

ПОБУДОВА АПРОКСИМАЦІЙНИХ ПОЛІНОМІВ ТА ІЗОЛІНІЙ ДИСПЕРСІЇ ПРОГНОЗОВАНИХ ЗНАЧЕНЬ ІЗ ВИКОРИСТАННЯМ СИСТЕМ КОМП'ЮТЕРНОЇ МАТЕМАТИКИ

І.В. ГАДЗАМАН, В.І. БАРАНЕЦЬКИЙ

Розглядаються матриці планування експериментів на основі трьохкомпонентних сполук. Описано метод побудови апроксимаційних поліномів та ізоліній дисперсії прогнозованих значень для експериментальних матриць планування як для випадку, коли експериментальні точки знаходяться у вершинах або на гранях симплекса, так і тоді, коли вони усі розташовуються всередині області, що досліджується. Наведено схему побудови ізоліній дисперсії прогнозованих значень у випадку регресійних поліномів другого, третього та четвертого порядків у декартовій системі координат. Для побудови тернарних графіків у декартовій системі координат використовуються формули переходу Дрейпера-Лоуренса. Цю схему апробовано на експериментальних матрицях *D*-оптимальних та симплекс-граткових планів із другого по четвертий порядок включно. За допомогою систем комп'ютерної математики отримано декілька оптимальних матриць для експериментальних планів другого порядку, та побудовано ізоліній дисперсії прогнозованих значень для цих планів.

ВСТУП

Класична концепція планування статистичних експериментів плідно використовується для розв'язання багатьох прикладних задач. Практика проведення прикладних робіт засвідчила високу ефективність застосування планування експериментів у хімії, фізиці, біології, матеріалознавстві та багатьох інших галузях науки і техніки [1, 2, 3]. Разом із цим ускладнення моделей і їхнє різноманіття ставить задачі оптимізації наукових досліджень, що у багатьох випадках виходять за рамки вже існуючих методів проведення та обробки результатів експериментів. Створення складних моделей пов'язане насамперед із використанням сучасної обчислювальної техніки, що дозволила вивчати великі системи, які залежать від багатьох факторів [2, 4].

Мета роботи — побудова рівнянь регресії для трьохкомпонентних експериментальних планів та ізоліній дисперсії прогнозованих значень для цих планів, у випадку коли всі експериментальні точки знаходяться всередині області досліджень.

ПОСТАНОВКА ЗАДАЧІ

Для побудови математичних моделей реальних процесів і явищ та розв'язку задач по діаграмах «склад-властивість» широко використовуються плани Шеффе, Кіфера, Бокса, Дрейпера-Лоуренса та ін. [1, 5, 6].

Основою побудови математичних моделей для цих планів є апарат регресійного аналізу. Оскільки в більшості задач по вивченню поверхні відклику, залежність між відкликом і незалежними змінними (факторами) невідома, то, на першому етапі, необхідно знайти апроксимуючу функцію, яка може бути представлена у вигляді звичайного поліному степені n від q змінних, що містить C_{q+n}^n коефіцієнтів:

$$Y = b_0 + \sum_{1 \leq i \leq q} b_i x_i + \sum_{1 \leq i \leq j \leq q} b_{ij} x_i x_j + \sum_{1 \leq i \leq j \leq k \leq q} b_{ijk} x_i x_j x_k + \sum_{1 \leq i \leq j \leq k \leq l \leq q} b_{ijkl} x_i x_j x_k x_l + \dots,$$

де Y — функція відклику (вихідний параметр); x_i — фактори (регульовані керуючі параметри); b_0, b_i, b_{ij} тощо — коефіцієнти рівняння регресії; n — число вхідних змінних.

Плануючи експеримент під час вивчення діаграм «склад-властивість» факторний простір являє собою правильний $(q - 1)$ — мірний симплекс, для якого в афінній системі координат виконується співвідношення

$$\sum_{i=1}^q x_i = 1.$$

Для $q = 3$ правильний симплекс являє собою рівносторонній трикутник, де кожна з точок трикутника відповідає одному складу потрійної системи, і, навпаки, кожний склад визначається однією точкою.

