

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ імені ІГОРЯ СІКОРСЬКОГО»

**Навчальний посібник**

# **ОСНОВИ ТЕОРІЇ ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ**

**РОЗДІЛ ДИСЦИПЛІНИ «МЕТОДИКА ТА  
ОРГАНІЗАЦІЯ НАУКОВИХ ДОСЛІДЖЕНЬ»**

*Рекомендовано Методичною радою КПІ ім. Ігоря Сікорського  
як навчальний посібник для здобувачів ступеня бакалавра  
за освітньою програмою «Комп'ютерні системи та мережі»  
спеціальності 123 «Комп'ютерна інженерія»,*

---

Київ  
КПІ ім. Ігоря Сікорського  
2022

УДК 519.23, 004.67

К

*Основи теорії планування експерименту: Розділ дисципліни «Методика та організація наукових досліджень» [Електронний ресурс] : навч. посіб. для студ. освітньої програми «Комп'ютерні системи та мережі» за спеціальністю 123 «Комп'ютерна інженерія» / А.М.Волокита, В.Л.Селіванов О. А; КПІ ім. Ігоря Сікорського. – Електронні текстові дані (1 файл: 1,58 Мбайт). – Київ : КПІ ім. Ігоря Сікорського, 2022. – 41 с.*

*Гриф надано Методичною радою КПІ ім. Ігоря Сікорського  
(протокол № 1 від 02.09.2022р.)  
за поданням Вченої ради факультету інформатики та обчислювальної техніки  
(протокол № 11 від 11.07.2022 р.)*

Електронне мережне навчальне видання

## ОСНОВИ ТЕОРІЇ ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ

### РОЗДІЛ ДИСЦИПЛІНИ «МЕТОДИКА ТА ОРГАНІЗАЦІЯ НАУКОВИХ ДОСЛІДЖЕНЬ»

Укладачі: Волокита Артем Миколайович, канд. техн. наук, доц,  
Селіванов Віктор Левович, канд. техн. наук, доц..

Відповідальний редактор **Кулаков Ю.О., д-р. техн. наук, проф.**

Рецензент: **Дичка І. А., д.т.н., професор, декан факультету прикладної математики, науковий керівник кафедри програмного забезпечення комп'ютерних систем КПІ ім. Ігоря Сікорського**

Навчальний посібник призначено для вивчення основ теорії планування експерименту та використання при проведенні лабораторних занять з дисципліни «Методика та організація наукових досліджень» для підготовки студентів за освітньою програмою «Комп'ютерні системи та мережі» спеціальності 123 «Комп'ютерна інженерія» та для інших спеціальностей і освітніх програм з галузей знань 11 «Математика та статистика» та 12 «Інформаційні технології». Розглядаються теоретичні та прикладні питання теорії планування експерименту, методи наукових досліджень. Список літератури.

© КПІ ім. Ігоря Сікорського, 2022

## ЗМІСТ

<b>ВСТУП.....</b>	<b>4</b>
<b>Загальні відомості та завдання експерименту .....</b>	<b>5</b>
<b>Основні поняття та визначення теорії планування експерименту.....</b>	<b>8</b>
<b>Критерій оптимізації.....</b>	<b>9</b>
<b>Незалежні змінні (фактори).....</b>	<b>10</b>
<b>Забезпечення необхідних та достатніх умов.....</b>	<b>11</b>
<b>для виконання регресійного аналізу .....</b>	<b>11</b>
<b>Статистична перевірка однорідності дисперсії.....</b>	<b>12</b>
<b>Вибір апроксимуючої функції.....</b>	<b>14</b>
<b>Плани факторних експериментів.....</b>	<b>16</b>
<b>Композиційні плани факторних експериментів .....</b>	<b>26</b>
<b>Розрахунок коефіцієнтів рівняння регресії.....</b>	<b>32</b>
<b>Перевірка значущості коефіцієнтів рівняння регресії (нуль-гіпотеза) .....</b>	<b>38</b>
<b>Перевірка адекватності моделі оригіналу .....</b>	<b>39</b>
<b>РЕКОМЕНДОВАНА ЛІТЕРАТУРА.....</b>	<b>41</b>

