left(0,k_2\right)}{\partial k_1}=\frac{-n}{4\left(1+k_2\right)}<0.$$

Для $k_1 \neq 0$ неравенство (10) имеет место, если $\sqrt{\left(1+k_2\right)^2-k_1^2}<1+k_2$, т. е. если $-k_1^2<0$.

Из (10) следует, что точка минимума (9) – это строго монотонно убывающая функция по аргументу k_1. При $k_1 = 0$ функция $f_n\left(s;0,k_2\right)$ имеет минимальное значение по s, если знаменатель в (7) принимает максимальное значение по s, т. е. если $s_n^0\left(0,k_2\right)=\frac{n}{2}$.

В задаче целочисленной оптимизации $f_n\left(s;k_1,k_2\right)$ по s точка минимума (5) будет также монотонно убывать с ростом значения k_1, но в отличие от непрерывного случая это изменение будет происходить ступенчато, т. е.

$$s_n^0\left(k_1,k_2\right)=n_1-t,\eta_{t-1}\le k_1\le\eta_t, t=0,\pm1,\pm2,\ldots, \tag{11}$$

где $n_1=\left[\frac{n}{2}\right]$ и $k_1=\eta_t$ – это корень относительно k_1 уравнения

$$t_n\left(n_1-t;k_1,k_2\right)=t_n\left(n_1-t-1;k_1,k_2\right). \tag{12}$$

Решая (12) относительно k_1, получаем

$$k_1=\eta_t=\frac{n\left(1+k_2\right)\left(n-2n_1+2t+1\right)}{2n_1\left(n_1-2t-n-1\right)+n\left(2t+n+1\right)+2t\left(t+1\right)}. \tag{13}$$

Значение (13) совпадает с (4), что и завершает доказательство теоремы.

Исследуем теперь, насколько робастными (устойчивыми) являются точные A-оптимальные планы экспериментов для линии регрессии (1) с равноточными наблюдениями $\left(k_1=k_2=0\right)$ относительно изменения дисперсии наблюдений (2).

В статье [4] было доказано, что для равноточных наблюдений (1) точные A-оптимальные планы экспериментов имеют структуру для четных $n = 2s$:

$$\varepsilon_{2s}^0=\begin{Bmatrix}-1, & 1\\ s, & s\end{Bmatrix}, \tag{14}$$

для нечетных $n = 2s + 1$:

$$\varepsilon_{2s+1}^0=\begin{Bmatrix}-1, & 1\\ s, & s+1\end{Bmatrix},\ \varepsilon_{2s+1}^0=\begin{Bmatrix}-1, & 1\\ s+1 & s\end{Bmatrix}. \tag{15}$$

Используя доказанную теорему и полагая в (3), (4) $t = 0$, получаем, что план (14) инвариантен относительно изменения дисперсии наблюдений (2), если

$$-\frac{1+k_2}{s}\le k_1\le\frac{1+k_2}{s}, k_2>-1. \tag{16}$$

Аналогично, полагая в (3), (4) $t = 0$, получаем, что первый из планов в (15) инвариантен относительно изменения дисперсии наблюдений (2), если

$$0\le k_1\le\frac{\left(2s+1\right)\left(1+k_2\right)}{s^2+s+1}, k_2>-1, \tag{17}$$

а для инвариантности второго из планов в (15) достаточно,чтобы выполнялось неравенство (18), полагая в (3), (4) $t = -1$:

$$-\frac{(2s+1)(1+k_2)}{s^2+s+1} \le k_1 \le 0, k_2 > -1. \tag{18}$$

Естественно, в (16) – (18) k_1, k_2 должны удовлетворять одному из ограничений 1) – 4).