При побудові регресійних співвідношень, що представляють спрощену математичну модель явища часто вдається обмежитися розкладанням за поліномами другого порядку.

У планах, запропонованих Дрейпером та Лоуренсом, на відміну від інших, експеримент проводиться тільки з q -компонентними сполуками, що є їх суттєвою перевагою, адже вони передбачають одночасне варіювання всіх факторів. Для зручності побудови тернарних графіків (трьохкомпонентні системи) Дрейпер і Лоуренс вводять нову систему координат. При цьому початок координат співпадає з центром ваги концентраційного трикутника (x_1, x_2, x_3) , одна з вершин лежить на осі z_2 , а дві інші симетричні відносно цієї осі [1, 7].

Для побудови полінома другого порядку стосовно трьохкомпонентних систем Дрейпер і Лоуренс побудували плани, які містять від 8 до 15 експериментальних точок [8].

Оцінку коефіцієнтів поліномів можна провести згідно матричного рівняння $\beta = (X^T X)^{-1} X^T Y$ або отримати, використовуючи властивості насиченості плану [5].

Оскільки точність прогнозування відклику неоднакова в різних точках симплекса, то однією із характеристик «якості» планів можна використовувати дисперсію прогнозованого значення відклику. Її можна визначити за законом накопичення помилок, адже вона, здебільшого, залежить від числа і розподілу паралельних дослідів між окремими точками плану. Збільшуючи кількість паралельних дослідів дисперсію $\sigma^2 \{\hat{y}\}$ можна зробити досить ма-

лою. Проте на практиці число паралельних дослідів завжди обмежене і для полегшення розрахунків часто обмежуються проведенням однакової кількості паралельних дослідів для всіх експериментальних точок.

У такому випадку дисперсію прогнозованого значення відклику шукають за формулою:

$$\sigma^2 \{\hat{y}\} = \sigma^2 \{y\} \frac{\xi}{r},$$

де $\sigma^2 \{y\}$ — дисперсія відтворюваності, r — число паралельних дослідів.

Для симплекс-граткових та D -оптимальних планів формули для розрахунку регресійних коефіцієнтів апроксимуючих поліномів (зведених поліномів Шеффе) отримані з урахуванням того, що вони є лінійними функціями відкликів у вершинах симплекса [1].

Коли ж експеримент проводиться тільки з q -компонентними сполуками, або використовуються ненасичені чи перенасичені плани, оцінку коефіцієнтів відповідних поліномів та значення дисперсії прогнозованого значення провести значно складніше. У таких випадках для розрахунків доцільно використовувати спеціалізовані програми для статистичного опрацювання даних.

У цій роботі для побудови графіків дисперсії прогнозованого значення пропонується така схема: будується матриця планування, далі, на її основі, використовуючи метод найменших квадратів, шукається апроксимаційний поліном та отримуються значення $\sigma^2 \{\hat{y}\}$. Програмна реалізація здійснюється за допомогою використання системи комп'ютерної математики «Mathematica».

ПОБУДОВА МАТЕМАТИЧНОЇ МОДЕЛІ

Покажемо реалізацію запропонованої схеми на прикладі моделі другого порядку. Зважаючи на те, що часто експериментатору важко оперувати складними сполуками із дробовими співвідношеннями компонентів та досить значним збільшенням дисперсії у вузлах симплекса, то обмежимося значеннями $x_1, x_2, x_3 \leq 0,8$.

Точки плану — вершини трикутників, подібних концентраційному, з центром в початку координат і сторонами p та g . У системі z_1 та z_2 координати точок обчислюються за формулами:

$$\left(0, \frac{1}{\sqrt{3}} p\right), \left(\pm \frac{1}{2} p, -\frac{\sqrt{3}}{6} p\right), \left(0, -\frac{1}{\sqrt{3}} g\right), \left(\pm \frac{1}{2} g, \frac{\sqrt{3}}{6} g\right).$$

Формули для переходу від декартової системи координат $z_1 z_2$ до ternaryної $x_1 x_2 x_3$ та навпаки наведені в [5].

