Математические методы исследования
Mathematical methods of investigation

DOI: https://doi.org/10.26896/1028-6861-2020-86-5-65-72

Q-ОПТИМАЛЬНЫЕ И БЛИЗКИЕ К НИМ ПЛАНЫ ЭКСПЕРИМЕНТА ДЛЯ ПОЛИНОМИАЛЬНОЙ РЕГРЕССИИ НА ОТРЕЗКЕ

© **Юрий Дмитриевич Григорьев**

Санкт-Петербургский государственный электротехнический университет («ЛЭТИ»), Россия, 197376, Санкт-Петербург, ул. Профессора А. Попова, 5; e-mail: yuri_grigoriev@mail.ru

Статья поступила 14 мая 2019 г. Поступила после доработки 24 июня 2019 г. Принята к публикации 28 августа 2019 г.

Рассмотрена задача построения Q-оптимальных планов эксперимента для полиномиальной регрессии на отрезке [–1, 1]. Показано, что известные планы Малютова – Федорова, использующие спектр D-оптимальных планов (спектр Лежандра), Q-оптимальными не являются. Этот вывод является непосредственным следствием замечания Шабадоса к гипотезе Эрдеша, опровергающим ее. Сама гипотеза Эрдеша заключалась в том, что спектр насыщенных D-оптимальных планов для полиномиальной регрессии на отрезке одновременно является и спектром насыщенных Q-оптимальных планов. Приведен насыщенный точный Q-оптимальный план для полиномиальной регрессии степени $s = 3$, подтверждающий замечание Шабадоса. Далее это утверждение переносится на непрерывные планы. Для случаев $s = 3, 4$ показано, что известная теорема Малютова – Федорова о непрерывных Q-оптимальных планах также неверна, хотя и остается справедливой для степеней $s = 1, 2$. Исследованы планы Малютова – Федорова со спектром Лежандра с точки зрения их близости к Q-оптимальным. На примерах показано, что они достаточно близки для малых степеней s полиномиальной регрессии. Найдено универсальное выражение для Q-оптимального распределения весов p_i опорных точек x_i в случае произвольного спектра. В качестве примера с помощью полученного выражения проведено табулирование распределения весов для планов Малютова – Федорова для $s = 3, ..., 6$. Отмечена общность полученного выражения для Q-оптимальных весов с A-оптимальным распределением весов (распределение Пукельсхайма) при той же постановке задачи. В заключение дана краткая рекомендация о численном построении Q-оптимальных планов. Отмечено, что помимо традиционных численных методов в данном случае могут быть использованы программные системы символьных вычислений, использующие методы результантов и исключения. Приводимые в статье примеры Q-оптимальных планов построены с использованием именно этих методов.

Ключевые слова: полиномиальная регрессия; критерий Q-оптимальности; план эксперимента; планы эксперимента Малютова – Федорова; интерполяционные полиномы Лагранжа; спектр плана; метод результантов.

Q-OPTIMAL EXPERIMENTAL DESIGNS AND CLOSE TO THEM EXPERIMENTAL DESIGNS FOR POLYNOMIAL REGRESSION ON THE INTERVAL

© **Yury D. Grigoriev**

St. Petersburg State Electrotechnical University (“LETI”), 5 Professor A. Popov str., St. Petersburg, 197376, Russia; e-mail: yuri_grigoriev@mail.ru

Received May 14, 2019. Revised June 24, 2019. Accepted August 28, 2019.

The problem of constructing Q-optimal experimental designs for polynomial regression on the interval [–1, 1] is considered. It is shown that well-known Malyutov – Fedorov designs using D-optimal designs (so-called Legendre spectrum) are other than Q-optimal designs. This statement is a direct consequence of Shabados remark which disproved the Erdős hypothesis that the spectrum (support points) of saturated D-optimal designs for polynomial regression on a segment appeared to be support points of saturated

Q-optimal designs. We present a saturated exact Q-optimal design for polynomial regression with $s = 3$ which proves the Shabados notion and then extend this statement to approximate designs. It is shown that when $s = 3, 4$ the Malyutov – Fedorov theorem on approximate Q-optimal design is also incorrect, though it still stands for $s = 1, 2$. The Malyutov – Fedorov designs with Legendre spectrum are considered from the standpoint of their proximity to Q-optimal designs. Case studies revealed that they are close enough for small degrees s of polynomial regression. A universal expression for Q-optimal distribution of the weights p_i for support points x_i for an arbitrary spectrum is derived. The expression is used to tabulate the distribution of weights for Malyutov – Fedorov designs at $s = 3, ..., 6$. The general character of the obtained expression is noted for Q-optimal weights with A-optimal weight distribution (Pukelsheim distribution) for the same problem statement. In conclusion a brief recommendation on the numerical construction of Q-optimal designs is given. It is noted that in this case in addition to conventional numerical methods some software systems of symbolic computations using methods of resultants and elimination theory can be successfully applied. The examples of Q-optimal designs considered in the paper are constructed using precisely these methods.

