

ВЫБОР НАЧАЛЬНЫХ СТРАТЕГИЙ ПОСЛЕДОВАТЕЛЬНОГО УПРАВЛЕНИЯ ЭКСПЕРИМЕНТОМ И ПРОБЛЕМА ИНВАРИАНТНОСТИ

В работе проведено исследование классических оптимальных планов экспериментов с позиции их возможного использования в задачах последовательного экспериментирования в рамках так называемой схемы эффективного управления экспериментом. Анализируются основные отличия задач оптимального планирования и эффективного управления. Выясняются условия инвариантности классических планов относительно линейных преобразований области экспериментирования.

1. ПОСТАНОВКА ЗАДАЧИ

Рассмотрим уравнение наблюдения за объектом (или системой) вида:

$$y(x) = \theta^T f(x) + \varepsilon(x),$$

где $E(y(x)|x) = \theta^T f(x)$ – уравнение регрессии (регрессионная модель, функция отклика), $\varepsilon(x)$ – случайная переменная со свойствами $E(\varepsilon(x)) = 0$ и $E(\varepsilon^2(x)) = \sigma^2, \sigma^2 > 0$, дисперсия σ^2 неизвестна (здесь и далее E – символ математического ожидания),

$f^T(x) = (f_1(x), f_2(x), \dots, f_k(x))$ – базисный вектор модели регрессии (вектор непрерывных на области экспериментирования X функций, $X \subset R^p$), x – вектор размерности p входных контролируемых переменных изучаемого объекта, $y(x)$ – выходная наблюдаемая переменная объекта, θ – вектор неизвестных параметров модели размерности k .

Пусть наблюдения за объектом производятся в точках x_1, x_2, \dots, x_m (в точках так называемого спектра плана экспериментов) в количестве n_1, n_2, \dots, n_m наблюдений в каждой из точек спектра плана соответственно.

Положим
$$N = \sum_{i=1}^m n_i,$$

$$M(\xi) = \sum_{i=1}^m n_i f(x_i) f^T(x_i),$$

$$Y = \sum_{i=1}^m n_i f(x_i) y_i, \quad y_i = \frac{1}{n_i} \left(\sum_{j=1}^{n_i} y_{ij} \right),$$

$$y_{ij} = y_j(x_i) = \sum_{l=1}^k \theta_l f_l(x_i) +$$

$$+ \varepsilon_j(x_i), \quad j = 1, 2, \dots, n_i; i = 1, 2, \dots, m.$$

Тогда в соответствии с обычным методом наименьших квадратов, оценки параметров θ можно найти по формуле: $\hat{\theta} = M^{-1}(\xi)Y$.

Планом эксперимента называется набор, состоящий из точек спектра плана и так называ-

емых весов, в соответствии с которыми распределяется общее число экспериментов N по этим точкам. Вид плана эксперимента следующий:

$$\xi = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ n_1 & n_2 & \dots & n_m \end{pmatrix}, \quad \text{где}$$

$$\xi \in \Xi = \left\{ \xi = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ n_1 & n_2 & \dots & n_m \end{pmatrix} \middle| x_i \in X, \right.$$

$$\left. n_i \geq 0, n_i - \text{целые}, i = 1, 2, \dots, m, \sum_{i=1}^m n_i = N \right\}.$$

Суть задачи построения наилучшего (оптимального) плана экспериментов [1] состоит в нахождении на множестве допустимых планов экспериментов Ξ такого, который максимизировал бы (или минимизировал) некоторый функционал от точностной характеристики оценок параметров модели (или от оценок самой модели). Например, если план эксперимента получается в соответствии с решением задачи максимизации некоторого критерия от матрицы $M(\xi)$ (так называемой информационной матрицы Фишера), т. е. максимизируется некоторый критерий от нее $\Phi(M(\xi))$, то соответствующий план называется Φ -оптимальным. Более точно, Φ -оптимальный план получается как решение следующей оптимизационной задачи (см. [1], [2], [3]):

$$\Phi(M(\xi^*)) = \max_{\xi \in \Xi} \Phi(M(\xi)). \quad (1)$$

Часто на практике вместо задачи (1) решают эквивалентную ей задачу:

$$\Phi(M^{-1}(\xi^*)) = \min_{\xi \in \Xi} \Phi(M^{-1}(\xi)). \quad (2)$$

Если наложено ограничение на общее число экспериментов N (или на общую стоимость экспериментов) и множество Ξ принимает, например, вид

$$\Xi = \left\{ \xi = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ n_1 & n_2 & \dots & n_m \end{pmatrix} \middle| x_i \in X, n_i \geq 0, \right. \\ \left. n_i \geq 0, n_i - \text{целые}, i = 1, 2, \dots, m, \sum_{i=1}^m n_i = N \leq N^* \right\},$$