## ВСТУП

Дисципліна «Методика та організація наукових досліджень» є **вибірковою** дисципліною підготовки фахівців ступеня бакалавр за освітньою програмою «Комп'ютерні системи та мережі» спеціальності 123 «Комп'ютерна інженерія» та для інших спеціальностей і освітніх програм з галузей знань 11 «Математика та статистика» та 12 «Інформаційні технології». Розділ «Основи теорії планування експерименту» призначений для вивчення теоретичних та прикладних питань теорії планування експерименту, методів наукових досліджень, які використовуються в комп'ютерній інженерії, інженерії програмного забезпеченні, інформаційних системах і технологіях. Дисципліна вивчається на четвертому курсі, тому вважається, що студенти засвоїли курси «Вища математика», «Алгоритми та методи обчислень», «Програмування», «Теорія ймовірності та математична статистика».

# ОСНОВИ ТЕОРІЇ ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ

## Загальні відомості та завдання експерименту

Широке застосування різних експериментальних методів призвело до створення теорії експерименту. Ця теорія повинна була надати експериментатору (досліднику) відповіді на наступні питання:

1. Яким чином необхідно зорганізувати експеримент, щоб найкращим чином (витративши мінімум часу, засобів та коштів) виконати необхідні дослідження.

2. Яким чином потрібно обробляти отримані результати експериментів, щоб отримати максимальну кількість інформації стосовно систем (процесів, явищ або об'єктів), що досліджується.

3. Які основні висновки можемо зробити відносно цього об'єкта (системи, явища чи процесу).

Завдання експерименту:

1. Визначення невідомих характеристик та властивостей об'єкта (системи, явища чи процесу).

2. Пошук оптимуму.

3. Перевірка гіпотези.

4. Отримання математичного опису об'єкта (системи, явища чи процесу), який в подальшому використовується для математичного моделювання.

**Моделювання** – це один з методів наукового дослідження, який передбачає використання **моделі** в якості засобу дослідження об'єкта (системи, явища чи процесу), тобто це метод дослідження систем (процесів, явищ або об'єктів) шляхом створення моделей і дослідження цих моделей.

**Модель** – це інший об'єкт, схема, система або навіть алгоритм, який замінює первинний об'єкт (явище чи процес) (**оригінал**) і є джерелом інформації стосовно оригіналу.

**Математичне моделювання** є різновидом моделювання, в основу якого покладено **ідентичність** (співпадіння за формою) математичних описів оригіналу (системи, об'єкта, явища чи процесу що досліджується) та обраної моделі і **однозначність та сталість співвідношень** між змінними оригіналу і моделі. Таким чином, практично в якості оригіналу при математичному моделюванні виступає не сам оригінал, а його математичний опис.

Припустимо, що маємо будь-який об'єкт (систему, явище чи процес), який має  $k$  входів і один вихід (рис.1), де  $\eta$  – вихідна (ендогенна) змінна, яка залежить від  $k$  вхідних (екзогенних) змінних (факторів)  $x_i$  ( $i=\overline{1,k}$ ).

**P.S.:** В подальшому будемо використовувати загальний термін «об'єкт», маючи на увазі, що це може бути як об'єкт, так і система або явище чи процес.



Рис.1

Діапазони, в яких може змінюватися кожна вхідна змінна  $x_i$  ( $i=\overline{1,k}$ ) відомі: від  $x_{i\min}$  до  $x_{i\max}$  ( $i=\overline{1,k}$ ), але внутрішня структура об'єкту невідома і, природньо, невідомий математичний опис, тобто невідомою також є залежність вихідної змінної  $\eta$  від вхідних змінних  $x_i$  ( $i=\overline{1,k}$ )  $\{\eta=f(x_1, x_2, \dots, x_k)\}$ . Оскільки  $f(x_1, x_2, \dots, x_k)$  є невідомою, неможливо її визначити на підставі будь-яких досліджень. Тому необхідно виконати заміну цієї невідомої функції  $f(x_1, x_2, \dots, x_k)$  (метод апроксимації) на іншу (апроксимуючу) функцію  $\eta=\varphi(x_1, x_2, \dots, x_k)$ , яка є відомою.