Побудуємо апроксимаційний поліном на прикладі моделі другого порядку із параметрами $p = 0,7$ та $g = 0,29$. Значення експериментальних точок наведені в табл. 1.

Таблиця 1. Матриця планування у тернарній системі координат

Номер досліджу	Значення експериментальних точок		
	x_1	x_2	x_3
1	0,1	0,1	0,8
2	0,1	0,8	0,1
3	0,8	0,1	0,1
4	0,43	0,43	0,14
5	0,14	0,43	0,43
6	0,43	0,14	0,43

Використовуючи метод найменших квадратів (функцію *Fit* у середовищі *Mathematica*), побудуємо апроксимаційний поліном:

$$\begin{aligned}
 f(x_1, x_2, x_3) = & 0,131665x_1^2 y_{s_1} + 0,685672x_1 x_2 y_{s_1} + 0,131665x_2^2 y_{s_1} - \\
 & - 1,78186x_1 x_3 y_{s_1} - 1,78186x_2 x_3 y_{s_1} + 1,99314x_3^2 y_{s_1} + \\
 & + 0,131665x_1^2 y_{s_2} - 1,78186x_1 x_2 y_{s_2} + 1,99314x_2^2 y_{s_2} + \\
 & + 0,685672x_1 x_3 y_{s_2} - 1,78186x_2 x_3 y_{s_2} + 0,131665x_3^2 y_{s_2} + \\
 & + 1,99314x_1^2 y_{s_3} - 1,78186x_1 x_2 y_{s_3} + 0,131665x_2^2 y_{s_3} - \\
 & - 1,78186x_1 x_3 y_{s_3} + 0,685672x_2 x_3 y_{s_3} + 0,131665x_3^2 y_{s_3} - \\
 & - 0,767133x_1^2 y_{s_4} + 7,89561x_1 x_2 y_{s_4} - 0,767133x_2^2 y_{s_4} - \\
 & - 1,50878x_1 x_3 y_{s_4} - 1,50878x_2 x_3 y_{s_4} + 0,277799x_3^2 y_{s_4} + \\
 & + 0,277799x_1^2 y_{s_5} - 1,508781x_1 x_2 y_{s_5} - 0,767133x_2^2 y_{s_5} - \\
 & - 1,50878x_1 x_3 y_{s_5} + 7,89561x_2 x_3 y_{s_5} - 0,767133x_3^2 y_{s_5} + \\
 & + 0,767133x_1^2 y_{s_6} - 1,50878x_1 x_2 y_{s_6} + 0,277799x_2^2 y_{s_6} + \\
 & + 7,89561x_1 x_3 y_{s_6} - 1,50878x_2 x_3 y_{s_6} - 0,767133x_3^2 y_{s_6} ,
 \end{aligned}$$

де y_{s_i} — значення відклику в експериментальних точках.

Звівши подібні доданки при y_{s_i} , отримаємо середні квадратичні відхилення, які піднесемо до квадрату. Оскільки математична модель не передбачає паралельних дослідів, то $y_{s_i} = 1$. Поліном для побудови ліній рівних значень ξ набуде вигляду:

$$\begin{aligned}
 \xi = & 5,26142x_1^4 - 18,029x_1^3 x_2 + 75,1226x_1^2 x_2^2 - 18,029x_1 x_2^3 + 5,26142x_2^4 - \\
 & - 18,029x_1^3 x_3 - 30,8238x_1^2 x_2 x_3 - 30,8238x_1 x_2^2 x_3 - 18,029x_2^3 x_3 + 75,1226x_1^2 x_3^2 -
 \end{aligned}$$

$$-30,8238x_1x_2x_3^2 + 75,1226x_2^2x_3^2 - 18,029x_1x_3^3 - 18,029x_2x_3^3 + 5,26142x_3^4.$$

Розташування експериментальних точок та лінії рівних значень ξ для даної моделі наведено на рис. 1.

