ISSN 1814-1196
Научный вестник НГТУ
том 56, № 3, 2014, с. 67–78

http://journals.nstu.ru/vestnik
Scientific Bulletin of NSTU
Vol. 56, No. 3, 2014, pp. 67–78

ОБРАБОТКА ИНФОРМАЦИИ INFORMATION PROCESSING

УДК 519.24

Об одном выпуклом критерии оптимальности планов эксперимента*

Ю.Д. ГРИГОРЬЕВ[1], В.Н. ЛАПТЕВ[2]

[1] 197376, РФ, г. Санкт-Петербург, ул. проф. Попова, 5, Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова, д. т. н., профессор, e-mail: yuri_grigoriev@mail.ru
[2] 630087, РФ, г. Новосибирск, пр. К. Маркса, 26, Сибирский университет потребительской кооперации, к. т. н., доцент

В качестве критериев оптимальности планов эксперимента обычно используют различные выпуклые функционалы от информационной матрицы плана. Как правило, это функционалы типа матричных средних. К числу таких функционалов, в частности, относятся такие хорошо известные функционалы, как критерии Д-, А- и Е-оптимальности.

В данной работе в качестве критерия оптимальности планов регрессионных экспериментов предлагается λ-функционал типа матричной дисперсии. Доказывается выпуклость этого функционала и устанавливаются необходимые и достаточные условия оптимальности соответствующих планов. При доказательстве выпуклости λ-функционала существенно используется понятие выпуклости по Шуру.

Для иллюстрации полученных результатов в работе рассмотрены два примера. В обоих случаях в качестве функции регрессии выбирается квадратичная парабола, заданная на отрезке [–1, 1]. В первом примере получена аналитическая зависимость целевого функционала от весов точек плана, подтверждающая его выпуклость. Во втором примере планы эксперимента для квадратичной параболы, оптимальные по различным критериям, сравниваются с аналитически найденным λ-планом.

В заключении предлагается еще один критерий оптимальности, который не входит в класс матричных средних и представляет собой коэффициент обусловленности информационной матрицы плана. Высказывается предположение о его выпуклости. Для данного критерия также построен соответствующий план, удовлетворяющий необходимому условию оптимальности. Однако достаточные условия оптимальности данного плана пока остаются неизвестными.

Ключевые слова: план эксперимента, выпуклый функционал, информационная функция, класс матричных средних, матричная дисперсия, λ-функционал, порядок Левнера, выпуклость по Шуру, квадратичная регрессия, неравенство Беллмана, условия оптимальности

ВВЕДЕНИЕ

В задачах планирования эксперимента (Федоров [8]) важную роль играют критерии оптимальности, определяющие выбор того или иного непрерывного плана ε. Одним из основных требований, предъявляемых к критериям оптимальности планов эксперимента, является их выпуклость. Функционал $\varphi : \mathcal{M} \to R^1$, заданный на выпуклом замкнутом конусе $\mathcal{M} \subset R^{m\times m}$ информационных матриц $M(\varepsilon)$, называется выпуклым, если для любых $A, B \in \mathcal{M}$ имеет место неравенство

$$\varphi[\alpha A + (1-\alpha)B] \le \alpha\varphi(A) + (1-\alpha)\varphi(B), \quad 0 < \alpha < 1. \tag{1}$$

* *Статья получена 09 марта 2013 г. Повторно после исправлений 18 июля 2014 г.*
Работа выполнена в СПбГЭТУ при финансовой поддержке Министерства образования и науки Российской Федерации в рамках договора №02.G25.31.0058 от 12.02.2013 г.

Хорошо известен класс матричных средних Кифера $\Phi=\{\phi_p\}_{p=-\infty}^{1}$ (Кифер [10]; Пукельсхайм [13]), элементами которого

$$\phi_p(M)=[m^{-1}tr(M^p(\varepsilon))]^{1/p}=(m^{-1}\sum_{i=1}^{m}\lambda_i^p[M(\varepsilon)])^{1/p} \qquad (2)$$

являются так называемые *информационные функции*. Они обладают многими положительными свойствами, которые позволяют использовать их в качестве критериев оптимальности $\varphi_p[M(\varepsilon)]\to\max$ планов эксперимента ε. В (2) величины $\lambda_i[M(\varepsilon)]$ – это собственные числа матрицы $M(\varepsilon)$, при этом в результате предельного перехода полагают

$$\varphi_{-\infty}(M)=\lim_{p\to-\infty}\varphi_p(M)=\lambda_{\min}[M].$$

Эти функции являются не только вогнутыми (при необходимости их приводят к выпуклым), но и *изотонными*, т. е. сохраняющими в диапазоне их значений порядок Левнера. Последнее означает, что если $A,B\in\mathcal{M}$, то

$$A\ge B\Rightarrow\phi_p(A)\ge\phi_p(B).$$

Три наиболее важных для приложений элемента ϕ_0, ϕ_{-1} и $\phi_{-\infty}$ класса Φ представляют критерии *D*-, *A*- и *E*-оптимальности соответственно. Изотонность этих критериев означает, что имеют место импликации: $A,B\in\mathcal{M}$ и $A\ge B$, то

$$|A|\ge|B|,\ trA\ge trB,\ \lambda_{\min}|A|\ge\lambda_{\min}|B|.$$

Однако наряду с изотонными критериями, образующими класс матричных средних Ф, встречаются и другие критерии оптимальности, которые также бывают востребованы на практике. К их числу относятся, например, критерии отрогональности, роттатабельности, экстраполяции в точку и многие другие. В предлагаемой работе исследуется λ-критерий оптимальности типа матричной дисперсии, для которого доказывается, что он является выпуклым на конусе $\mathcal{M}$.