где N^* – верхняя граница для N , то задачи (1) и (2) можно переписать с учетом того, что достижение желаемого качества в соответствии с критерием $\Phi(M^{-1}(\xi))$ может осуществляться при меньшем, чем N^* , числе экспериментов.

Таким образом, если главным критерием для экспериментатора является $\Phi(M^{-1}(\xi))$, а общее число экспериментов (или общая их стоимость) – это ограничение, то задачу построения планов экспериментов (стратегий экспериментирования) можно записать, например, в виде:

$$\Delta^+ \Phi(M^{-1}(\xi)) \rightarrow \min_{\xi \in \Xi}. \quad (3)$$

Здесь

$$\Delta^+ \Phi(M^{-1}(\xi)) = (\Phi(M^{-1}(\xi)) - \Phi^*)_+ = \begin{cases} \Phi(M^{-1}(\xi)) - \Phi^*, & \text{если } \Phi(M^{-1}(\xi)) > \Phi^*; \\ 0, & \text{иначе.} \end{cases}$$

Здесь Φ^* – желаемое значение функционала $\Phi(M^{-1}(\xi))$. Очевидным образом задача (3) обобщается на случай, когда известны стоимость экспериментов (затраты на экспериментирование) в точках области X .

Как было отмечено в работах [4], [5], более привлекательными и интересными с практической точки зрения методами решения задачи (3) являются последовательные методы эффективного управления экспериментами. В этом случае задача нахождения эффективных стратегий экспериментирования формализуется следующим образом. Если обозначить через $\xi_{(i)}^*$ эффективную стратегию экспериментирования (управления экспериментом) на i -ом шаге, то задача по ее нахождению (синтезу) выглядит следующим образом:

$$\xi_{(i)}^* = \text{Arg min}_{\xi_{(i)}} \Delta^+ \Phi(M^{-1}(\xi_{(i)}^0, y_{(i)}^0)), \quad (4)$$

где $\xi_{(i)} \in \Xi_{(i)}$, $M(\xi_{(i)}, y_{(i)}) = F_{1(i)}^T D_{2(i)}^{-1} F_{1(i)}$;

$\Phi(\bullet)$ – как и выше, выпуклый вниз на множестве матриц $M^{-1}(\xi, y)$ функционал;

$$D_{2(i)} = \text{diag}(\hat{\sigma}_{\bar{y}(i)}^2(x_j));$$

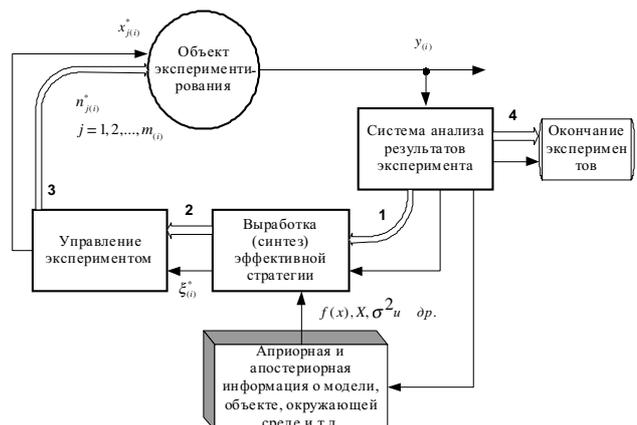
$$\hat{\sigma}_{\bar{y}(i)}^2(x_j) = \frac{1}{n_{j(i)}^0(n_{j(i)}^0 - 1)} \sum_{l=1}^{n_{j(i)}^0} (\bar{y}_j - y_{jl})^2, \quad j = 1, 2, \dots, m_{(i)}^0;$$