Позначимо кількість невідомих коефіцієнтів цієї відомої апроксимуючої функції  $\eta=\varphi(x_1, x_2, \dots, x_k)$  як  $K$ . Тоді для знаходження  $K$  невідомих коефіцієнтів, тобто для знаходження функції  $\eta=\varphi(x_1, x_2, \dots, x_k)$  (математичного опису об'єкту), необхідно провести  $K$  експериментів, для кожного з яких обирати значення вхідних змінних  $x_{ij}$  ( $i=\overline{1,k}; j=\overline{1,K}$ ), і визначити  $K$  значень вихідної змінної  $\eta_j$  ( $j=\overline{1,K}$ ). Після цього необхідно знайти значення  $K$  невідомих коефіцієнтів відомої апроксимуючої функції  $\eta=\varphi(x_1, x_2, \dots, x_k)$  з системи  $K$  рівнянь  $\{y_j=\varphi(x_{1j}, x_{2j}, \dots, x_{kj})\}$  ( $j=\overline{1,K}$ ) з  $K$  невідомими. Але при цьому невідомо, яким чином обирати апроксимуючу функцію  $\eta=\varphi(x_1, x_2, \dots, x_k)$ ; як обирати значення  $K$ , щоб забезпечити необхідну точність апроксимації; які значення необхідно обирати для кожної вхідної змінної  $x_i$  ( $i=\overline{1,k}$ ), тобто значення  $x_{ij}$  ( $i=\overline{1,k}; j=\overline{1,r}$ ) [ $x_{i\min} \leq x_{ij} \leq x_{i\max}$  ( $i=\overline{1,k}; j=\overline{1,r}$ ),] та скільки значень  $r$  необхідно обирати для вхідних змінних  $x_i$  ( $i=\overline{1,k}$ ); які  $K$  комбінацій значень  $x_{ij}$  ( $i=\overline{1,k}; j=\overline{1,r}$ ) необхідно обирати з усіх можливих комбінацій, загальна кількість яких складає  $r^k$ . Але це ще не всі труднощі. Уся справа в тому, що окрім вхідних змінних  $x_i$  ( $i=\overline{1,k}$ ), що контролюються, на об'єкт впливають

різноманітні неконтрольовані зовнішні збудження  $z_v$  ( $v=\overline{1,w}$ ) (змінювання температури, коливання напруги тощо), які мають випадковий характер, а вплив цих збуджень на об'єкт призводить до того, що вихідна змінна об'єкту  $\eta$  стає випадковою, і поведінка об'єкту дослідження носить випадковий характер (рис.2).

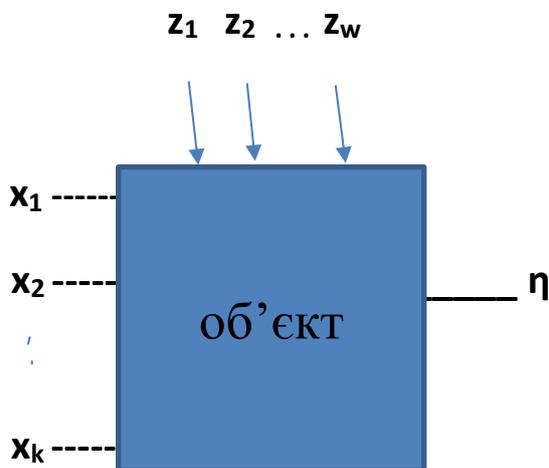


Рис.2

Таким чином, якщо дослідник декілька разів задає одну і ту ж комбінацію рівнів факторів на входи об'єкту, він отримає різні значення вихідної змінної об'єкту  $\eta$ . В даному випадку результати експериментів (отримані значення вихідної змінної) розглядаються як випадкові величини і для аналізу отриманих експериментальних даних необхідно використовувати математичну статистику, яка базується на теорії ймовірності. Під впливом математичної статистики змінювались методи аналізу результатів досліджень, почала змінюватись стратегія проведення експериментів. Слід зазначити, що до певного часу при вивченні складних багатofакторних систем використовувався традиційний спосіб проведення експериментів: для функції  $\eta=f(x_1, x_2, \dots, x_k)$  знаходились залежності  $\eta=f(x_i)$  ( $i=\overline{1,k}$ );  $x_{i\min} \leq x_i \leq x_{i\max}$  для  $x_j = \text{const}$  ( $i=\overline{1,k}$ );  $j \neq i$ ). Отримані при цьому сім'ї характеристик не дозволяли відтворити математичний опис, тобто отримати аналітичну форму запису функції  $\eta=f(x_1, x_2, \dots, x_k)$ .