Однако полной робастности планов (14), (15) относительно изменения дисперсии наблюдений, описываемого соотношениями (16) – (18), нет. Во-первых, интервал изменения k_1 в (16) шире, чем в интервалах (17), (18) для $s \ge 2$. Во-вторых, для нечетных n только один из планов (15) остается робастным в зависимости от того, положительное значение k_1 или отрицательное. Очевидно, что при $k_1 = 0$, $k_2 > -1$ A-оптимальные планы экспериментов (14), (15) для равноточных наблюдений останутся полностью неизменными (робастными) относительно изменения дисперсии наблюдений:

$$D\{y\} = a_0\left(k_2x^2+1\right), \ a_0 > 0, k_2 > -1, \ -1 \le x \le 1. \tag{19}$$

Для того чтобы планы (14), (15) с дисперсией наблюдений (19) были также и D-оптимальными, надо чтобы в соответствии с [5] дисперсия (19) принадлежала классу функций $d(x)$:

$$d(x) \ge d\frac{1+x^2}{2}, d(-1) = d(1) = d. \tag{20}$$

Для функции (19) значение $d = a_0(k_2+1)$ и неравенство (20) для этой функции сводится к неравенству

$$(k_2-1)x^2 \ge k_2 - 1, x \in [-1, 1],$$

которое выполняется, если $-1 < k_2 \le 1$.

Итак, можно утверждать, что планы экспериментов (14), (15) являются одновременно D- и A-оптимальными, робастными к изменению дисперсии наблюдений (19) для $-1 < k_2 \le 1$.

Как правило, точные значения k_1, k_2 исследователю неизвестны. Однако если есть априорная информация о том, что эти коэффициенты принадлежат некоторому замкнутому, ограниченному множеству B и на этом множестве k_1, k_2 имеют определенную плотность распределения вероятностей, то можно вычислить вероятность применения A-оптимальных планов экспериментов, определяемых формулами (3), (4).

Например, при $n = 4$, предполагая, что k_1, k_2 имеют равномерное распределение вероятностей на множестве $-1{,}4 \le k_1 \le 1{,}4$; $0{,}5 \le k_2 \le 4$, следует использовать A-оптимальные планы

$$\varepsilon_4^0 = \begin{Bmatrix} -1, & 1 \\ 1, & 3 \end{Bmatrix}, \varepsilon_4^0 = \begin{Bmatrix} -1, & 1 \\ 3, & 1 \end{Bmatrix} \tag{21}$$

с равными вероятностями 0,043 112 244 и план

$$\varepsilon_4^0 = \begin{Bmatrix} -1, & 1 \\ 2, & 2 \end{Bmatrix} \tag{22}$$

с вероятностью 0,913 775 51. Поскольку вероятность применения плана (22) намного больше, чем для планов (21), то план (22) в данной ситуации предпочтительнее для A-оптимального оценивания неизвестных параметров модели (1).

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Федоров В. В. Теория оптимального эксперимента. М., 1968.
2. Кирлица В. П. О структуре точных D- и A-оптимальных планов экспериментов для линии регрессии с неравноточными наблюдениями // Вестн. БГУ. Сер. 1, Физика. Математика. Информатика. 2012. № 3. С. 78–81.
3. Акинфин А. В., Кирлица В. П. Анализ устойчивости точных A-оптимальных планов экспериментов для линии регрессии при квадратичном изменении дисперсии наблюдений // Вестн. Белорус. ун-та. Сер. 1, Физика. Математика. Информатика. 1998. № 3. С. 49–51.
4. Кирлица В. П. Точные D- и A-оптимальные планы для линии регрессии // Вестн. Белорус. ун-та. Сер. 1, Физика. Математика. Механика. 1990. № 2. С. 36–38.
5. Кирлица В. П. D-оптимальные планы экспериментов для линии регрессии, робастные относительно изменения дисперсии наблюдений // Вестн. БГУ. Сер. 1, Физика. Математика. Информатика. 2008. № 3. С. 89–92.

Поступила в редакцию 18.12.2013.

Валерий Петрович Кирлица – кандидат физико-математических наук, доцент кафедры математического моделирования и анализа данных.