Keywords: polynomial regression; Q-optimality criterion; experimental design; Malyutov – Fedorov experimental designs; Lagrangian interpolation polynomials; design spectrum; method of the resultants.

Введение

При построении оптимальных планов эксперимента используют как аналитические [1 – 3], так и численные методы [4], при этом приоритет, если это возможно, отдают аналитическим методам. Такое часто имеет место в задачах нелинейного планирования (см., например, [5, 6]), а также для определенных классов линейно параметризованных моделей. К числу данных моделей относится, в частности, класс полиномиальных моделей.

Задача построения оптимальных планов для полиномиальной регрессии на отрезке [–1, 1] в настоящее время хорошо изучена. В этом направлении получено много аналитических и численных результатов. Тем не менее на некоторые вопросы ответы до сих пор не получены.

Работа посвящена уточнению некоторых теоретических результатов, касающихся Q-оптимальных планов, из которых, в частности, следует возможность рассмотрения планов типа Малютова – Федорова применительно к A-критерию.

Суть рассматриваемой проблемы состоит в следующем. В серии работ [7 – 9] Малютов и Федоров сформулировали результат, согласно которому планы со спектром D-оптимальных планов (спектр Лежандра) и соответствующим ему распределением весов являются Q-оптимальными. Однако оказалось, что для степеней полинома $s \geq 3$ данный результат неверен. Это является следствием сделанного в 1966 г. замечания Шабадоса [10] к гипотезе Эрдеша [11], опровергающим ее.

В связи с этим обстоятельством в работе построены точные и непрерывные Q-оптимальные планы для $s = 3, 4$ и проведено их сравнение с точными планами, соответствующими гипотезе Эрдеша, и непрерывным планам Малютова – Федорова. Поскольку для данных значений s различие между Q-оптимальными и приближенными ξ_{MF}-планами со спектром Лежандра оказалось крайне незначительным, то в работе построены планы ξ_{MF} для последующих степеней $s = 5, 6$ в предположении, что и для этих степеней различие между ними и Q-оптимальными планами хотя и будет возрастать, но останется в разумных пределах.

Планы ξ_{MF} интересны в том отношении, что согласно найденному в работе аналитическому соотношению для весов, имеющему более общий характер, чем результат Малютова – Федорова, выражения для весов точек планов ξ_{MF} могут быть протабулированы для любого спектра. Соответствующее аналитическое выражение аналогично результату Малютова – Федорова и связано с фундаментальными интерполяционными полиномами Лагранжа.

Универсальное выражение для весов ξ_{MF}-планов с произвольным спектром оказалось совершенно аналогичным распределению весов Пукельсхайма для A-оптимальных планов с произвольным спектром. Распределение Пукельсхайма определяется главной диагональю матрицы $(FF^{\mathrm{T}})^{-1}$, где F — матрица плана, являющаяся в нашем случае матрицей Вандермонда.

Статья включает постановку задачи, необходимые обозначения. В ней приведены результаты из теории интерполяции, сформулирована гипотеза Эрдеша и представлен опровергающий ее пример Шабадоса. Рассмотрены непрерывные Q-оптимальные и связанные с ними планы Малютова – Федорова ξ_{MF}. Приведены табулированные значения весов планов ξ_{MF} и оптимальный план для квадратичной регрессии на отрезке с областью усреднения Z, отличной от области планирования $X = [–1, 1]$.

Отметим, что критерий Q-оптимальности в западной литературе называется I-критерием [12, 13], но мы сохраняем за ним название, более привычное отечественному читателю [9, с. 153].

Постановка задачи

Как критерий выбора интерполяционного базиса в задачах численного анализа критерий Q-оптимальности появляется в работе [11]. С развитием теории оптимального планирования эксперимента он появляется в ней наряду с критерием A-оптимальности как один из многих других линейных критериев планирования. По его поводу В. В. Федоров пишет [9, с. 153]: «Если экспериментатора интересует общая закономерность зависимости изучаемой величины от контролируемых переменных, т.е. он может пожертвовать точностью описания в малых областях ради хорошего описания во всей области, то разумно потребовать, чтобы минимизировалась величина

$$\left(\int_Z \mathrm{d}z\right)^{-1} E\left(\int_Z [\eta(z,\hat{\theta})-\eta(z,\theta)]^2\,\mathrm{d}z\right)=\left(\int_Z dz\right)^{-1}\int_Z d(z,\xi)\mathrm{d}z,$$

где область Z необязательно совпадает с областью планирования X. Планирование эксперимента заключается в отыскании плана ξ^*, минимизирующего величину

$$Q[D(\xi)] = \int_Z d(z,\xi)\mathrm{d}z, \tag{1}$$

где $\mathrm{d}(x, \xi)$ — дисперсия оценки функции отклика $\eta(x, \theta)$». Далее будем рассматривать только случай $Z = X = [-1, 1]$. Введем необходимые определения.

