

А.В. АКИНФИН, В.П. КИРЛИЦА, Е.А. ЧЕРКАС

НЕПРЕРЫВНЫЕ D -ОПТИМАЛЬНЫЕ ПЛАНЫ ЭКСПЕРИМЕНТОВ ДЛЯ КВАДРАТИЧНОЙ РЕГРЕССИИ С ЛИНЕЙНЫМ ИЗМЕНЕНИЕМ ДИСПЕРСИИ НАБЛЮДЕНИЙ

D -optimal continuous design of experiments for quadratic regression under linear variance of observations has been constructed. Stability of D -optimal design of experiments for homoscedastical observations respect linear variance of observations was investigated.

Рассмотрим следующую модель наблюдений:

$$y_i = \theta_0 + \theta_1 x_i + \theta_2 x_i^2 + \xi(x_i), \quad i = 1, 2, \dots, \quad (1)$$

где y_i – наблюдаемые значения, x_i – контролируемые переменные, $x_i \in [-1; 1]$; $\theta = (\theta_0, \theta_1, \theta_2)^T$ – подлежащий оцениванию вектор параметров модели (1); $\xi(x_i)$ – некоррелированные ошибки наблюдений с нулевым средним и дисперсией

$$D\{\xi(x)\} = d(x) = a_0 + a_1 x > 0, \quad x \in [-1; 1], \quad a_0 > 0, \quad |a_1| < a_0. \quad (2)$$

Для равнооточных наблюдений ($a_1=0$) модели (1), (2) непрерывный D -оптимальный план экспериментов имеет вид [1]:

$$\varepsilon^0 = \left\{ \begin{array}{ccc} -1; & 0; & 1 \\ \frac{1}{3}; & \frac{1}{3}; & \frac{1}{3} \end{array} \right\}, \quad (3)$$

где $-1, 0, 1$ – точки спектра плана, в которых нужно проводить наблюдения, а $\frac{1}{3}$ – мера, или вес, наблюдений в этих точках.

Однако для неравнооточных наблюдений ($a_1 \neq 0$) с линейным изменением дисперсии (2) для модели (1), (2) непрерывные D -оптимальные планы экспериментов до сих пор не были построены.

В данной статье приводится правило построения таких планов экспериментов и рассматривается робастность, устойчивость классических D -оптимальных планов экспериментов (3) с равнооточными наблюдениями относительно линейного изменения дисперсии наблюдений.

Проблема построения оптимальных планов экспериментов с неравнооточными наблюдениями исследовалась в статьях [2–4].

Справедлива следующая

Теорема 1. Для модели неравнооточных наблюдений (1), (2) непрерывный D -оптимальный план экспериментов имеет вид:

$$\varepsilon^0(k) = \left\{ \begin{array}{ccc} -1; & \frac{-2 + \sqrt{4 - 3k^2}}{3k}; & 1 \\ \frac{1}{3}; & \frac{1}{3}; & \frac{1}{3} \end{array} \right\}, \quad (4)$$

где $k = \frac{a_1}{a_0}$, $|k| < 1$.

Доказательство. Для равнооточных наблюдений ($k=0$) план экспериментов (4) обращается в классический план (3) для равнооточных наблюдений.

Поскольку функция отклика для модели наблюдений (1) является полиномом второй степени относительно контролируемой переменной x , то в

соответствии с теорией построения непрерывных D -оптимальных планов для полиномиальной одномерной регрессии [1] непрерывный D -оптимальный план для модели наблюдений (1) сосредоточен в трех точках, если дисперсия наблюдений является полиномом степени не выше четвертой. А так как дисперсия наблюдений (2) является полиномом первой степени относительно контролируемой переменной x , то для модели (1), (2) непрерывный D -оптимальный план экспериментов следует искать среди планов, сосредоточенных в трех различных точках. Далее, так как число точек спектра непрерывного D -оптимального плана для модели наблюдений (1), (2) совпадает с числом неизвестных параметров $\theta_0, \theta_1, \theta_2$ в (1), то вес наблюдений такого плана должен быть равным $\frac{1}{3}$. Итак, непрерывный D -оптимальный план экспериментов для модели наблюдений (1), (2) должен иметь вид:

$$\varepsilon = \begin{Bmatrix} x_1; & x_2; & x_3 \\ \frac{1}{3}; & \frac{1}{3}; & \frac{1}{3} \end{Bmatrix}, \quad (5)$$

где x_1, x_2, x_3 – точки спектра плана, выбираемые на отрезке $[-1; 1]$, $x_1 < x_2 < x_3$.