1. ПОСТАНОВКА ЗАДАЧИ

Согласно определению $M(\varepsilon)\in\mathcal{M}$, где $M(\varepsilon)$ – информационная матрица любого непрерывного плана ε (Федоров [8]). Если $A\in\mathcal{M}$, то записываем $A\ge 0$. В частности, если $A-B\in\mathcal{M}$, то записываем $A\ge B$. Аналогично, если $\mathcal{K}\subset R^m$ – выпуклый замкнутый конус неотрицательных векторов *x*, то записываем $x\ge 0$, при этом если $x-y\in\mathcal{K}$, то записываем $x\ge y$.

Пусть $\mathcal{M}_0$ – открытый выпуклый конус положительно определенных информационных матриц, $\mathcal{M}_0\subset\mathcal{M}$, Ξ – множество всех непрерывных планов. В работе исследуется критерий λ-оптимальности $\varphi:\mathcal{M}_0\to(0,\infty)$, представляющий, в отличие от матричных средних, «дисперсию» собственных чисел $\lambda_i=\lambda_i[M^{-1}(\varepsilon)]$ дисперсионной матрицы $D(\varepsilon)=M^{-1}(\varepsilon)$:

$$\varphi[M(\varepsilon)]\overset{def}{=}\sum_{i=1}^{m}(\lambda_i-\bar\lambda)^2\to\min_{\varepsilon\in\Xi},\quad \bar\lambda=m^{-1}\sum_{i=1}^{m}\lambda_i. \qquad (3)$$

Критерий (2) не входит в класс Φ, но, как будет показано, является выпуклым, что в порядке гипотезы было высказано в (Лаптев [6]).

Целью работы является доказательство выпуклости критерия (3). Оно необходимо для обоснования предложенных в (Лаптев [6]) достаточных условий оптимальности λ-планов ε^*.

Кроме доказательства выпуклости в работе рассмотрены два примера. Первый из них имеет чисто иллюстративное значение. Он демонстрирует выпуклость λ-функционала для двумерной линейной регрессии. В качестве второго примера рассмотрена одномерная квадратичная регрессия, для которой проведено сравнение λ-оптимального плана с планами, оптимальными по другим критериям, в том числе не входящими в класс матричных средних.

2. ВСПОМОГАТЕЛЬНЫЕ РЕЗУЛЬТАТЫ

Для доказательства основного результата, т. е. выпуклости λ-функционала (3), нам понадобится несколько вспомогательных утверждений.

1. Прежде всего отметим, что возможна другая форма записи функционала (3)

$$\sum_{i=1}^{m}(\lambda_i-\bar{\lambda})^2 = m^{-1}\sum_{1\le i<j\le m}(\lambda_i-\lambda_j)^2, \tag{4}$$

которая следует из тождества Лагранжа (Беллман [3], с. 74)

$$(\sum_i x_i^2)(\sum_i y_i^2)-(\sum_i x_i y_i)^2 = \frac{1}{2}\sum_i\sum_j (x_i y_j - x_j y_i)^2, \quad x,y\in R^m. \tag{5}$$

Для этого в тождестве (5) нужно положить $y=m^{-1}1_m$, $1_m=(1,...,1)\in R^m$. Однако в большей степени нам нужна другая эквивалентная форма записи функционала (3).

Обозначим $\mathcal{H}_m$ – множество всех симметричных матриц $A\in R^{m\times m}$. Определим на нем *ковариацию* матриц

$$\langle C,D\rangle \overset{def}{=} trCD-m^{-1}(trC)(trD),\ C,D\in\mathcal{H}_m. \tag{6}$$

Тогда

$$\mathbb{D}\mathrm{C} = trC^2 - m^{-1}tr^2C \tag{7}$$

– *дисперсия* матрицы *C*. Эта терминология оправдана в том смысле, что согласно неравенству Беллмана

$$trC^2 \ge m^{-1}tr^2C, \tag{8}$$

которое справедливо для любых симметричных матриц (Беллман [3, с. 168], имеем $\mathbb{D}\mathrm{C}\ge 0$.

Аналогично введем дисперсию вектора λ в евклидовом пространстве R^m, отождествив ее с λ-функционалом (3).

Будем говорить, что множества векторов $\mathcal{R}_m\subseteq R^m$ и всех симметричных матриц $\mathcal{H}_m$ *однотипны*, если между их элементами установлено соответствие – каждой матрице $C\in\mathcal{H}_m$ сопоставляется вектор $\lambda=(\lambda_1,...,\lambda_m)\in\mathcal{R}_m$ ее собственных значений, т. е. $\lambda_i=\lambda_i[C]$, $i = 1,\dots, m$. Отметим, что данное соотношение не является биективным.

Лемма 1. *Пусть* $\lambda_i=\lambda_i[A]$ *– собственные числа матрицы* $A\in\mathcal{H}_m\subset R^{m\times m}$, *т. е. множества* $\mathcal{R}_m$ *и* $\mathcal{H}_m$ *однотипны. Тогда*

$$\mathbb{D}\lambda = \mathbb{D}\mathrm{A}. \tag{9}$$

Д о к а з а т е л ь с т в о. Утверждение (9) следует из равенств

$$trA=\sum_{i=1}^{m}\lambda_i,\quad trA^2=\sum_{i=1}^{m}\lambda_i^2.$$

Первое равенство очевидно. Представим A в виде $U\Lambda U^T$, где U – ортогональная матрица, $\Lambda = diag\{\lambda_1,...,\lambda_m\}$. Исходя из такого представления матрицы A определим функцию $g(A)$ формулой $g(A) = U\Lambda_g U^T$, где $\Lambda_g = diag\{g(\lambda_1),...,g(\lambda_m)\}$. Отсюда следует, что собственные числа матрицы $g(A)$ получаются из собственных чисел матрицы A в результате «воздействия» на них функции g. В нашем случае $g : x \to x^2$. Отсюда следует второе равенство. □

2. Для однотипных множеств $\mathcal{R}_m$ и $\mathcal{H}_m$ введем общий термин. Именно, будем рассматривать их как гильбертово пространство L^2 случайных величин (св.) со скалярным произведением $\langle\xi,\eta\rangle = M\xi\eta$. Тогда $cov(\xi,\eta) = \langle\xi,\eta\rangle - \mathbb{M}\xi\mathbb{M}\eta$ – ковариация (св.) ξ и η, $\mathbb{D}\xi = cov(\xi,\xi)$ – дисперсия св. ξ.