Пусть задана модель наблюдений

$$y = \eta(x, \theta) + e = f(x)^{\mathrm{T}}\theta + e, \tag{2}$$

где $y \in \mathrm{R}^n$ — вектор наблюдений; $\eta(x, \theta) = f(x)^{\mathrm{T}}\theta$ — линейная по параметрам функция отклика; $\theta \in \mathrm{R}^m$ — вектор оцениваемых параметров; $n \geq m$; $e = N(0, \sigma^2 I_n)$ — вектор ошибок, имеющих нормальное распределение с нулевым средним $E[e] = 0$ и дисперсией $D[e] = \sigma^2 I_n$. Далее полагаем $\sigma^2 = 1$, а в качестве функции отклика рассмотрим полиномиальную модель регрессии степени s:

$$\eta(x,\theta) = \sum_{i=0}^{s}\theta_i x^{s-i},\ x \in X = [-1, 1]. \tag{3}$$

Из (3) следует, что $m = s + 1$ — количество оцениваемых параметров. Положим $n = m = s + 1$ и пусть

$$\xi = (supp\,\xi, p) = \{(x_i, p_i)_{i=1}^n\},\ x_i \in supp\,\xi,$$

$$p_i \geq 0,\ \sum_{i=1}^{n} p_i = 1 —$$

насыщенный план эксперимента, где $supp\,\xi = (x_1, ..., x_n)$ — спектр плана, удовлетворяющий условию

$$-1 \leq x_1 < x_2 < ... < x_n \leq 1, \tag{4}$$

а $p = (p_1, ..., p_n)$ — вероятностная мера, заданная на $supp\,\xi$.

Если $p_1 = ... = p_n = n^{-1}$, то соответствующий план называется точным, в противном случае — непрерывным. Требуется найти план ξ^*, минимизирующий функционал (1), который определяется следующим образом.

Обозначим $f(x)^{\mathrm{T}} = (1, x, ..., x^s)$, P — диагональную матрицу с элементами $P_{ii} = p_i$ на главной диагонали. Пусть

$$F = [f(x_1)^{\mathrm{T}}, ..., f(x_n)^{\mathrm{T}}] = x_i^{j-1},\ i, j = 1, ..., n,$$

$$M(\xi) = F^{\mathrm{T}}PF — \tag{5}$$

матрица планирования, составленная из строк $f(x_i)^{\mathrm{T}}$, и информационная матрица плана ξ соответственно. В случае (5) определитель $W = |F|$ сводится к известному определителю Вандермонда [14, с. 13], который не равен нулю в силу (4).

Поскольку в данном случае оценка наименьших квадратов

$$\hat{\theta} = M(\xi)^{-1}F^{\mathrm{T}}Py = F^{-1}y,$$

то

$$D[\eta(x,\hat{\theta})] = f(x)^{\mathrm{T}} D[\hat{\theta}] f(x) = f(x)^{\mathrm{T}} M(\xi)^{-1} f(x) — \tag{6}$$

дисперсия оценки $\eta(x, \hat{\theta})$. Полагая $D(\xi) := D[\hat{\theta}] = M(\xi)^{-1}$, определяем функцию дисперсии

$$d(x, \xi) = f(x)^{\mathrm{T}}D(\xi)f(x) = f(x)^{\mathrm{T}}M(\xi)^{-1}f(x),$$

$$x \in X = [-1, 1], \tag{7}$$

играющую основную роль в последующем изложении. План ξ^*, минимизирующий функционал

$$Q[D(\xi)] = \int_{-1}^{1} d(x,\xi)\mathrm{d}x, \tag{8}$$

называется Q-оптимальным. Наша задача — исследовать структуру Q-оптимальных планов, остановившись, в частности, на свойствах планов Малютова – Федорова [8], которые, как оказалось, не являются Q-оптимальными в классе непрерывных планов Ξ_n для $s \geq 3$.

Связь с теорией интерполяции

Пусть на отрезке [–1, 1] заданы $n = s + 1$ произвольных точек (4). Положим $\omega(x) = c(x - x_1)(x - x_2)\ldots(x - x_n)$ $(c \neq 0)$ и пусть

$$l_i = \frac{\omega(x)}{\omega'(x)(x - x_i)},\ i = 1, \ldots, n, \text{ —}$$

фундаментальные интерполяционные многочлены Лагранжа, удовлетворяющие условиям $l_i(x_j) = \delta_{ij}$. Множество последовательностей (4), как отмечено выше, обозначаем Ξ_n.

Существует несколько задач наилучшей интерполяции, связанных с последовательностями (4) и порождаемыми ими многочленами $l = (l_1, \ldots, l_n)^{\text{т}}$. Нас будут интересовать две из них, первая из которых имеет отношение к D-, а вторая — к Q-оптимальному планированию эксперимента.