Саме тому і було розроблено і створено **теорію планування експерименту (ТПЕ)**, яка дозволила вирішити цю проблему, тобто дозволила отримувати математичний опис для будь-яких складних об'єктів. Окрім того, теорія планування експерименту побудована таким чином, що вона забезпечує мінімум затрат (часу, матеріалів, обладнання та коштів) при виконанні досліджень. Також ТПЕ надає чіткі рекомендації стосовно обрання апроксимуючої функції  $\eta=f(x_1, x_2, \dots, x_k)$  та отримання з неї рівняння регресії  $\hat{y}=f(x_1, x_2, \dots, x_k)$ ; визначає, яку

кількість значень обирати для кожної вхідної змінної  $x_i$  ( $i=\overline{1,k}$ ) та як обирати ці значення; як обирати кількість необхідних комбінацій та які саме комбінації обирати. Крім того, ТПЕ пропонує дуже зручну форму для планування досліджень. При цьому стратегія експерименту побудована таким чином, що жодного зайвого експерименту не буде проводитись і уся отримана інформація буде використовуватись для необхідних розрахунків.

Основоположником **теорії планування експерименту** вважається англійський вчений-статистик Р.Фішер (R.A.Fisher), який ще у 1918 році розробив дисперсійний аналіз в якості статистичного методу планування експерименту, та вперше довів доцільність одночасного варіювання усіма факторами. Однак, лише в 1951 році англійські математики Бокс та Уїлсон (G.E.P. Box & K.B. Wilson) розробили правила, згідно з якими і здійснюється одночасне варіювання усіма факторами.

### **Основні поняття та визначення теорії планування експерименту**

**$\eta$**  – параметр процесу, що підлягає оптимізації або параметр (критерій) оптимізації, який повинен мати чітке тлумачення (фізичне, економічне тощо), легко вимірюватись та бути однозначною функцією.

**$x_i$  ( $i=\overline{1,k}$ )** – вхідні змінні (фактори);  **$k$**  – кількість факторів;

**$r$**  – кількість значень кожного фактору (кількість рівнів фактору);

**$N$**  – кількість комбінацій значень факторів, які використовуються; слід пам'ятати, що кількість комбінацій не повинна бути меншою, ніж кількість невідомих коефіцієнтів, тобто  $N \geq K$ , але для мінімізації витрат при дослідженнях необхідно обирати мінімально можливе значення  $N$ .

**$m$**  – кількість повторів кожної комбінації [ $m_i$  ( $i=\overline{1,N}$ ) - кількість повторів  $i$ -ої комбінації];

**$n$**  – загальна кількість експериментів;

**$\bar{x}_i$  ( $i=\overline{1,k}$ )** – кодовані значення факторів.

**Факторний простір** – це декартова система координат, осі якої – кодовані значення факторів  $\bar{x}_k, \bar{x}_{k1} \dots \bar{x}_2, \bar{x}_1$ . Будь-якій комбінації значень факторів відповідає точка факторного простору. При цьому точка з «нульовими» координатами (центр експерименту) відповідає основним рівням факторів  $x_{i0}$  ( $i=\overline{1,k}$ ).

**ПФЕ** – повний факторний експеримент,- це коли використовуються усі можливі комбінації рівнів факторів; при ПФЕ кількість комбінацій є  $N_n=r^k$ .

**ДФЕ** – дробовий факторний експеримент,- це коли використовуються частка ПФЕ, яка кратна степеню двійки, тобто  $N_d=2^{-1}N_n$  (напіврепліка),  $N_d=2^{-2}N_n$  (чвертьрепліка) тощо.

**РКП** – ротатабельний композиційний план (ротатабельний центрально-композиційний план другого порядку),- це коли використовується частка ПФЕ для  $r=5$ . РКП отримують додаванням до ПФЕ ( $N_n=2^k$ ) спеціальних «зоряних» точок.

**ЦОКП** – центральний ортогональний композиційний план (ортагональний центрально-композиційний план другого порядку),- це коли використовується частка ПФЕ для  $r=5$ . ЦОКП отримують додаванням до ПФЕ ( $N_n=2^k$ ) спеціальних «зоряних» точок та додаванням певної кількості експериментів в центрі плану («нульових» точок).