Лемма 2. *Дисперсия* $\varphi : \xi \to \mathbb{D}\xi$ *– выпуклая функция, т. е. если* $\xi,\eta \in L^2$, то

$$\varphi[\alpha\xi + (1-\alpha)\eta] \le \alpha\varphi(\xi) + (1-\alpha)\varphi(\eta), \quad 0 < \alpha < 1. \tag{10}$$

Д о к а з а т е л ь с т в о. Используя свойства дисперсии и определение ковариации, последовательно получаем

$$\begin{aligned}\varphi[\alpha\xi + (1-\alpha)\eta] &= \alpha^2\varphi(\xi) + 2\alpha(1-\alpha)cov(\xi,\eta) + (1-\alpha)^2\varphi(\eta)\\ &= \alpha\varphi(\xi) + (1-\alpha)\varphi(\eta) - \alpha(1-\alpha)\varphi(\xi-\eta)\\ &\le \alpha\varphi(\xi) + (1-\alpha)\varphi(\eta),\end{aligned}$$

что совпадает с (10). □

3. Следующая лемма устанавливает взаимосвязь между векторным упорядочением, индуцируемым выпуклым конусом $\mathcal{K} \subset R^m$ собственных чисел симметричных матриц $\mathcal{H}_m$, и матричным упорядочением, индуцируемым выпуклым конусом $\mathcal{M} \subset \mathcal{H}_m \subset R^{m\times m}$ этих матриц. Это упорядочение, необходимое для определения монотонности матричнозначных функций от матриц, является частичным порядком и называется *порядком Левнера*. Соответствующая запись $A \le B$, как уже отмечалось, означает, что $B - A$ является неотрицательно определенной матрицей.

Лемма 3. (*Маршалл, Олкин* [7], с. 480). *Пусть выполнены следующие условия*:
(1) $A, B \in \mathcal{M}$;
(2) $\lambda = (\lambda_i[A])_{i=1}^m$, $\mu = (\mu_i[B])_{i=1}^m \in \mathcal{K}$ *– векторы собственных чисел матриц A и B*;
(3) координаты векторов λ *и* μ *упорядочены по невозрастанию*

$$\forall i = 1,...,m-1 : \lambda_i \ge \lambda_{i+1}, \quad \mu_i \ge \mu_{i+1}.$$

Тогда
$$A \le B \Rightarrow \lambda \le \mu. \tag{11}$$

В (Беккенбах, Беллман [2], с. 109) отмечается, что это утверждение непосредственно следует из минимаксной теоремы Куранта – Фишера (Беллман [3], с. 143). Доказательство леммы 3 содержится в (Левнер [11]). □

4. Пусть $\varphi : R^1 \to R^1$ – некоторая функция. Продолжим φ до функции вида $\psi : \mathcal{H}_m \to \mathcal{H}_m$ способом, описанным выше при доказательстве леммы 1, а именно: для всякой $A = U\Lambda U^T \in \mathcal{H}_m$ полагаем $\psi(A) = U\Lambda_\varphi U^T$. Произвольные матричнозначные функции ψ от матриц вводятся естественным образом с помощью стандартных матричных операций (Маршалл, Олкин [7], с. 466).

Пусть $\mathcal{M} \subset \mathcal{H}_m$. Функция $\varphi : \mathcal{M} \to \mathcal{H}_m$ называется матричновозрастающей на $\mathcal{M}$, если отношение $A \le B$ имплицирует неравенство $\varphi(A) \le \varphi(B)$ всякий раз, когда $A, B \in \mathcal{M}$. В контексте нашей задачи конус $\mathcal{M}$ – это множество всех неотрицательных определенных матриц, φ дисперсия матрицы $A \in \mathcal{M}$, неубывание которой можно доказать.

Функция $g : \mathcal{M} \to \mathcal{H}_m$ называется матрично выпуклой, если

$$\forall A, B \in \mathcal{M} : g[\alpha A + (1-\alpha)B] \le \alpha g(A) + (1-\alpha) g(B), \quad 0 \le \alpha \le 1. \tag{12}$$

Если неравенство (12) строгое, $0 < \alpha < 1$ и матрица $B - A$ имеет полный ранг, то функция g называется *матрично строго выпуклой*. Нас будет интересовать функция $g(A) = A^{-1}$, заданная на открытом конусе $\mathcal{M}_0$ положительно определенных матриц. Очевидно, $\mathcal{M}_0 \to \mathcal{M} \to \mathcal{H}_m$.

Лемма 4. (*Маршалл, Олкин* [7], с. 474). *На конусе $\mathcal{M}_0$ функция $g(A) = A^{-1}$ является матрично строго выпуклой, т. е. если $A, B \in \mathcal{M}_0$, то*

$$[\alpha A + (1-\alpha)B]^{-1} < \alpha A^{-1} + (1-\alpha)B^{-1}, \quad 0 < \alpha < 1, \tag{13}$$

причем знак равенства имеет место, если только $A = B$.