Задача Фейера. Задача состоит в отыскании такой последовательности $(x_1, \ldots, x_n) \in \Xi_n$, для которой

$$\max_{x \in [-1,1]} \sum_{i=1}^{n} l_i^2(x)$$

достигает минимума. В 1932 году Фейер доказал [15], что этот минимум достигается на последовательности $(x_1, \ldots, x_n) \in \Xi_n$, элементы которой являются корнями многочлена $(1 - x^2)P'_{n-1}(x)$, где $P_n(x)$ — n-й многочлен Лежандра, при этом

$$\inf_{\xi \in \Xi_n} \max_{x \in [-1,1]} \sum_{i=1}^{n} l_i(x)^2 = 1.$$

Это следует из имеющего место с точностью до константы $c > 0$ представления

$$\omega(x) = \int_{-1}^{x} P_{n-1}(t)\mathrm{d}t = (x^2 - 1)P'_n(x),$$

$$P_n(x) = \frac{1}{n!\,2^n} \frac{\mathrm{d}^n}{\mathrm{d}x^n}(x^2 - 1)^n,\ n \geq 1. \qquad (9)$$

Выражения для $P_n(x)$ и $P'_n(x)$ приводятся в разных источниках, в частности [16, с. 271].

В современной постановке решение задачи Фейера эквивалентно задаче построения насыщенных D-оптимальных планов для полиномиальной регрессии степени s. В связи с этим (для краткости) спектр D-оптимальных планов назовем спектром Лежандра. Для $s = 2, \ldots, 6$ эти спектры представлены в табл. 1.

Гипотеза Эрдеша. В работе [11] при тех же условиях, что и выше, была сформулирована задача определения последовательности (4), для которой величина

$$Q(x_1, \ldots, x_n) = \int_{-1}^{1} \sum_{i=1}^{n} l_i^2(x)\mathrm{d}x,\ n = 2, 3, \ldots, \qquad (10)$$

минимальна.

Очевидно, что в терминах планирования эксперимента мы имеем задачу Q-оптимального планирования [9, с. 153]. Эрдеш предположил, что ее решение достигается на спектре Лежандра. Однако в [10] эта гипотеза была опровергнута для значений $s \geq 3$.

Пример 1. Пусть $s = 3$. Для точного плана $\xi = \{x_{1,4} = \pm 1, x_{2,3} = \pm u\}$ с $p_i = \xi(x_i) = 1/4$ согласно (8) запишем

$$Q(\xi) = 4\int_{-1}^{1} \sum_{i=1}^{4} l_i^2(x)\mathrm{d}x\, \frac{16}{105} \frac{35u^6 - 28u^4 + 23u^2 + 2}{(u^2 - 1)^2 u^2}.$$

Обозначим $a = \sqrt[3]{36 + 21\sqrt{3}}$. Стандартной оптимизацией по $v = u^2$ получаем, что оптимальные значения u^* и $Q(\xi^*)$ имеют вид:

$$v^* = \frac{2a^2 - 3a - 6}{21a},\ u^* = 0{,}4307,\ Q(\xi^*) = 6{,}8430.$$

Таблица 1. Спектры Лежандра, или нули полиномов $(x^2 - 1)P'_n(x)$, s — степень полинома

Table 1. Legendre spectra (support points) or the roots of polynomials $(x^2 - 1)P'_n(x)$, s – polynomial degree

s	Узлы D-оптимальных планов			
	$x_{1,s+1}$	$x_{2,s}$	$x_{3,s-1}$	$x_{4,s-2}$
2	±1	0		
3	±1	$\pm\frac{1}{\sqrt{5}} = \pm 0{,}4472$		
4	±1	$\pm\frac{\sqrt{21}}{7} = \pm 0{,}6546$	0	
5	±1	$\pm\frac{\sqrt{147 + 42\sqrt{7}}}{21} = 0{,}7650$	$\pm\frac{\sqrt{147 - 42\sqrt{7}}}{21} = \pm 0{,}2852$	
6	±1	$\pm\frac{\sqrt{495 + 66\sqrt{15}}}{33} = \pm 0{,}8302$	$\pm\frac{\sqrt{495 - 66\sqrt{15}}}{33} = \pm 0{,}4688$	0

В то же время для D-оптимального спектра $\xi = \{x_{1,4} = \pm 1, x_{2,3} = \pm 1/\sqrt{5}\}$, тогда

$$u = 0{,}4472,\ Q(\xi) = 48/7 = 6{,}8571.$$

Поскольку $Q(\xi) > Q(\xi^*)$, то вывод Шабадоса подтверждается [10]. Отметим, что различие между D- и Q-оптимальными планами — ξ и ξ^* — соответственно крайне незначительно.