$f(x_1, x_2 \dots x_k)$  – **функція (поверхня) відгуку** або **реакція відгуку** або **цільова функція**. Функція відгуку  $f(x_1)$  для однієї змінної вироджується в лінію  $\eta(x_1)$ ; для двох змінних поверхня відгуку  $\eta(x_1, x_2)$  – поверхня в тривимірному просторі; для  $k$  змінних - поверхня відгуку  $\eta(x_1, x_2 \dots x_k)$  – поверхня в  $(k+1)$ -одновимірному просторі. Для двох змінних цільова функція (поверхня відгуку)  $\eta(x_1, x_2)$  графічно може бути зображена на площині у вигляді ліній однакового рівня  $x_2(x_1)$  для  $\eta = C_1 = \text{const}$ ,  $\eta = C_2 = \text{const}$ ,  $\eta = C_3 = \text{const} \dots$  Таке зображення цільової функції дуже зручно для ілюстрування особливостей поверхні відгуку і процесів пошуку екстремуму.

$\varphi(x_1, x_2 \dots x_k)$  – **апроксимуюча функція, яка замінює функція (поверхня) відгуку, і розглядається як математичний опис об'єкту;  $K$  – кількість коефіцієнтів цієї функції.** Апроксимуюча функція  $\varphi(x_1, x_2 \dots x_k)$  – скінченне число членів ряду розкладання функції  $f(x_1, x_2 \dots x_k)$  в ряд.

### **Критерій оптимізації**

Обираючи критерій (параметр) оптимізації  $\eta$  необхідно брати до уваги багато міркувань. По перше, критерій оптимізації бажано мати таким, щоб він однозначно та з достатньою повнотою характеризував об'єкт дослідження. По друге, критерій оптимізації повинен бути єдиним, мати чітке фізичне значення та оцінюватися кількісно з максимальною ефективністю. Якщо є декілька критеріїв оптимізації, необхідно розглянути можливість зменшення їх кількості. Якщо не

вдається зменшити кількість критеріїв до одного,-- необхідно замінити задачу з декількома критеріями послідовними задачами з одним критерієм.

Критерій оптимізації в технічних системах може бути:

1. Технічним.
2. Економічним.
3. Статистичним.
4. Психологічним.
5. Техніко-економічним.
6. Тактико-технічним тощо.

### Незалежні змінні (фактори)

Кількість обраних факторів  $k$  обумовлює вимірність відповідного факторного простору. Якщо  $k=2$ ,-- маємо двовимірний факторний простір (осі декартової системи координат  $\bar{x}_2$  та  $\bar{x}_1$ ); для  $k=3$  -- тривимірний факторний простір (осі декартової системи координат  $\bar{x}_3$ ,  $\bar{x}_2$  та  $\bar{x}_1$ ) тощо. Бажано, щоб фактори мали кількісну оцінку, хоча планування експерименту інколи можливо навіть тоді, коли деякі з факторів мають якісну оцінку. Фактори  $x$  повинні допускати можливість зміни значень та можливість вимірювання цих значень.

Вимоги до факторів:

1. Керованість. Це означає, що дослідник може встановити потрібний рівень будь-якого фактору.
2. Незалежність. Це означає, що будь-який рівень фактору (будь-яке значення фактору) можемо встановити незалежно від рівнів інших факторів.
3. Сумісність. Це означає, що можемо встановити будь-яку комбінацію рівнів факторів і ця комбінація буде безпечною.
4. Однозначність. Це означає, що будь-який фактор безпосередньо впливає на об'єкт дослідження.
5. Точність вимірювання. Це означає, що дослідник має можливість здійснити вимірювання будь-якого фактору з необхідною точністю.

Для кожного фактору необхідно з'ясувати межі, в яких змінюються значення факторів, тобто значення  $x_{i\min}$  ( $i=\overline{1,k}$ ) та  $x_{i\max}$  ( $i=\overline{1,k}$ ). По цим значенням розрахо-

вуються для кожного фактору наступні значення: значення основних рівнів факторів  $x_{i0}=0.5(x_{imax}+x_{imin})$  ( $i=\overline{1,k}$ ), шага варіювання факторів  $x_{i\Delta}=0.5(x_{imax}-x_{imin})$  ( $i=\overline{1,k}$ ) та діапазони варіювання  $\Delta x_i=x_{imax}-x_{imin}$  ( $i=\overline{1,k}$ ).