Утверждение (13) можно найти в (Федоров [8], с. 24). Доказательство леммы приводится в (Олкин, Пратт [12]; Уиттл [14]). □

3. ВЫПУКЛОСТЬ λ-ФУНКЦИОНАЛА

Прежде чем приступить к доказательству выпуклости λ-функционала (3), введем необходимое в дальнейшем понятие выпуклости по Шуру (Маршалл, Олекин [7]). Пусть $\le$ – упорядочение Левнера на некотором множестве $\mathcal{A} \subset R^m$. Вещественная функция $\varphi : \mathcal{A} \to R^1$ называется выпуклой в смысле Шура, или *изотонной*, если для всех $x, y \in \mathcal{A}$

$$x \le y \Rightarrow \varphi(x) \le \varphi(y).$$

Выпуклость по Шуру, очевидно, отличается от выпуклости по Йенсену (1).

Как уже отмечалось выше, выпуклость по Шуру является желательным свойством критериев оптимальности планов эксперимента. Теперь мы можем приступить к доказательству основного результата – выпуклости λ-функционала (4).

Теорема 1. Пусть выполнены следующие условия:
(1) $M_i = M(\varepsilon_i) \in \mathcal{M}_0$ – информационные матрицы невырожденных планов ε_i, $i = 1, 2$;
(2) $\varphi : \mathcal{M}_0 \to (0, \infty)$ – λ-функционал (4).

Тогда
$$\varphi\{[\alpha M_1 + (1-\alpha)M_2]^{-1}\} \le \alpha\varphi(M_1^{-1}) + (1-\alpha)\varphi(M_2^{-1}), \quad 0 < \alpha < 1, \tag{14}$$

т. е. функционал φ – выпуклый.

Доказательство. Положим $A = M(\varepsilon_1)$, $B = M(\varepsilon_2)$. Обозначим

$$C = [\alpha A + (1-\alpha)B]^{-1}, \quad D = \alpha A^{-1} + (1-\alpha)B^{-1}.$$

Согласно (13) имеем $C \le D$. Пусть $\lambda = (\lambda_i[C])_{i=1}^m$ и $\mu = (\mu_i[D])_{i=1}^m$ – векторы монотонно упорядоченных собственных чисел матриц C и D. Тогда согласно лемме **3** получаем $\lambda \le \mu$. Поскольку дисперсия $\varphi : \mathcal{R}_m \to [0, \infty)$ является выпуклой по Шуру (Маршалл, Олкин [7], с. 417), то отсюда следует, что $\varphi(\lambda) \le \varphi(\mu)$.

С другой стороны, согласно (9) заключаем,что

$$\varphi(\lambda) = \varphi(C), \tag{15}$$

где $\lambda \in \mathcal{R}_m$ – вектор собственных значений матрицы $C \in \mathcal{H}_m$.Тем самым равенство (15) означает, что дисперсия $\varphi : \mathcal{H}_m \to [0,\infty)$ также является выпуклой функцией. Отсюда, возвращаясь к определению матриц C и D, получаем

$$\varphi\{[\alpha A + (1-\alpha)B]^{-1}\} \le \varphi[\alpha A^{-1} + (1-\alpha)B^{-1}]. \qquad (16)$$

Положим $\xi = A^{-1}$, $\eta = B^{-1}$ и будем рассматривать правую часть неравенства (16) как дисперсию св. $\tau = \alpha\xi + (1-\alpha)\eta \in \mathcal{H}_m$, где $\mathcal{H}_m$ рассматриваем как гильбертово пространство L^2. Поскольку дисперсия $\mathbb{D}\tau$ согласно лемме **2** выпукла, т. е. удовлетворяет неравенству (10), то отсюда следует требуемый результат (14). □

Пример 1. Пусть $\eta(x,\theta) = \theta_0 + \theta_1 x_1 + \theta_2 x_2$ – функция регрессии, заданная на квадрате $X = [-1,1]^2$. Обозначим $A(1,1)$, $B(-1,1)$, $C(-1,-1)$ вершины квадрата X, $O(0,0)$ – его центр. Для оценивания параметра $\theta \in R^3$ выберем два плана эксперимента

$$\varepsilon_1 = \left\{\xi(A) = \xi(B) = \xi(C) = \frac{1}{3}\right\}, \quad \varepsilon_2 = \left\{\xi(A) = \xi(B) = \xi(C) = \xi(O) = \frac{1}{4}\right\}$$

с матрицами Якоби

$$F_1 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & -1 & -1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Вычислим $M_1 = M(\varepsilon_1) = \frac{1}{3}F_1^T F_1$, $M_2 = M(\varepsilon_2) = \frac{1}{4}F_2^T F_2$ и $M = \alpha M_1 + (1-\alpha)M_2$. Затем найдем $D_i = M_i^{-1}$ и $D = M^{-1}$, $i = 1,2$. Зная матрицы D_i и D, найдем их собственные значения, причем для матрицы D эти значения будут функциями α. После этого определим хорду $G(\alpha) = \alpha\varphi(D_1) + (1-\alpha)\varphi(D_2)$, $\alpha \in [0,1]$ и функцию $L(\alpha) = \varphi(D)$, где $\varphi(D)$ – λ-функционал. В результате проведенных вычислений получаем

$$L(\alpha) = \frac{18(5\alpha^2 + 14\alpha + 45)}{(\alpha+3)^2(\alpha-9)^2}, \quad G(\alpha) = \frac{1}{9}\left(10 + \frac{\alpha}{8}\right), \quad \alpha \in [0,1].$$