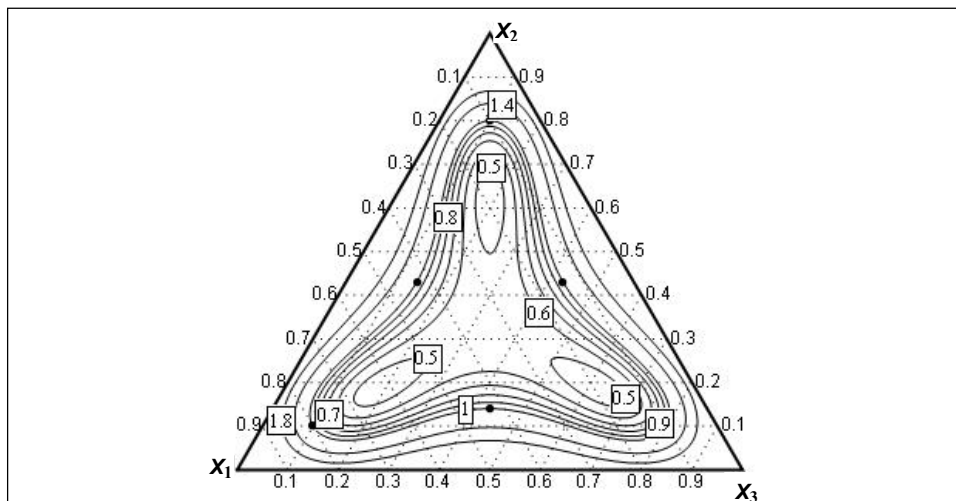


Рис. 1. Ізолінії дисперсії прогнозованого значення ($p = 0,7$; $g = 0,29$)

Таким чином, можна провести оцінку коефіцієнтів регресійних поліномів та побудувати контурний графік ξ для моделей і вищих порядків.

АПРОБАЦІЯ МОДЕЛІ

На рис. 2–6 зображено лінії рівних значень ξ для симплекс-ґраткових та D-оптимальних планів. Зліва наведені графіки, що взяті з літературних джерел [1, 5], справа — змодельовані.

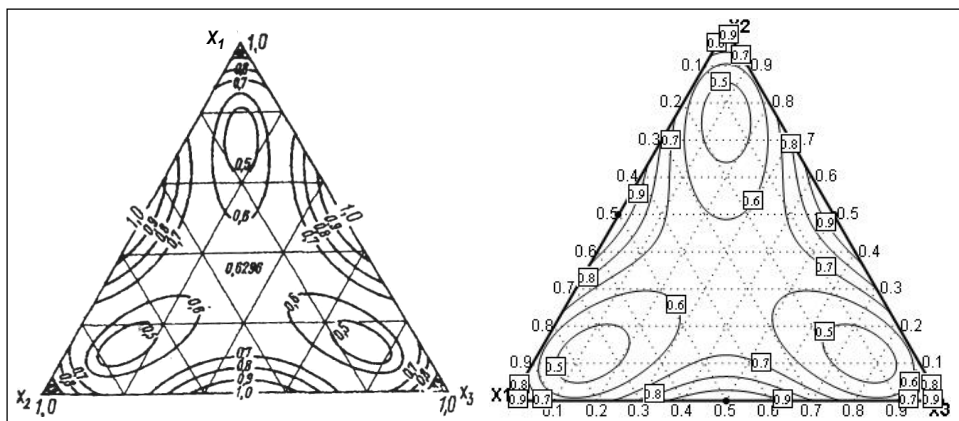


Рис. 2. Ізолінії ξ для симплекс-ґраткового плану 2-го порядку

Як видно з рис. 2–6 характер ізоліній ξ є однаковим, тому запропонована схема може мати місце при обчисленні коефіцієнтів регресійних поліномів та побудові дисперсії прогнозованих значень.