Непрерывные Q-оптимальные планы

Гипотезу Эрдеша можно распространить и на непрерывные планы. Возможно, руководствуясь ею, авторы работ [7 – 9, с. 155] сформулировали теорему о непрерывных Q-оптимальных планах для полиномиальной регрессии.

Теорема 1 [8]. Пусть выполнены следующие условия:

1) $\eta(x, \theta) = \sum_{i=0}^{s} \theta_i x^{s-i}$ — полиномиальная регрессия степени $s \geq 1$;

2) $X = [-1, 1]$ — область планирования и усреднения функции дисперсии $d(x, \xi)$.

Тогда имеют место следующие утверждения:

1) спектр $supp\,\xi = (x_1, ..., x_n)$ непрерывных Q-оптимальных планов совпадает с множеством корней полинома $(1 - x^2)P_s'(x)$, где $P_s(x)$ — s-й многочлен Лежандра;

2) веса точек $x_i \in supp\,\xi$ Q-оптимальных планов ξ определяются выражением

$$p_i = \frac{|P_s(x_i)|^{-1}}{\sum_{j=1}^{s+1} |P_s(x_j)|^{-1}},\ i = 1, ..., s + 1. \qquad (11)$$

Назовем планы ξ со спектром Лежандра и распределением весов (11) планами Малютова – Федорова и обозначим их ξ_{MF}. Оказывается, что в общем случае (для $s \geq 3$) теорема 1 неверна. Контрпример для $s = 3$ приведен в [16, с. 160]. В качестве примера в табл. 2 представлены Q-оптимальные планы ξ^* и планы Малютова – Федорова ξ_{MF} для $s = 3, 4$, что дает возможность сравнить их между собой.

Сравнивая планы ξ^* и ξ_{MF} при $s = 3, 4$, заключаем, что различие между ними крайне незначительно, хотя с ростом s разность $Q(\xi_{MF}) - Q(\xi^*)$ начинает возрастать. Возникает вопрос, каким на самом деле является спектр Q-оптимальных планов? Существует ли для него аналитическое выражение? Ответ на этот вопрос до сих пор не получен.

Тем не менее определенное продвижение в вопросе построения Q-оптимальных планов дает следующая теорема.

Теорема 2. Пусть выполнены следующие условия:

1) $\eta(x, \theta) = \sum_{i=0}^{s} \theta_i x^{s-i}$ — полиномиальная регрессия степени $s \geq 1$;

2) $X = [-1, 1]$ — область планирования и усреднения функции дисперсии $d(x, \xi)$;

3) $l = (l_1, ..., l_{s+1})^{\text{т}}$ — фундаментальные интерполяционные многочлены Лагранжа.

Тогда распределение весов $p = (p_1, ..., p_{s+1})^{\text{т}}$ узлов плана ξ, Q-оптимальное относительно спектра $supp\,\xi = (x_1, ..., x_{s+1})$, имеет вид

$$p_i = \frac{\sqrt{L_i}}{\sum_{j=1}^{s+1}\sqrt{L_j}},\ L_i = \int_{-1}^{1} l_i^2(x)\mathrm{d}x,\ i = 1, ..., s + 1. \qquad (12)$$

Доказательство. Согласно условиям 1) – 3) имеет место равенство [14, с. 334]

$$l(x) = (F^{\text{т}})^{-1} f(x),$$

где $f(x) = (1, x, ..., x_s)^{\text{т}}$ — вектор, определяемый согласно (2), а само равенство следует из того, что слева и справа в нем стоят многочлены степени s, совпадающие в $s + 1$ точках, т.е. они тожде-

Таблица 2. Q-оптимальные планы ξ^* (первые строки) и планы Малютова – Федорова ξ_{MF} (вторые строки), s — степень полинома

Table 2. Q-optimal designs ξ^* (the first row) and Malyutov – Fedorov designs ξ_{MF} (the second row), s – polynomial degree

s	Узлы и веса планов ξ^* и ξ_{MF}			$Q(\xi^*)$, $Q(\xi_{MF})$
	$(x_{1,s+1}, p_{1,s+1})$	$(x_{2,s}, p_{2,s})$	$(x_{3,s-1}, p_{3,s-1})$	
3	$(\pm 1,\ 0{,}1549)$	$(\pm 0{,}4366,\ 0{,}3451)$		5,9796
	$\left(\pm 1, \frac{\sqrt{5}-1}{8} = 0{,}1545\right)$	$\left(\frac{\pm 1}{\sqrt{5}}, \frac{5-\sqrt{5}}{8}\right)$		$\frac{24+8\sqrt{5}}{8} = 5{,}9841$
4	$(\pm 1,\ 0{,}1076)$	$(\pm 0{,}6436,\ 0{,}2501)$	$(0,\ 0{,}2847)$	7,7351
	$\left(\pm 1, \frac{3}{28} = 0{,}1071\right)$	$\left(\pm\sqrt{\frac{3}{7}}, \frac{1}{4}\right)$	$\left(\frac{0{,}2}{7} = 0{,}2857\right)$	$\frac{3136}{405} = 7{,}7433$

ственно равны. Поскольку $M(\xi) = F^{\text{т}}PF$, то отсюда и согласно (7) получим

$$d(x, \xi) = f(x)^{\text{т}}M(\xi)^{-1}f(x) = \sum_{i=1}^{s+1} p_i^{-1} l_i^2(x). \qquad (13)$$

Функционал (8) является выпуклым на множестве информационных матриц. Это означает, что при фиксированном спектре он имеет единственный минимум по p. Стандартной оптимизацией функционала (8) по p с использованием (13) приходим к результату (12), что и требовалось доказать.