Легко проверить, что $L''(\alpha) > 0$. Поэтому функция $L : \alpha \to R^1$ выпукла. Отсюда следует $G(\alpha) \ge L(\alpha)$, $\alpha \in [0,1]$. Значения функций L и G протабулированы и представлены в табл. 1. Минимум $L(\alpha)$ достигается в точке

$$\alpha_0 = \frac{1}{5}\left(2\sqrt[3]{484 + 20\sqrt{1241}} - \frac{128}{\sqrt[3]{484 + 20\sqrt{1241}}} - 7\right) = 0.4203,$$

при этом $L(0.4203) = 1.082073$, в то время как $L(0) = \frac{10}{9} = 1.1111$, $L(1) = \frac{9}{8} = 1.1250$. Таким образом, результаты вычислений не противоречат справедливости теоремы **1** о выпуклости λ-функционала (3). □

Таблица 1

λ -критерий: табулированные значения функций *L* и *G*

	$\alpha = 0.1\,(0.1)\,0.9$								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
$L(\alpha)$	1.0984	1.0895	1.0842	1.0821	1.0830	1.0865	1.0927	1.1012	1.1120
$G(\alpha)$	1.1125	1.1139	1.1153	1.1167	1.1181	1.1194	1.1208	1.1222	1.1236

4. УСЛОВИЯ ОПТИМАЛЬНОСТИ λ -ПЛАНОВ

Поскольку λ-функционал является выпуклым, то к нему применима теорема Федорова о выпуклых функционалах.

Теорема 2 (Федоров [9]). Положим $\mathcal{M}$ – конус информационных матриц $M(\varepsilon)$, $\Psi : \mathcal{M} \to R^1$ – критерий оптимальности. Пусть выполнены следующие условия:

(1) $\eta(x,\theta) = f^T(x)\theta$ – функция регрессии, $x \in X \subset R^k$, $\theta \in \Theta \subset R^m$;

(2) $f \in C^1(X)$, X – компакт, Θ – открытое множество;

(3) $\Psi \in C^1(\mathcal{M})$ – выпуклый функционал, т. е.

$$\Psi[\alpha M(\varepsilon_1) + (1-\alpha)M(\varepsilon_2)] \le \alpha\Psi[M(\varepsilon_1)] + (1-\alpha)\Psi[M(\varepsilon_2)], \quad 0 \le \alpha \le 1;$$

(4) $\psi(x,\varepsilon) = f^T(x)\dfrac{\partial\Psi}{\partial M} f(x)\,|_{M=M(\varepsilon)}$.

Тогда план ε^* оптимален, если и только если

$$\min_{x\in X}\psi(x,\varepsilon^*) = \mathrm{tr}M\left.\frac{\partial\Psi}{\partial M}\right|_{M=M(\varepsilon^*)}, \tag{17}$$

причем в узлах плана (за исключением, быть может, множества точек меры нуль) функция $\psi(x,\varepsilon^*)$ достигает своей нижней грани.

Опираясь на теорему **2**, можно сформулировать необходимые и достаточные условия оптимальности для непрерывных λ -планов ε^*. Именно, имеет место

Теорема 3 (Лаптев [6]). Пусть выполнены условия:

(1) $\eta(x,\theta) = f^T(x)\theta$ – функция регрессии, $x \in X \subset R^k$, $\theta \in \Theta \subset R^m$;

(2) $f \in C^1(X)$, X – компакт, Θ – открытое множество;

(3) $\nu = m^{-1}trD(\varepsilon)$, $G(\varepsilon) = D(\varepsilon) - \nu I_m$.

Тогда следующие утверждения эквивалентны:

(1) $\varepsilon^* = \arg\min\limits_{\varepsilon\in\Xi}[trD^2(\varepsilon) - m^{-1}tr^2D(\varepsilon)]$ – λ -оптимальный план;

(2) имеет место равенство

$$\max_{x\in X} f^T(x)D(\varepsilon^*)G(\varepsilon^*)D(\varepsilon^*)f(x) = \mathrm{tr}G^2(\varepsilon^*). \tag{18}$$

Доказательство. *Необходимость.* $(1) \Rightarrow (2)$. В силу теоремы **1** функционал (3) является выпуклым. Поэтому согласно теореме **2** для Ψ-оптимального плана выполняется равенство (17):

$$\min_{x\in X} f^T(x)\frac{\partial\Psi}{\partial M} f(x) = trM\frac{\partial\Psi}{\partial M}. \tag{19}$$

Поскольку $D = M^{-1}$ и $\dfrac{\partial trD}{\partial M} = -D^2$, то

$$\frac{\partial trD^2}{\partial M} = -2D^3, \quad \frac{\partial tr^2 D}{\partial M} = -2(trD)D^2.$$

Отсюда находим

$$\frac{\partial \Psi}{\partial M} = \frac{\partial [trD^2(\varepsilon) - m^{-1} tr^2 D(\varepsilon)]}{\partial M} = -2D(\varepsilon)G(\varepsilon)D(\varepsilon),$$

$$trM(\varepsilon)\frac{\partial \Psi}{\partial M} = -2[trD^2(\varepsilon) - m^{-1} tr^2 D(\varepsilon)].$$

Теперь с учетом равенства $trG^2 = trD^2 - m^{-1}tr^2 D$, перемены в двух последних равенствах знака минус на плюс и изменения в (19) операции min на max, приходим к (18).