Следовательно, чтобы построить непрерывный D -оптимальный план для модели наблюдений (1), (2), необходимо указать точки спектра x_1, x_2, x_3 в плане (5), для которых определитель информационной матрицы плана (5) был бы максимальным.

Численные расчеты с использованием ЭВМ показали, что максимальное значение определителя информационной матрицы плана (5) достигается при $x_1 = -1, x_3 = 1$. Поэтому непрерывный D -оптимальный план экспериментов для модели наблюдений (1), (2) должен находиться среди планов

$$\varepsilon(x_2) = \begin{Bmatrix} -1; & x_2; & 1 \\ \frac{1}{3}; & \frac{1}{3}; & \frac{1}{3} \end{Bmatrix}. \quad (6)$$

Определитель информационной матрицы $M(\varepsilon(x_2))$ плана (6) для модели наблюдений (1), (2) равен

$$M(\varepsilon(x_2)) = f(x_2, k) = \frac{4(x_2^2 - 1)^2}{27(1 + kx_2)}. \quad (7)$$

Задача поиска непрерывного плана, «подозрительного» на D -оптимальность, свелась к нахождению точек максимума по $x_2, -1 < x_2 < 1$, функции (7). Для того чтобы найти максимум функции $f(x_2, k)$ по $x_2, -1 < x_2 < 1$, поступим стандартным образом: найдем точки, в которых производная этой функции обращается в ноль. В результате придем к уравнению $(x_2^2 - 1)(3k^2x_2^2 + 4x_2 + k) = 0$. Корни $x_2 = \pm 1$ этого уравнения являются граничными точками отрезка $[-1; 1]$, и в этих точках определитель (7) информационной матрицы плана (6) обращается в ноль, т. е. эти точки не могут доставлять максимальное значение функции $f(x_2, k)$, $x_2 \in (-1, 1)$. Две другие точки, в которых производная функции $f(x_2, k)$ может обращаться в ноль, находятся как корни квадратного уравнения $3k^2x_2^2 + 4x_2 + k = 0$:

$$x_{21}(k) = \frac{-2 - \sqrt{4 - 3k^2}}{3k}, \quad x_{22}(k) = \frac{-2 + \sqrt{4 - 3k^2}}{3k}.$$

Очевидно, что для $-1 < k < 1$ корень $x_{21}(k)$ по модулю больше 1, а $x_{22}(k) \in \left(-\frac{2}{3}, \frac{2}{3}\right)$ и $f(x_{22}(k), k) > 0$. Итак, значение $x_2 = x_{22}(k)$ доставляет максимум функции $f(x_2, k)$, $x_2 \in (-1, 1)$, и при таком значении x_2 план экспериментов (6) обращается в план (4).

Чтобы завершить доказательство теоремы, осталось показать, что план экспериментов (4), построенный таким полуинтуитивным способом, действительно является D -оптимальным, т. е. выполняется критерий D -оптимальности Кифера – Вольфовитца [1]: дисперсия поверхности отклика $d(x, \varepsilon^0(k)) < 3$ для $x \in [-1, 1]$, $k \in (-1, 1)$, и в точках спектра плана $\varepsilon^0(k)$ это неравенство обращается в равенство. Здесь $d(x, \varepsilon^0(k))$ – дисперсия поверхности отклика:

$$d(x, \varepsilon^0(k)) = \frac{1}{d(x)} (1, x, x^2) M^{-1}(\varepsilon^0(k)) (1, x, x^2)^T.$$

Функция $d(x, \varepsilon^0(k))$ как функция аргумента x является отношением полинома четвертой степени к полиному первой степени. Коэффициенты полинома четвертой степени зависят от k и имеют сложные, громоздкие выражения, что не позволяет строго аналитически проверить выполнимость критерия D -оптимальности Кифера – Вольфовитца. Поэтому вместо проверки данного критерия оптимальности для всех $x \in [-1, 1]$, $k \in (-1, 1)$ проверялось его выполнение в узлах достаточно мелкой сетки с шагом $h=0,01$ по оси изменения x и по оси изменения k . При этом для всех значений k узел $(s^0(k), k)$ включался в сетку. Численные расчеты на ЭВМ подтвердили выполнение критерия D -оптимальности непрерывного плана (4), что и завершает доказательство теоремы.