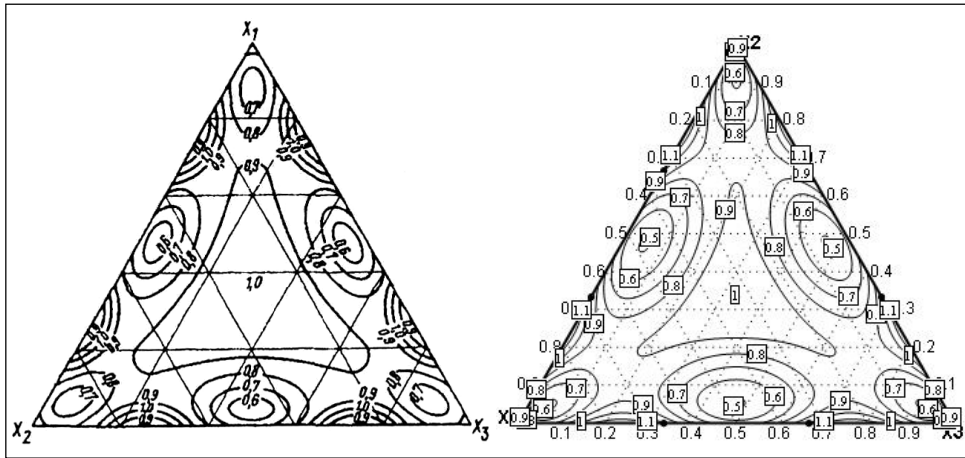


Рис. 3. Ізолінії ξ для симплекс-граткового плану 3-го порядку

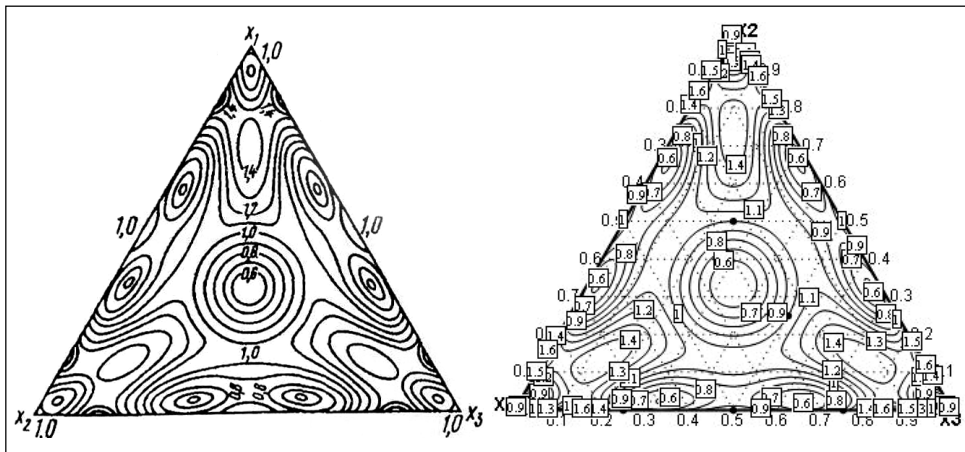


Рис. 4. Ізолінії ξ для симплекс-граткового плану 4-го порядку

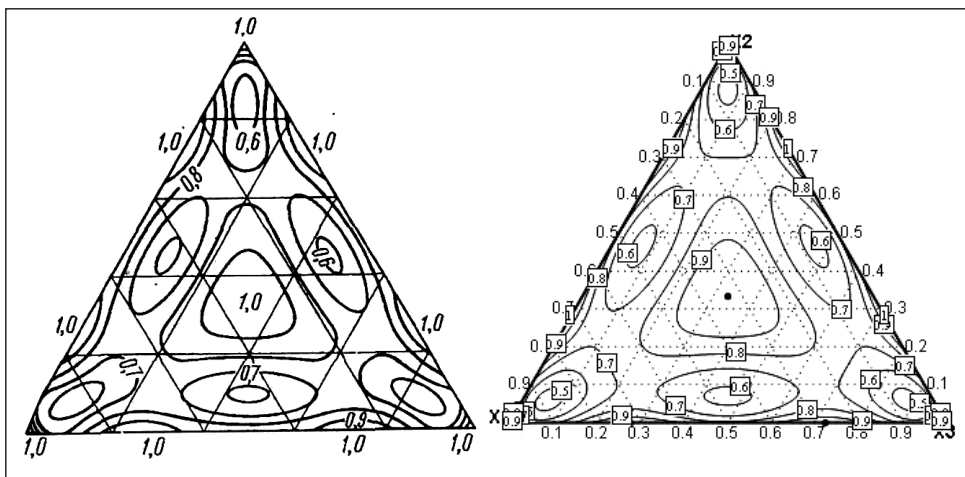


Рис. 5. Ізолінії ξ для D -оптимального плану 3-го порядку

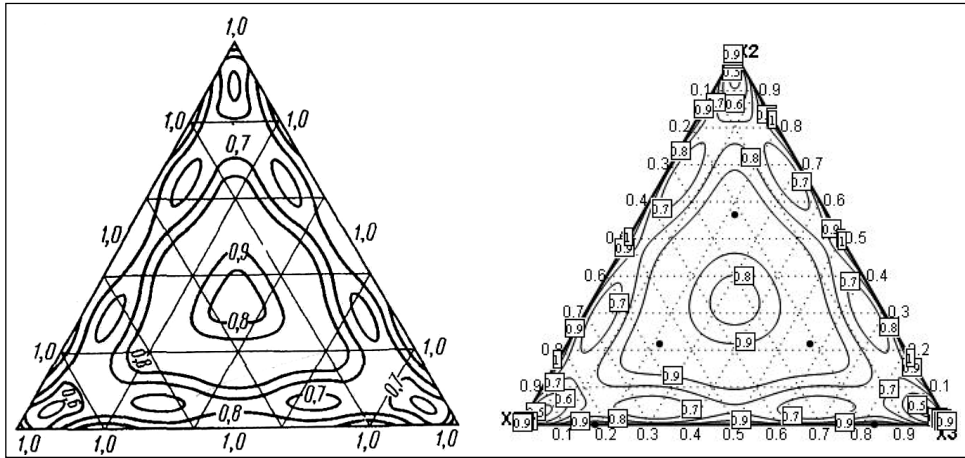


Рис. 6. Ізолінії ξ для D -оптимального плану 4-го порядку

РЕЗУЛЬТАТИ І ВИСНОВКИ