Достаточность. $(2) \Rightarrow (1)$. Доказательство этой части теоремы никак не зависит от конкретного вида выпуклого функционала Ψ. Поэтому здесь применим ход доказательства теоремы **2**, изложенный в (Федоров [9]), а именно: доказательство от противного импликации $\neg(1) \Rightarrow \neg(2)$ (если план ε^* неоптимален, то для него не выполнено условие (19). □

5. КВАДРАТИЧНАЯ РЕГРЕССИЯ

Для наглядной характеризации различных критериев оптимальности, включая λ-критерий, рассмотрим одномерную квадратичную регрессию

$$\eta(x,\theta) = \theta_0 + \theta_1 x + \theta_2 x^2, \quad x \in X = [-1,1]. \tag{20}$$

Для регрессии (20) построим оптимальные по различным критериям планы эксперимента в классе симметричных планов Ξ_0. Последнее означает, что предполагаемые оптимальные планы должны иметь вид

$$\varepsilon = \{\xi(\pm 1) = p, \quad \xi(0) = 1 - 2p\}, \tag{21}$$

где $\xi(x)$ – вероятностная мера, или вес, точки $x \in S(\varepsilon)$, $S(\varepsilon)$ – известный спектр плана ε. Таким образом, требуется определить величину p, которая для различных критериев должна быть разной.

В общем случае оптимальность планов $\varepsilon \in \Xi_0$ для различных критериев оптимальности проверяется с помощью соответствующих теорем эквивалентности. Однако для критериев из класса матричных средних эти теоремы являются следствием общего решения полярного уравнения (Пукельсхайм [13]). С помощью этого уравнения доказана оптимальность планов $\varepsilon \in \Xi_0$ по критериям из класса матричных средних Φ для полиномиальной регрессии любой степени $r \ge 1$ на $X = [-1,1]$.

1. Информационная матрица плана $M(\varepsilon)$ и обратная к ней $D(\varepsilon) = M^{-1}(\varepsilon)$ имеют вид

$$M(\varepsilon) = \begin{pmatrix} 1 & 0 & 2p \\ 0 & 2p & 0 \\ 2p & 0 & 2p \end{pmatrix}, \quad D(\varepsilon) = \begin{pmatrix} \dfrac{1}{1-2p} & 0 & -\dfrac{1}{1-2p} \\ 0 & \dfrac{1}{2p} & 0 \\ -\dfrac{1}{1-2p} & 0 & \dfrac{1}{2p(1-2p)} \end{pmatrix}.$$

Вычисляем собственные числа $\lambda_i = \lambda_i[D(\varepsilon)]$, полагая $\lambda_1 \geq \lambda_2 \geq \lambda_3$:

$$\lambda_1 = \frac{1+2p+\sqrt{D}}{4p(1-2p)}, \quad \lambda_2 = \frac{1}{2p}, \quad \lambda_3 = \frac{1+2p-\sqrt{D}}{4p(1-2p)}, \quad D = \sqrt{20p^2-4p+1}.$$

Отсюда

D-критерий: $$|D(\varepsilon)| = \lambda_1\lambda_2\lambda_3 = \frac{1}{4p^2(1-2p)};$$

A-критерий: $$trD(\varepsilon) = \lambda_1+\lambda_2+\lambda_3 = \frac{1}{p(1-2p)};$$

E-критерий: $$\lambda_{\max}[D(\varepsilon)] = \lambda_1 = \frac{1+2p+\sqrt{D}}{4p(1-2p)};$$

λ-критерий: $$\lambda[D(\varepsilon)] = \sum_{i=1}^{3}(\lambda_i - \frac{1}{3}\sum_{j=1}^{3}\lambda_j)^2 = \frac{24p^2-6p+1}{6p^2(1-2p)^2};$$

C-критерий: $$cond[D(\varepsilon)] = \frac{\lambda_{\max}[D(\varepsilon)]}{\lambda_{\min}[D(\varepsilon)]} = \frac{\lambda_1}{\lambda_3} = \frac{1+2p+\sqrt{D}}{1+2p-\sqrt{D}};$$

G-критерий: $$G[D(\varepsilon)] = \max_{-1\leq x\leq 1} d(x,\varepsilon^*) = \begin{cases} \dfrac{1}{1-2p}, & x=0, \\ \dfrac{1}{p}, & x=1. \end{cases}$$

Отметим, что

$$\begin{aligned} Q[D(\varepsilon)] &= \int_{-1}^{1} d(x,\varepsilon)dx \\ &= \int_{-1}^{1}[D_{11}+(2D_{13}+D_{22})x^2+D_{33}x^4]dx = \frac{8}{15p(1-2p)}, \end{aligned}$$

т. е. A- и Q-оптимальные планы для квадратичной регрессии (20) совпадают. Оптимальные по различным критериям планы ε^* представлены в табл. 2, а значения соответствующих им функционалов – в табл. 3.

Таблица 2

Непрерывные оптимальные планы $\varepsilon^* = \{\xi(\pm 1) = p,\ \xi(0) = 1-2p\}$ для квадратичной регрессии (20)

	Критерий оптимальности				
	$\lvert D(\varepsilon)\rvert$	$trD(\varepsilon)$	$\lambda_{\max}$	$\sum_{i=1}^{3}(\lambda_i - \frac{1}{3}\sum_{j=1}^{3}\lambda_j)^2$	$\frac{\lambda_{\max}}{\lambda_{\min}}$
p	$\frac{1}{3} = 0.3334$	$\frac{1}{4} = 0.2500$	$\frac{1}{5} = 0.2000$	0.1823	$\frac{1}{6} = 0.1667$

Таблица 3

Характеристики планов $\varepsilon^* = \{\xi(\pm 1) = p,\ \xi(0) = 1 - 2p\}$ для квадратичной регрессии (20),
$\lambda_{\max} = \lambda_1[D(\varepsilon)]$, $\lambda_{\min} = \lambda_3[D(\varepsilon)]$, $\mathrm{cond}[D(\varepsilon)] = \lambda_1 / \lambda_3$