Выражения для L_i могут быть протабулированы для различных спектров. В частности, это дает возможность продолжить табл. 2 применительно к планам ξ_{MF} (для спектров Лежандра) для значений $s \geq 5$. В табл. 3 они представлены для $s = 2, ..., 6$. Отметим, что в силу симметрии весов $p = (p_1, ..., p_{s+1})$ выражения для L_i также симметричны и связаны соотношениями:

$$L_i = L_{n+1-i},\ i = 1, ..., n - 1 - [n/2],$$

$$L_{n-[n/2]} = L_{[n/2]+1},\ n = s + 1,$$

где $[a]$ — целая часть числа a.

Пример 2. Пусть $s = 2$, $X = Z = [-1, 1]$. Спектр Лежандра в этом случае имеет вид $supp\,\xi = (-1, 0, 1)$. Согласно (12) и табл. 3 для $n = 3$ получаем $p = (1/4, 1/2, 1/4)$. С использованием теоремы эквивалентности Кифера – Вольфовица легко проверяется, что план Малютова – Федорова $\xi_{MF} = \{(\pm 1, 1/4), (0, 1/2)\}$ в данном случае является Q-оптимальным, что согласуется с полученными выше выводами.

С помощью метода результантов в [16, с. 159] построен Q-оптимальный план ξ^* для области усреднения $Z = [0, 2]$. Спектр для него оказался совпадающим со спектром Лежандра, а вот распределение весов p оказалось иным:

$$p_1 = \frac{10 - \sqrt{46}}{27} = 0{,}1192,\ p_2 = \frac{5\sqrt{46} - 23}{27} = 0{,}4041,$$

$$p_3 = \frac{40 - 4\sqrt{46}}{27} = 0{,}4767.$$

Таким образом, замечание Шабадоса можно распространить и на случай, когда области планирования X и усреднения Z не совпадают.

В заключение отметим, что теорема 2 вполне аналогична ситуации с A-оптимальными планами ξ, минимизирующими след дисперсионной матрицы $trD(\xi)$. Для данных планов также неизвестен A-оптимальный спектр $supp\,\xi$, а при произвольном спектре A-оптимальное распределение весов p имеет вид [17]

$$p_i = \frac{\sqrt{B_i}}{\sum_{j=1}^{s+1}\sqrt{B_j}},\ B_i = (FF^{\text{т}})_{ii}^{-1},\ i = 1, ..., s + 1.$$

Величины B_i также могут быть протабулированы для конкретных спектров, включая спектр Лежандра.

Вычислительный аспект

Численное построение планов эксперимента для полиномиальной регрессии на отрезке для различных критериев оптимальности, включая Q-оптимальность, возможно с помощью программных систем символьных вычислений типа MAPLE. При этом используются различные варианты метода результантов и метода исключений.

Обзор этих методов, включая вычисление базиса Гребнера [18] и метод Безу, приведен в [19, 20]. Краткое описание метода результантов и

Таблица 3. Веса $p = (p_1, ..., p_n)^{\text{т}}$ точек планов ξ_{MF}, выраженные в координатах $L = (L_1, ..., L_n)$, $n = s + 1$

Table 3. The weights $p = (p_1, ..., p_n)^{\text{T}}$ of design support points ξ_{MF} in the coordinates $L = (L_1, ..., L_n)$, $n = s + 1$

n	$L = (L_1, ..., L_n)$						
	L_1	L_2	L_3	L_4	L_5	L_6	L_7
3	$\frac{4}{15}$	$\frac{16}{15}$	$\frac{4}{15}$				
4	$\frac{1}{7}$	$\frac{5}{7}$	$\frac{5}{7}$	$\frac{1}{7}$			
5	$\frac{4}{45}$	$\frac{196}{405}$	$\frac{256}{405}$	$\frac{196}{405}$	$\frac{4}{45}$		
6	$\frac{2}{33}$	$\frac{14-\sqrt{7}}{33}$	$\frac{14+\sqrt{7}}{33}$	$\frac{14+\sqrt{7}}{33}$	$\frac{14-\sqrt{7}}{33}$	$\frac{2}{33}$	
7	$\frac{4}{91}$	$\frac{744-42\sqrt{15}}{2245}$	$\frac{744+42\sqrt{15}}{2245}$	$\frac{1024}{2245}$	$\frac{744+42\sqrt{15}}{2245}$	$\frac{744-42\sqrt{15}}{2245}$	$\frac{4}{91}$

примеры его использования для построения оптимальных планов даны в [16]. В частности, Q-оптимальные планы ξ^* для $s = 3, 4$ в табл. 2 построены с помощью метода результантов.