При плануванні дослідження дуже зручно використовувати не абсолютні значення рівнів факторів  $x_i$  ( $i=\overline{1,k}$ ) (кожен з факторів має свою розмірність), а так звані кодовані значення рівнів факторів  $\bar{x}_i$  ( $i=\overline{1,k}$ ), які не мають розмірності. Кодування здійснюється за формулою:  $\bar{x}_i=(x_i-x_{i0})/x_{i\Delta}$ . При цьому абсолютному значенню будь-якого фактору  $x_{imax}$  ( $i=\overline{1,k}$ ) завжди відповідає кодоване значення  $\bar{x}_{imax}=+1$  ( $i=\overline{1,k}$ ); абсолютному значенню будь-якого фактору  $x_{imin}$  ( $i=\overline{1,k}$ ) завжди відповідає кодоване значення  $\bar{x}_{imin}=-1$  ( $i=\overline{1,k}$ ); абсолютному значенню основного рівня будь-якого фактору  $x_{i0}$  ( $i=\overline{1,k}$ ) відповідає кодоване значення  $\bar{x}_{i0}=0$  ( $i=\overline{1,k}$ ), а абсолютному значенню шагу варіювання будь-якого фактору  $x_{i\Delta}$  ( $i=\overline{1,k}$ ) відповідає кодоване значення  $\bar{x}_{i\Delta}=1$  ( $i=\overline{1,k}$ ). Для розрахунку абсолютного значення зоряного плеча  $\Delta x_{ii}$  ( $i=\overline{1,k}$ ) для будь-якого фактору слід використовувати формулу  $\Delta x_{ii}=l(x_{imax}-x_{i0})$ , де  $l$  – значення «зоряного» плеча, значення якого залежить від того композиційного плану, який будемо використовувати. Для розрахунку абсолютного значення фактору, яке потрібно задавати на відповідний вхід об'єкту дослідження, якщо кодове значення дорівнює  $+1$ , розраховується за формулою  $+x_{ii}=x_{i0}+\Delta x_{ii}$ , а для кодового значення  $-1$ , -- за формулою  $-x_{ii}=x_{i0}-\Delta x_{ii}$ .

### **Забезпечення необхідних та достатніх умов для виконання регресійного аналізу**

Згідно з вимогами регресійного аналізу правильна обробка та використання результатів експериментальних досліджень можливі лише тоді, коли:

1. Результати вимірювань функції відгуку для  $N$  комбінацій являють собою реалізацію нормально розподіленої випадкової величини.

2. Дисперсії реалізацій (вимірювань функції відгуку)  $\sigma^2_j$  ( $j=\overline{1,N}$ ) для нормально розподіленої випадкової величини однакові для усіх комбінації, тобто не залежать від абсолютного значення функції відгуку для кожної комбінації. Ця властивість нормально розподіленої випадкової величини має назву «однорідність дисперсії».

3. Фактори  $x_i (i=\overline{1,k})$  – незалежні та вимірюються з точністю, яка суттєво вища, ніж точність вимірювання значень вихідної змінної  $y_i (i=\overline{1,n})$ .

Слід зазначити, що третя умова забезпечується завжди, а якщо забезпечується друга умова, тоді автоматично виконується і перша умова. Отже потрібно перевіряти другу умову з допомогою спеціальних статистичних критеріїв.

Для кожної комбінації рівнів факторів необхідно обирати таку кількість повторів  $\{m$  або  $m_i (i=\overline{1,N})\}$ , яка забезпечить нормальний закон розподілу випадкової величини. Для підтвердження нормального закону розподілу випадкової величини необхідно здійснювати статистичну перевірку однорідності дисперсії за обраним критерієм. Теорія планування експерименту пропонує послідовно збільшувати кількість повторів, кожного разу перевіряючи однорідність дисперсії, яка є дуже простою, замість обрання одразу великої кількості повторів, що дозволить суттєво зменшити загальну кількість дослідів  $n \{n=Nm$  або  $n=\sum_1^N m_i\}$ . Можемо обрати спочатку невелику кількість повторів, провести дослідження і здійснити перевірку однорідності дисперсії. Якщо перевірка не підтвердить однорідність дисперсії (з обраною ймовірністю),- тоді необхідно збільшити кількість повторів, провести додаткові дослідження і знову здійснити перевірку ... І так до тих пір, доки не отримаємо підтвердження однорідності дисперсії.