Функционалы	Оптимальные планы ε^*.				
	D-оптим.	A-оптим.	E-оптим.	λ-оптим.	cond-оптим.
$\vert D(\varepsilon^*)\vert$	6.7500	8.0	10.417	11.316	13.500
$trD(\varepsilon^*)$	9.0	8.0	8.3334	8.5201	9.0
$\lambda_{\max}[D(\varepsilon^*)]$	6.8423	5.2361	5.0	5.0142	5.1213
$\lambda[D(\varepsilon^*)]$	22.500	10.667	8.7963	8.7227	9.0
$\mathrm{cond}[D(\varepsilon^*)]$	10.404	6.8541	6.0	5.9019	5.8284
$G[D(\varepsilon^*)]$	3.0	4.0	5.0	5.3125	6.0

Анализ табл. 3 позволяет формулировать различные предпочтения на множестве функционалов от $M(\varepsilon)$. По существу, мы имеем дело с многокритериальной задачей (задачей с векторным критерием), и выбор соответствующих предпочтений определяется конкретными требованиями, предъявляемыми к решению. В частности, заметим, что критерии в табл. 2 перечислены в порядке убывания веса крайних точек.

Поскольку оптимальному плану ε^* соответствует минимальное значение выбранного критерия, то минимальный элемент каждой строки табл. 3 находится на главной диагонали квадратной табл. 3, в которой, для краткости, опущен шестой столбец, соответствующий критерию G-оптимальности. Это объясняется тем, что согласно теореме эквивалентности Кифера – Вольфовица (Федоров [8], с. 77) непрерывные D- и G-оптимальные планы ε^* совпадают.

ЗАКЛЮЧЕНИЕ

Критерий λ-оптимальности аналогичен принципу дисперсии назначения премий в страховании, в то время как критерий A-оптимальности соответствует принципу математического ожидания [5]. Эти принципы различны в том отношении, что первый из них инвариантен к сдвигу $\lambda_j \mapsto \lambda_j + c$, а второй, как, впрочем, и все остальные рассмотренные критерии – нет. Это означает, что λ-критерий занимает в этом отношении особое место и, возможно, более предпочтителен с точки зрения помехоустойчивости вычислений, чем E-критерий и, тем более, A-критерий. Именно это свойство E- и мажорирующих его критериев отмечалось в (Берсенев, Саблин [4]), см. по этому поводу также (Батасова, Гаврилов [1]).

Что касается C-критерия, то он представляет собой число обусловленности матрицы $D(\varepsilon)$ и также не относится к числу матричных средних. Следуя логике табл. 2 и статьи (Берсенев, Саблин [4]), можем заключить, что с точки зрения вычислительной устойчивости он представляет еще больший интерес, чем λ-критерий. Однако вопрос о выпуклости C-критерия и способах построения соответствующих планов остается открытым.

СПИСОК ЛИТЕРАТУРЫ

1. *Батасова В.С., Гаврилов В.А.* Оптимальное расширение локальных сейсмотелеметрических сетей Камчатки // Вулканология и сейсмология. – 1990. – Вып. 1. – С. 91–103.

2. *Беккенбах Э., Беллман Р.* Неравенства. – М.: Мир, 1965. – 276 с.

3. *Беллман Р.* Введение в теорию матриц. – М.: Наука, 1969. – 368 с.

4. *Берсенев С.М., Саблин Н.И.* Планирование помехоустойчивых регрессионных экспериментов // Заводская лаборатория. – 1976. – Т. 42, вып. 3. – С. 314–316.

5. Современная актуарная теория риска / Р. Каас, М. Гувертс, Ж. Дэнэ, М. Денут. – М.: Янус-К, 2008. – 376 с.

6. *Лаптев В.Н.* Некоторые вопросы синтеза оптимальных планов регрессионных экспериментов с помощью ЭВМ: дис. ... канд. техн. наук; Новосиб. электротехн. ин-т. – Новосибирск, 1974. – 129 с.

7. *Маршалл А., Олкин И.* Неравенства: теория мажоризации и ее приложения. – М.: Мир, 1983. – 576 с.

8. *Федоров В.В.* Теория оптимального эксперимента. – М.: Наука, 1971. – 312 с.
9. *Федоров В.В.* Численные методы построения оптимальных планов для регрессионных экспериментов // Кибернетика. 1975. – Вып. 1. – С. 124–130.
10. *Kiefer J.C.* General equivalence theory for optimum designs (approximate theory) // The Annals of Statistics. – 1974. – Vol. 2, № 5. – P. 849–879.
11. *Loevner C.* Über monotone Matrixfunctionen // Mathematische Zeitschrift. – 1934. – Vol. 38. – P. 177–216.
12. *Olkin I., Pratt J.* A multivariate Tschbychev's inequality // Annals of Mathematical Statistics. – 1958. –Vol. 29. – P. 226–234.
13. *Pukelsheim F.* Optimal Design of Experiments. – Philadelphia, USA: SIAM, 2006. – 453 p. – (Classics in Applied Mathematics; vol. 50).
14. *Whittle P.* A multivariate generalization of Tschbychev's inequality // The Quarterly Journal of Mathematics. – 1958. – Vol. 9, no. 1. – P. 232–240.

Григорьев Юрий Дмитриевич, доктор технических наук, профессор кафедры математического обеспечения и применения ЭВМ Санкт-Петербургского государственного электротехнического университета («ЛЭТИ»), зав. лабораторией планирования эксперимента, теории риска и информатики. Основные направления научных исследований – планирование эксперимента, теория нелинейной регрессии, математическая теория гармонии. Имеет более 120 публикаций. E-mail: yuri_grigoriev@mail.ru

Лаптев Валерий Николаевич, кандидат технических наук, доцент. Основные направления научных исследований – планирование эксперимента, теория автоматического регулирования. Имеет более 20 публикаций.