$$\Xi_{(i)} = \left\{ \xi \middle| \xi = \begin{cases} x_j \\ n_j \end{cases} \right._{j=1,2,\dots,m_{(i)}} \left. , x_j \in X, j = 1, 2, \dots, m_{(i)}; \right. \\ \left. n_j \geq 0, n_j - \text{целые}, j = 1, 2, \dots, m_{(i)}, \sum_{j=1}^{m_{(i)}} n_j = N_{(i)} \right\},$$

$$y_{(i)}^0 = \begin{pmatrix} y_{(i-1)}^0 \\ y_{(i)} \end{pmatrix}; \xi_{(i)}^0 = \xi_{(i-1)}^* \dot{\cup} \xi_{(i)}, \xi_{(i)}^* = \xi_{(i-1)}^* \dot{\cup} \xi_{(i)}^*, \\ N_{(i)}^0 = N_{(i-1)}^0 + N_{(i)}, N_{(i)} = \sum_{j=1}^{m_{(i)}} n_{j(i)}, \quad i = 1, 2, \dots$$

Здесь i – номер шага экспериментирования; $\xi_{(0)}^* = \xi_{(0)}^*$ – известная начальная стратегия экспериментирования, $\xi_{(i)}^*$ – эффективная стратегия экспериментирования на i -ом шаге; $\xi_{(i)}^0$ – «итоговая» (сводная) стратегия управления экспериментом, «накопленная» (составленная) из стратегий $\xi_{(0)}^*, \xi_{(1)}^*, \dots, \xi_{(i-1)}^*, \xi_{(i)}^*$ и «собранная» в единую стратегию $\xi_{(i)}^0$; $y_{(i)}^0$ – вектор всех значений выходной переменной, полученных за i шагов экспериментирования.

Схема управления экспериментом в этом случае имеет вид (см. рис. 1).



$\square \rightarrow_1$ – команды на поиск очередной эффективной стратегии; $\square \rightarrow_2$ – на систему управления экспериментами в соответствии с $\xi_{(i)}^*$; $\square \rightarrow_3$ – на объект экспериментирования; $\square \rightarrow_4$ – на прекращение экспериментов; \rightarrow – каналы передачи информации (информационные каналы).

Рисунок 1. Система управления экспериментами.

2. СРАВНИТЕЛЬНЫЙ АНАЛИЗ ДВУХ ПОДХОДОВ К УПРАВЛЕНИЮ ЭКСПЕРИМЕНТАМИ

Проведем сравнение классических оптимальных планов экспериментов (которые строятся до проведения экспериментов в соответствии с решением задач (1) или (2)) и эффективных стратегий (синтезируемых последовательно в соответствии с решением задачи (4)) по нескольким характеристикам.

Критерий оптимальности (эффективности)

Критерий оптимальности в случае классического оптимального планирования экспериментов (КОПЭ) является детерминированным, а для эффективного управления экспериментами (ЭУЭ) – стохастическим функционалом. В соответствии с этим и задача построения эффективных стратегий превращается в стохастическую задачу математического программирования. А это означает, что и эффективное управление – стохастическое управление, а задача по синтезу стратегий имеет не единственное решение (как это было в случае КОПЭ), а бесконечное множество решений.

Характеристическое равенство в теореме эквивалентности

Утверждение теоремы эквивалентности, сформулированной и доказанной для КОПЭ (см. [1], [2]), в общем случае для эффективных стратегий экспериментирования не выполняется. Для классического случая максимальное значение функции дисперсии оценки модели ($\hat{d}(x, \xi)$) достигается в точках спектра оптимально плана ξ^* . Попробуем разобраться в этом свойстве оптимального плана с практической точки зрения. После того, как в соответствии с планом ξ^* будут проведены эксперименты, по полученным данным будут оценены параметры $\hat{\theta}$ модели $y(x)$. При этом наибольшие значения дисперсии $\hat{d}(x, \xi)$ оценки $\hat{y}(x) = f^T(x)\hat{\theta}$ (для D - и G -оптимальных планов), т. е. наименьшая точность построенной модели, будут достигаться в точках, в которых были проведены наблюдения над объектом исследования. Некоторое время тому назад нам трудно было объяснить экспериментатору-исследователю, оставаясь на позициях классического подхода к планированию экспериментов, чем объясняется такая особенность оптимальных планов. Нам часто при-

ходилось выслушивать критику со стороны экспериментаторов относительно подобных свойств классических планов. Тогда мы не могли объяснить им этого феномена. Теперь причина ясна. Оказывается, что в общем случае оптимальных планов экспериментов не существует. А для эффективной стратегии управления экспериментом теорема, аналогичная классической теореме эквивалентности, не выполняется.