Основным недостатком методов символьных вычислений является то, что их сложность и требуемый объем вычислений резко возрастают с ростом размерности задачи. Поэтому обычно при их использовании приходится ограничиваться тремя-четырьмя переменными, которыми являются точки спектра плана и их веса.

В этом смысле построение планов типа Малютова – Федорова не представляет затруднений, так как спектр планов задан, а распределение весов легко вычисляется. Это делает планы ξ_{MF} удобными в применении без большой потери эффективности планирования.

Заключение

В работе отмечена неточность в формулировке теоремы Малютова – Федорова о непрерывных Q-оптимальных планах, заключающаяся в том, что описываемые ею планы построены на базе спектра Лежандра, который на самом деле спектром Q-оптимальных планов в общем случае не является. Утверждение теоремы перестает быть справедливым для $s \geq 3$, где s — степень полиномиальной регрессии.

Показано, что планы Малютова – Федорова ξ_{MF} являются достаточно эффективными для малых значений s, что делает их удобными при использовании на практике, так как их вычисление не представляет затруднений.

В работе найдено универсальное выражение для Q-оптимального распределения затрат, не связанное с видом спектра. Частные случаи этого выражения протабулированы для значений $s = 2, ..., 6$. Отмечена общность этого выражения с соответствующим результатом Пукельсхайма для A-оптимальных планов.

Методом результантов построено несколько непрерывных Q-оптимальных планов и отмечена трудность их построения методами символьных вычислений для больших значений s.

ЛИТЕРАТУРА

1. **Guchenko R., Melas V. B.** Efficient computation of Bayesian optimal discriminating designs / J. Comput. and Graphic. Statist. 2017. Vol. 24. N 2. P. 424 – 433.
2. **Melas V. B., Shpilev P.** T-optimal discriminating designs for Fourier regression models / Comput. Statis. and Data Analysis. 2017. Vol. 113. P. 196 – 206.
3. **Guchenko R., Melas V. B., Wong W. K.** Optimal discrimination designs for semi-parametric models / Biometrika. 2018. Vol. 105. N 1. P. 185 – 197.
4. **Ермаков С. М., Семенчиков Д. Н.** О методах оптимизации в задачах планирования эксперимента / Заводская лаборатория. Диагностика материалов. 2019. Т. 85. № 1. Ч. 1. С. 72 – 77.
5. **Григорьев Ю. Д., Мелас В. Б., Шпилев П. В.** Избыточность локально D-оптимальных планов и гомотетии / Вестник СПбГУ. Сер. 1. Математика. Механика. Астрономия. 2017. Т. 4(62). № 4. С. 552 – 562.
6. **Grigoriev Yu. D., Melas V. B., Shpilev P. V.** Excess of locally D-optimal designs for Cobb-Douglas model / Statistical Papers. 2018. Vol. 59. N 4. P. 1425 – 1439.
7. **Fedorov V. V., Malytov M. B.** On the designs for certains weighted polynomial regression minimizing the average variance. Preprint. N 8, LSR. — Moscow: State University Press, 1969. — 12 p.
8. **Малютов М. Б., Федоров В. В.** О планах взвешенной полиномиальной средней регрессии / Теория вероят. и примен. 1971. Т. 16. № 4. С. 734 – 738.
9. **Федоров В. В.** Теория оптимального эксперимента. — М.: Наука, 1971. — 312 с.
10. **Szabados J.** On a problem of P. Erdos / Acta Math. Hungar. 1966. Vol. 17. P. 155 – 157.
11. **Erdos P.** Problems and results on the theory of interpolation. II / Acta Math. Acad. Sci. Hung. 1961. Vol. 12. P. 235 – 244.
12. **Bandemer H., Bellman A., Jung W., Richter K.** Optimale Versuchsplanung. Bd. 131. — Berlin: Akademie-Verlag GmbH, Wissenschaftliche Taschenbücher, 1973. — 180 s.
13. **Imhof L.** Optimum exact designs for polynomial regression. Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften, Der Reinisch-Westfälischen Technischen Hochschule Aachen, 1997. — 40 p.
14. **Карлин С., Стадден В.** Чебышевские системы и их применение в анализе и статистике. — М.: Наука, 1976. — 568 с.
15. **Fejer L.** Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein möglichst kleines Maximum besitzt / Ann. Scuola Norm. Sup. 1932. Pisa Ser. II. N 1. P. 263 – 276.
16. **Григорьев Ю. Д.** Методы оптимального планирования эксперимента: линейные модели. — СПб.: Лань, 2015. — 320 с.
17. **Pukelsheim F.** Optimal design of experiments. — Philadelphia: Society for Industrial and Applied Mathematics, 2006. — 454 p.
18. **Бухбергер Б.** Базисы Гребнера. Алгоритмический метод в теории полиномиальных идеалов. — В кн.: Компьютерная алгебра. Символьные и алгебраические вычисления. — М.: Мир, 1986. С. 331 – 372.
19. **Калинина Е. А., Утешев А. Ю.** Теория исключения: Учеб. пособие. — СПб: Изд-во НИИ химии СПбГУ, 2002. — 72 с.
20. **Bikker P., Uteshev A. Yu.** On the Bezout construction of the resultant / J. Symbolic Computation. 1999. Vol. 28. N 1. P. 45 – 88.