On some convex optimal criterion of an experimental designs[*]

Yu.D. GRIGOR'EV, $\boxed{\text{V.N. LAPTEV}}$

[1] *Saint-Petersburg state electrotechnical University, 5, Professor Popov street, St. Petersburg, 197376, Russian Federation, D.St, professor, e-mail: yuri_grigoriev@mail.ru*
[2] *Siberian University of consumer cooperation20 prospect Karla Marksa, Novosibirsk, 197376, Russian Federation, C.tS, associate professor*

As an optimality criteria of experimental designs various convex functionals an information matrix dependent usually are used. As a rule, this functionals to a matrix averages type are related. Such well-known functionals as criteria of D-, A- and E-optimality pertain to the case of such functionals.

In the given work as criterion of an optimality criterion for the regression experiments designs λ -functional of a matrix covariance type is suggested. The convexity of this functional is proved and the necessary and sufficient optimality conditions for this designs are established. At the proof of λ -functional the concept of Schur's convexity essentially is used.

In order to more clearly explain the obtained results two examples are considered. In a both cases the quadratic parabolas given on a segment [–1, 1] as a regression function are choosed. In the first example the analytical expression of a target functional as a function of a design knots weights is obtained. This example confirm the convexity of suggested λ -functional. In the second example the experimental designs for a quadratic parabola optimal by various criteria with analytically computed λ -design are compared.

In summary, one more optimality criterion which does not belong to a matrix averages class is suggested. This functional is a conditionality coefficient of the information matrix. The supposition about its convexity is stated. For the given criterion the corresponding design, satisfying to a necessary condition of an optimality also is constructed. However sufficient optimality conditions for the designs of this type at present are unknown.

Keywords: experimental design, convex functional, information function, class of matrix means, matrix variance, λ -functional, Loevner's ordering, Schur's convexity, quadratic regression, Bellman's inequality, optimality conditions

REFERENCES

1. Batasova V.S., Gavrilov V.A. Optimal'noe rasshirenie lokal'nykh seismotelemetricheskikh setei Kamchatki [Optimum expansion of local seismotelemetric networks of Kamchatka]. *Vulkanologiya i seismologiya – Journal of Volcanology and Seismology*, 1990, iss. 1, pp. 91-103.
2. Beckenbach E.F., Bellman R. *Inequalities*. Berlin, Göttingen, Heidelberg, Springer-Verlag, 1961. xii+198 p. (Russ. ed.: Bekkenbakh E., Bellman R. *Neravenstva*. Moscow, Mir Publ., 1965. 276 p.).
3. Bellman R.E. *Introdaction to Matrix Analysis*. New York, Toronto, London, McGraw-Hill Book Co. Inc., 1960. 348 p. (Russ. ed.: Bellman R.E. *Vvedenie v teoriyu matrits*. – M.: Nauka, 1969. 368 p.).

[*] *Received 09 March 2014. Repeatedly 18 July 2014.*

Work is performed in Saint-Petersburg state electrotechnical University with financial support of the Ministry of Education and Science of the Russian Federation within the contract № 02.G25.31.0058 of 12.02.2013.

4. Bersenev S.M., Sablin N.I. *Planirovanie pomekhoustoichivykh regressionnykh eksperimentov* [Planning of noiseproof regression experiments]. *Zavodskaya laboratoriya – Factory laboratory,* 1976, vol. 42, iss. 3, pp. 314-316.

5. Kaas R., Goovaerts M., Dhaene J., Denuit M. *Modern Actuarial Risk Theory.* Boston, Kluwer Academic Publishers, 2001. 309 p. ISBN 0-7923-7636-6 (Russ. ed.: Kaas R., Guverts M., Dene Zh., Denut M. Sovremennaya aktuarnaya teoriya riska. Moscow, Yanus-K Publ., 2008. 376 p.).

6. Laptev V.N. *Nekotorye voprosy sinteza optimal'nykh planov regressionnykh eksperimentov s pomoshch'yu EVM. Diss. kand. tekhn. nauk* [Some questions of synthesis of optimum plans of regression experiments by means of the computer. Cand. eng. sci. diss.]. Novosibirsk, 1974. 129 p.

7. Marshall A.W., Olkin I. *Inequalities: Theory of Majorization and its Applications*. New York, Academic Press, 1979. (Russ. ed.: Marshall A., Olkin I. *Neravenstva: teoriya mazhorizatsii i ee prilozheniya*. M., Mir Publ., 1983. 576 p.).

8. Fedorov V.V. *Teoriya optimal'nogo eksperimenta* [Theory of optimum experiment]. – Moscow, Nauka Publ., 1971. 312 p.

9. Fedorov V.V. Chislennye metody postroeniya optimal'nykh planov dlya regressionnykh eksperimentov [Numerical methods of creation of optimum plans for regression experiments]. *Kibernetika – Cybernetics,* 1975, iss. 1, pp. 124-130.

10. Kiefer J.C. General equivalence theory for optimum designs (approximate theory). The Annals of Statistics, 1974, vol. 2, no. 5, pp. 849-879.

11. Loevner C. Über monotone Matrixfunctionen. Mathematische Zeitschrift. 1934, vol. 38, pp. 177-216.

12. Olkin I., Pratt J. A multivariate Tschbychev's inequality. Annals of Mathematical Statistics, 1958, vol. 29, pp. 226-234.

13. Pukelsheim F. Optimal Design of Experiments. Philadelphia, USA, SIAM, 2006. 453 p. Classics in Applied Mathematics; vol. 50.

14. Whittle P. A multivariate generalization of Tschbychev's inequality. The Quarterly Journal of Mathematics. – 1958, vol. 9, no. 1, pp. 232-240.

ISSN 1814-1196, http://journals.nstu.ru/vestnik
Scientific Bulletin of NSTU
Vol. 56, No .3, 2014, pp. 67–78