Влияние вида области экспериментирования на стратегии управления экспериментом

Хорошо известно, что для задач КОПЭ используется подход, связанный с переводом (преобразованием) исходной области экспериментирования (X) в некоторую стандартную область (X_{cm}). В качестве последней обычно служат единичный гиперкуб, единичная гиперсфера, единичный шар, единичный симплекс и другие. Именно для таких стандартных областей построены таблицы оптимальных (для различных классов моделей и критериев оптимальности) классических планов экспериментов. Предполагается, что если план эксперимента инвариантен относительно преобразований области X в область X_{cm} , то можно воспользоваться оптимальным планом, построенным для X_{cm} , затем пересчитать его на область X и затем, в соответствии с пересчитанным планом, произвести эксперименты на объекте. Это очень конструктивная идея, позволяющая унифицировать процедуру выбора (а самое главное – построения) классических оптимальных планов, поскольку одному оптимальному (на области X_{cm}) плану экспериментов соответствовало бесконечное множество задач с фактическими областями экспериментирования X . Можно сказать, что в классическом случае была создана своеобразная система стандартизации (унификации) оптимальных планов, не привязанная к конкретной задаче экспериментатора. Таким образом, можно объяснить появление многих видов кодировок уровней факторов (+1, -1, 0; a, b, c и т. д.) для классических планов.

3. ИНВАРИАНТНОСТЬ СТРАТЕГИЙ ЭКСПЕРИМЕНТИРОВАНИЯ ОТНОСИТЕЛЬНО ПРЕОБРАЗОВАНИЙ ОБЛАСТИ ЭКСПЕРИМЕНТИРОВАНИЯ X

Поскольку в качестве точек спектра эффективных стратегий экспериментирования на на-

чальном этапе могут быть выбраны точки спектра КОПЭ, то представляется логичным проверить КОПЭ на их инвариантность относительно преобразований (например, линейных) области X . Если окажется, что КОПЭ будут неинвариантными относительно этих преобразований, то в этом случае использовать таблицы оптимальных планов и пользоваться разработанными программными комплексами нужно с определенной осторожностью. В случае синтеза эффективных стратегий управления экспериментом актуальность этого вопроса (инвариантности стратегии экспериментирования относительно преобразования области X) нивелируется, т. к. эффективные стратегии строятся (синтезируются) всегда на области X , а преобразование этой области в некоторую стандартную X_{cn} только усложняет задачу и необходимость в нем отпадает. Однако, как было замечено выше, точки спектра КОПЭ могут быть использованы на начальных этапах последовательной процедуры синтеза стратегий экспериментирования, например при формировании стратегий $\xi_{(0)}^*$.

Пример. Пусть для уравнения регрессии $\theta^T f(x), f^T(x) = (1, x), x \in X = [-1; 1]$ при выполнении $E(\varepsilon(x)) = 0, E(\varepsilon^2(x)) = \sigma^2$ построен Φ -оптимальный классический план эксперимента ξ^* . Вопрос заключается в следующем: можно ли использовать этот план в качестве оптимального на отрезке $[-a, a]$? Другими словами, изменится ли оптимальность плана ξ^* при использовании оператора растяжения-сжатия: $P_x : X \rightarrow X_1 = [-a; a]$?

Можно записать, что $\bar{x} = a \cdot x$, где $\bar{x} \in X_1, x \in X$. Тогда уравнение регрессии на X_1 имеет вид:

$$y_1 = \theta_1^T \cdot f_1(\bar{x}), f_1^T(\bar{x}) = (1, a \cdot x) = P_f \cdot \begin{pmatrix} 1 \\ x \end{pmatrix},$$

где $P_f = \begin{pmatrix} 1 & 0 \\ 0 & a \end{pmatrix}$ – матричное представление линейного оператора преобразования $f(x) \rightarrow f_1(\bar{x})$. Здесь θ_1 – вектор параметров