Використовуючи запропоновану схему, було отримано ряд матриць планування 2-го порядку. У табл. 2–3 наведено дві вибрані матриці, а на рис. 7–8 — ізолінії ξ для цих матриць.

Таблиця 2. Значення експериментальних точок ($p = 0,685; g = 0,44$)

n	x_1	x_2	x_3	x_1
1	0,105	0,105	0,79	1
2	0,105	0,79	0,105	1
3	0,79	0,105	0,105	1
4	0,48	0,48	0,04	1
5	0,04	0,48	0,48	1
6	0,48	0,04	0,48	1

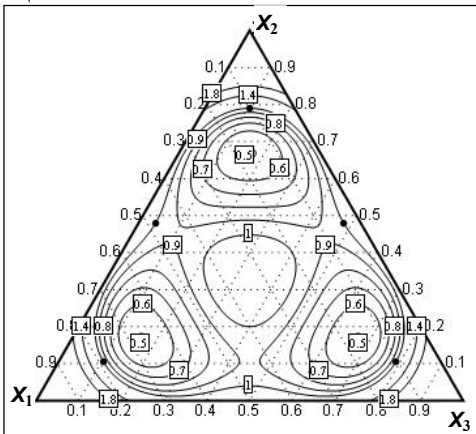


Рис. 7. Ізолінії ξ ($p = 0,685; g = 0,44$)