Для практических расчетов D -оптимальный план (4) применим только тогда, когда коэффициенты a_0, a_1 дисперсии наблюдений (2) заданы. Однако обычная ситуация, с которой сталкивается исследователь, состоит в том, что он знает только лишь качественный характер изменения дисперсии (линейная зависимость), но не знает точных количественных характеристик этой зависимости, поскольку дисперсия выражается через моменты второго порядка наблюдаемых значений. Но даже моменты первого порядка (функция отклика) ему точно неизвестны. Какой же выход из подобной ситуации?

В этом случае можно использовать непрерывный D -оптимальный план (3) для равноточных наблюдений, априори зная, что дисперсия наблюдений изменяется линейно, при этом нужно оценить относительную величину ошибки γ определителя информационной матрицы плана (3):

$$\frac{|M(\varepsilon^0(k))| - |M(\varepsilon^0)|}{|M(\varepsilon^0)|} < \gamma, \quad \gamma > 0. \quad (8)$$

Исходя из неравенства (8), можно решать и обратную задачу. При заданном уровне искажения γ величины определителя информационной матрицы плана экспериментов – определить, в каких пределах должен изменяться уровень искажения k равноточных наблюдений, чтобы вместо плана $\varepsilon^0(k)$ можно было бы использовать план экспериментов ε^0 .

Например, при уровне искажения величины определителя информационной матрицы в 10 % ($\gamma=0,1$) вместо плана $\epsilon^0(k)$ можно использовать план ϵ^0 , если уровень искажения k постоянного значения дисперсии будет принадлежать интервалу $[-0,8023; 0,8023]$. Данные расчеты были проведены с использованием ЭВМ и неравенства (8).

1. Федоров В. В. Теория оптимального эксперимента. М., 1968.
2. Кирлица В. П. // Вестн. Белорус. ун-та. Сер. 1. 1990. № 2. С. 36.
3. Кирлица В. П. // Там же. 1995. № 1. С. 42.
4. Kirlictsa V. P. // Computer data analysis and modeling. Minsk, 1995. Vol. 1. P. 64.

Поступила в редакцию 01.02.2001.

Андрей Владимирович Акинфин – научный сотрудник НИЛ СТАМ.

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

Евгения Анатольевна Черкас – студентка 5-го курса факультета прикладной математики и информатики.

УДК 519.1

О.И. МЕЛЬНИКОВ, В.Н. КОСОЛАПОВ

РАСПОЗНАВАНИЕ P_4 -СТРУКТУРЫ ГУСЕНИЦЫ

An $O(n)$ time-bounded algorithm to decide whether a given 4-uniform hypergraph represents the P_4 -structure of a caterpillar is proposed.

Реализацией гиперграфа $H(V, E)$ [1] называется такой граф $G(V, E)$, каждое ребро которого принадлежит некоторому ребру гиперграфа, и каждый подграф, порожденный вершинами любого гиперребра e гиперграфа H , связан. Этот подграф называется реализацией гиперребра e .

Пусть задано некоторое множество графов $\{S\}$. *S -реализацией гиперграфа H называется такой граф $G(V, E)$, каждый подграф которого, порожденный вершинами любого гиперребра, принадлежит S , и любой порожденный подграф графа G , принадлежащий S , определяет некоторое гиперребро гиперграфа H .*

Для заданного графа G рассмотрим гиперграф H , гиперребра которого образуются всеми порожденными P_4 -цепями графа. С одной стороны, гиперграф H называется P_4 -структурой графа G . С другой – граф G является P_4 -реализацией гиперграфа H .

Изучение P_4 -структур графов представляет интерес, поскольку доказано [2], что два графа, имеющие одинаковую P_4 -структуру, оба или совершенны, или несовершенны.

В [3] построен линейный алгоритм распознавания P_4 -структуры дерева.

В настоящей работе приведен алгоритм распознавания, является ли 4-униформный гиперграф P_4 -структурой некоторой гусеницы. Трудоемкость алгоритма – $O(|E|)$.

Гусеницей называется дерево, которое после удаления всех висячих вершин превращается в цепь.

Теорема 1. Пусть T – гусеница, диаметр которой не меньше 6, H – ее P_4 -структура. Тогда T – единственная P_4 -реализация гиперграфа H гусеницей.

Доказательство. Пусть $L(v_1, v_2, \dots, v_l)$ – диаметральная цепь гусеницы T ($l \geq 7$); (v_1, v_2, v_3, v_4) , (v_2, v_3, v_4, v_5) , ..., $(v_{l-3}, v_{l-2}, v_{l-1}, v_l)$ – цепи P_4 , опре-