преобразованной регрессии. Пусть $\xi_1^* = \left\{ \begin{matrix} \bar{x}_i \\ n_i \end{matrix} \right\}, i = 1, 2, \dots, m$ – пересчитанный с помощью P_x оптимальный план ξ^* , $\bar{x}_i = a \cdot x_i, i = 1, 2, \dots, m$. Поскольку

$$\hat{\theta} = M^{-1}(\xi^*) \cdot Y, M(\xi^*) = \sum_{i=1}^m n_i f(x_i) f^T(x_i),$$

$$Y = \sum_{i=1}^m n_i f(x_i) \bar{y}_i \text{ и } \hat{\theta}_1 = M_1^{-1}(\xi_1^*) \cdot Y_1,$$

$$M_1(\xi_1^*) = \sum_{i=1}^m n_i f_1(\bar{x}_i) f_1^T(\bar{x}_i), Y_1 = \sum_{i=1}^m n_i f_1(\bar{x}_i) \bar{y}_i,$$

то последовательно получаем:

$$f_1^T(\bar{x}) = [P_f \cdot f(x)]^T = f^T(x) P_f,$$

$$f^T(x) = f_1^T(\bar{x}) P_f^{-1};$$

$$M(\xi^*) = \sum_{i=1}^m n_i [f_1^T(\bar{x}_i) P_f^{-1}]^T [f_1^T(\bar{x}_i) P_f^{-1}] =$$

$$= P_f^{-1} \left(\sum_{i=1}^m n_i f_1(\bar{x}_i) f_1^T(\bar{x}_i) \right) P_f^{-1} =$$

$$= P_f^{-1} M_1(\xi_1^*) P_f^{-1} \neq M_1(\xi_1^*);$$

$$M^{-1}(\xi^*) = P_f M_1^{-1}(\xi_1^*) P_f \neq M_1^{-1}(\xi_1^*).$$

Поскольку матрицы $M^{-1}(\xi^*)$ и $M_1^{-1}(\xi_1^*)$ – это с точностью до множителя σ^2 дисперсионные матрицы оценок параметров $\hat{\theta}$ и $\hat{\theta}_1$, то справедлив следующий вывод.

Вывод 1. Для линейной модели на отрезке дисперсионная матрица оценок параметров и, следовательно, значение критерия эффективности от нее ($\Phi(M^{-1}(\xi^*))$) не являются инвариантными относительно преобразования P_x .

Далее можно записать, что

$$\hat{\theta} = M^{-1}(\xi^*) \cdot Y =$$

$$= P_f M_1^{-1}(\xi_1^*) P_f \cdot \left(\sum_{i=1}^m n_i P_f^{-1} f_1(\bar{x}_i) \bar{y}_i \right) =$$

$$= P_f M_1^{-1}(\xi_1^*) Y_1 = P_f \hat{\theta}_1.$$

Отсюда следует еще один вывод.

Вывод 2. Оценки параметров линейной на отрезке модели не инвариантны относительно преобразования P_x .

Найдем дисперсию оценки модели:

$$\hat{d}_f(x, \xi^*) = \sigma^2 f^T(x) M^{-1}(\xi^*) f(x) =$$

$$= \sigma^2 f_1^T(\bar{x}) P_f^{-1} P_f M_1^{-1}(\xi_1^*) P_f P_f^{-1} f_1(\bar{x}) =$$

$$= \sigma^2 f_1^T(\bar{x}) M_1^{-1}(\xi_1^*) f_1(\bar{x}) = \hat{d}_{f_1}(x, \xi_1^*),$$

где $\hat{d}_{f_1}(x, \xi_1^*)$ – дисперсия оценки модели $\hat{\theta}_1^T \cdot f_1(\bar{x})$.

Естественно, на основании только что полученного соотношения можно сделать еще один вывод.

Вывод 3. Дисперсии оценок линейных на отрезке моделей являются инвариантными от-

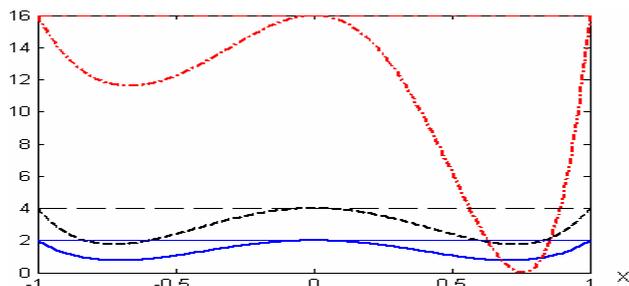


Рисунок 2. Функции $\hat{d}(x, \xi)$ на интервале $[-1; 1]$ для А-оптимального плана (непрерывная кривая), С-оптимального плана (штрих-пунктирная линия), D-оптимального плана (пунктирная линия) при $\sigma = 2$.