Таблиця 3. Значення експериментальних точок ($p = 0,7; g = 0,44$)

n	x_1	x_2	x_3	x_1
1	0,1	0,1	0,8	1
2	0,1	0,8	0,1	1
3	0,8	0,1	0,1	1
4	0,48	0,48	0,04	1
5	0,04	0,48	0,48	1
6	0,48	0,04	0,48	1

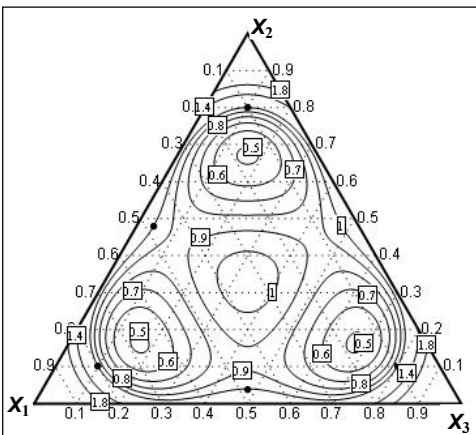


Рис. 8. Ізолінії ξ ($p = 0,7; g = 0,44$)

Вибірка матриць проводилася за критерієм $S_{\xi_{1,4}} - S_{\xi_{0,6}} \rightarrow \max$, де $S_{\xi_{1,4}}$ та $S_{\xi_{0,6}}$ — площі фігур, обмежені лініями та $\xi = 0,6$ відповідно.

Як видно із рисунків 7–8, досить великі ділянки досліджуваної області обмежені значеннями ξ , що близькі до одиниці; а отже, за допомогою наведених у табл. 2–3 отриманих оптимальних матриць планування можна більш достовірно описати взаємозв'язок між складом і властивістю.

Наведена схема оцінки коефіцієнтів регресійних поліномів та побудови ізоліній дисперсії прогнозованого значення для випадків ненасичених чи перенасичених експериментальних планів та у випадку, коли всі експериментальні точки лежать усередині області дослідження дає змогу проводити оцінку оптимальності матриць планування. Розроблена програма для побудови ліній рівних значень дисперсії пройшла апробацію на матрицях планування симплекс-граткових і D-оптимальних планів другого, третього та четвертого порядків. Отримано декілька матриць планування для 2-го порядку з оптимальними значеннями дисперсії.

ЛІТЕРАТУРА

1. Ахназарова С.Л., Кафаров В.В. Оптимизация эксперимента в химической технологии. — М.: Высш. школа, 1985. — 327 с.
2. Лапач С.Н., Губенко А.В., Бабич П.Н. Статистические методы в медико-биологических исследованиях с использованием Excel. — 2-е изд., перераб. и доп. — К: МОРИОН, 2001. — 408 с.
3. Спиридонов А.А. Планирование эксперимента при исследовании технологических процессов. — М.: Машиностроение, 1981. — 184 с.
4. Дьяконов В.П. Mathematica 4.1/4.2/5.0 в математических и научно-технических расчетах. — М.: СОЛОН-Пресс, 2004. — 316 с.
5. Зедгинидзе И.Г. Планирование эксперимента для исследования многокомпонентных систем. — М.: Наука, 1976. — 390 с.
6. Налимов В.В. Применение математической статистики при анализе вещества. — М.: Гос. изд-во лит-ры, 1960. — 430 с.
7. Баранецький В.І., Гадзаман І.В. Використання середовища Mathematica для оптимізації аналітичного запису рівнянь регресії n -го порядку // Актуальні проблеми фізики, математики та інформатики. — 2009. — № 1. — С. 38–40.
8. Draper N.R. and Lawrence W. Mixture designs for three factors // Journal of the Royal Statistical Society. — 1965.— Series B, 27, № 3. — P. 450–465.

Надійшла 07.06.2011