REFERENCES

1. **Guchenko R., Melas V. B.** Efficient computation of Bayesian optimal discriminating designs / J. Comput. and Graphic. Statist. 2017. Vol. 24. N 2. P. 424 – 433.
2. **Melas V. B., Shpilev P.** T-optimal discriminating designs for Fourier regression models / Comput. Statis. and Data Analysis. 2017. Vol. 113. P. 196 – 206.
3. **Guchenko R., Melas V. B., Wong W. K.** Optimal discrimination designs for semi-parametric models / Biometrika. 2018. Vol. 105. N 1. P. 185 – 197.
4. **Ermakov S. M., Semenchikov D. N.** On optimization methods in the problems of experiment design / Zavod. Lab. Diagn. Mater. 2019. Vol. 85. N 1. Part I. P. 72 – 77 [in Russian].
5. **Grigoriev Yu. D., Melas V. B., Shpilev P. V.** Excess of Locally D-optimal Designs and Homothetic Transformations / Vestn. St. Petersburg University. Mathematics. 2017. Vol. 50. N 4. P. 329 – 336.
6. **Grigoriev Yu. D., Melas V. B., Shpilev P. V.** Excess of locally D-optimal designs for Cobb-Douglas model / Statistical Papers. 2018. Vol. 59. N 4. P. 1425 – 1439.
7. **Fedorov V. V., Malytov M. B.** On the designs for certains weighted polynomial regression minimizing the average vari-

ance. Preprint. N 8, LSR. — Moscow: State University Press, 1969. — 12 p.

8. **Malyutov M. B., Fedorov V. V.** On weighted polynomial regression designs with minimum average variance / Theory Probab. Appl. 1971. Vol. 16. P. 716 – 720.
9. **Fedorov V. V.** Theory of Optimal Experiments / Translated by W. J. Studden and E. M. Klimko. — New York: Academic Press, 1972. — 292 p.
10. **Szabados J.** On a problem of P. Erdos / Acta Math. Hungar. 1966. Vol. 17. P. 155 – 157.
11. **Erdos P.** Problems and results on the theory of interpolation. II / Acta Math. Acad. Sci. Hung. 1961. Vol. 12. P. 235 – 244.
12. **Bandemer H., Bellman A., Jung W., Richter K.** Optimale Versuchsplanung. Bd. 131. — Berlin: Akademie-Verlag GmbH, Wissenschaftliche Taschenbücher, 1973. — 180 s.
13. **Imhof L.** Optimum exact designs for polynomial regression. Dissertation zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften, Der Reinisch-Westfälischen Technischen Hochschule Aachen, 1997. — 40 p.
14. **Karlin S., Studden W. J.** Tchebycheff Systems: With Applications in Analysis and Statistics. — New York: Wiley, 1966. — 586 p.
15. **Fejer L.** Bestimmung derjenigen Abszissen eines Intervalles, für welche die Quadratsumme der Grundfunktionen der Lagrangeschen Interpolation im Intervalle ein möglichst kleines Maximum besitzt / Ann. Scuola Norm. Sup. 1932. Pisa Ser. II. N 1. P. 263 – 276.
16. **Grigoriev Yu. D.** The Methods of the Optimal Experimental Design: Linear Models. — St. Petersburg: Lan’, 2015. — 320 p. [in Russian].
17. **Pukelsheim F.** Optimal design of experiments. — Philadelphia: Society for Industrial and Applied Mathematics, 2006. — 454 p.
18. **Buchberger B.** Gröbner Bases: An Algorithmic Method in Polinomial Ideal Theory. In: Computer Algebra. Symbolic and Algebraic Computation. Publ. N 83-29.0. Nov. 1983.
19. **Kalinina E. A., Uteshev A. Yu.** Exclusion Theory: Study Guide. — St. Petersburh: Izd. NII khimii SPbGU, 2002. — 72 p. [in Russian].
20. **Bikker P., Uteshev A. Yu.** On the Bezout construction of the resultant / J. Symbolic Computation. 1999. Vol. 28. N 1. P. 45 – 88.