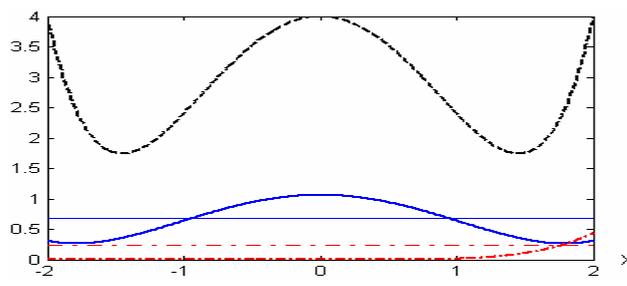


Рисунок 3. Функции $\hat{d}(x, \xi)$ на интервале $[-2; 2]$ для А-оптимального плана (непрерывная кривая), С-оптимального плана (штрих-пунктирная линия), D-оптимального плана (пунктирная линия) при $\sigma = 2$.

$$\xi^* = \begin{Bmatrix} -2 & 0 & 2 \\ 1 & 1 & 1 \end{Bmatrix} \text{ (D-оптимальный),}$$

$$\xi^* = \begin{Bmatrix} -2 & 0 & 2 \\ 2 & 3 & 6 \end{Bmatrix} \text{ (} x_0=4 \text{) (C-оптимальный).}$$

А для интервала $X = [0; 2.0]$ –

$$\xi^* = \begin{Bmatrix} 0 & 1 & 2 \\ 2 & 4 & 2 \end{Bmatrix} \text{ (A), } \xi^* = \begin{Bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{Bmatrix} \text{ (D),}$$

$$\xi^* = \begin{Bmatrix} 0 & 1 & 2 \\ 2 & 3 & 6 \end{Bmatrix} \text{ (} x_0=3 \text{) (C).}$$

Можно заметить (см. рис. 3), что при растяжении области экспериментирования условие инвариантности выполняется только для критерия D-оптимальности, а при смещении (см.

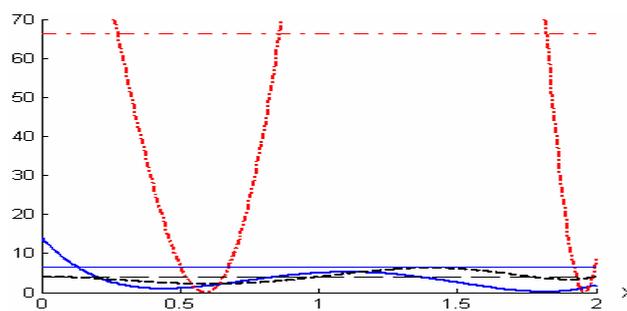


Рисунок 4. Функции $\hat{d}(x, \xi)$ на интервале $[0; 2]$ для А-оптимального плана (непрерывная кривая), С-оптимального плана (штрих-пунктирная линия), D-оптимального плана (пунктирная линия) при $\sigma = 2$.

рис. 4) условие инвариантности не выполнилось ни для одного из критериев.

Инвариантность D-оптимальных планов относительно преобразований области экспериментирования общего вида исследована в работе [2].

Список использованной литературы:

1. Kiefer J. General equivalence theory for optimum designs (approximate theory) // The Annals of Mathematical Statistics. – 1974. – v. 2. – №5. – P. 849-879.
2. Федоров В.В. Теория оптимального эксперимента. – М.: Наука, 1971.
3. Vandemer H. et al. Theorie und Anwendung der optimalen Versuchsplanung I. Handbuch zur Theorie. – Berlin: Akademie-Verlag, 1977.
4. Naumov A. Investigation of Experiment Design Effectiveness and First Steps to Random Experimenting Strategy. I. // Наука и образование в условиях социально-экономической трансформации общества: Материалы V Международной научной конференции, 30-31 мая 2002. – Гродно: Изд-во ГПУ, 2002. – Ч. 2. – С. 217-222.
5. Naumov A. Classical Design of Experiments Approach May Be Non-Effective For Active Identification Problems // Automation, Control, and Information Technology. Proceedings of the IASTED International Conference, June 10– 13, 2002, Novosibirsk. – Anaheim: IASTED, 2002. – P. 262-266.