### Статистична перевірка однорідності дисперсії

Перш, ніж розраховувати значення коефіцієнтів рівняння регресії по результатах досліджень, тобто по значенням  $y_i (i=\overline{1,n})$ , необхідно переконатися, що дисперсії реалізацій (вимірювань функції відгуку)  $\sigma^2_i (i=\overline{1,N})$  однакові для усіх комбінацій  $\{\sigma^2_i=\sigma^2 (i=\overline{1,N})\}$ , тобто переконатися в однорідності дисперсії.

Для цього необхідно спочатку знайти середньоарифметичне значення дослідів  $\bar{y}_j (j=\overline{1,m})$  (математичне сподівання  $m_{y_j}$ ) в кожній точці факторного простору:  $\bar{y}_j=(1/m)\sum_1^m y_{js} (j=\overline{1,N})$ .

Оскільки теоретичні значення дисперсії  $\sigma^2_j (j=\overline{1,N})$  невідомі, то перевірка однорідності дисперсії виконується на основі аналізу статистичних оцінок дисперсії  $S^2_j (i=\overline{1,N})$  для усіх точок факторного простору.

Статистичні оцінки дисперсії  $S^2_j (j=\overline{1,N})$  для кожної точки факторного простору розраховують за формулою:  $S^2_j=\{1/(m-1)\}\{\sum_1^m (y_{js}-\bar{y}_j)^2\} (j=\overline{1,N})$ .

Отже, перевірка однорідності дисперсії – це перевірка гіпотези стосовно належності  $N$  значень статистичних оцінок дисперсії  $S^2_j$  ( $i=\overline{1,N}$ ) одній генеральній сукупності. Якщо зазначена гіпотеза підтверджується для обраного значення ймовірності  $p$  (рівня значущості  $q=1-p$ ),-- це означає, що обрана кількість повторів  $m$  забезпечує нормальний закон розподілу випадкової величини з ймовірністю  $p$ . Якщо зазначена гіпотеза не підтверджується для обраного значення ймовірності  $p$  (рівня значущості  $q=1-p$ ),-- це означає, що обрана кількість повторів  $m$  не забезпечує нормальний закон розподілу випадкової величини з ймовірністю  $p$  і необхідно збільшити кількість повторів, провести додаткові дослідження, після чого знову перевірити гіпотезу стосовно однорідності дисперсії і так продовжувати до того часу, поки гіпотеза не буде підтверджена.

Якщо  $N \geq 3$  ( $k \geq 2$ ),-- тоді для перевірки однорідності дисперсії використовується критерій Кохрена:

1.Серед знайдених статистичних оцінок дисперсії  $S^2_j$  ( $j=\overline{1,N}$ ) знаходять оцінку з максимальним значенням  $S^2_{\max} = \text{Max}\{S^2_j$  ( $j=\overline{1,N}$ )}.

2.Розраховують значення критерію Кохрена  $G = S^2_{\max} / \sum_1^N S^2_j$

3.Визначають числа ступенів свободи  $f_1$  та  $f_2$ :  $f_1 = m - 1$ ;  $f_2 = N$ .

4.Обирають рівень значущості  $q$ .

5.По спеціальним таблицям Кохрена знаходять критичне (табличне) значення критерія Кохрена  $G_{кр}$ , яке відповідає значенням  $q$ ,  $f_1$  та  $f_2$ .

6.Порівнюють значення  $G$  та  $G_{кр}$ .

Якщо  $G \leq G_{кр}$ ,-- тоді вважається, що гіпотеза стосовно однорідності дисперсії підтверджується з ймовірністю  $p$  ( $p=1-q$ ) і ми можемо виконувати розрахунок коефіцієнтів рівняння регресії. Якщо  $G \geq G_{кр}$ ,-- тоді вважається, що гіпотеза стосовно однорідності дисперсії не підтверджується з ймовірністю  $p$  ( $p=1-q$ ), -- тоді необхідно збільшити кількість повторів, провести додаткові дослідження і знову здійснити перевірку однорідності дисперсії. Підтвердження гіпотези стосовно однорідності дисперсії дозволяє отримати більш точну статистичну оцінку дисперсії функції відгуку  $S^2 = (1/N) \sum_1^N S^2_j$ .

Якщо  $N=2$  ( $k=1$ ),-- тоді може використовуватися критерій Фішера або критерій Романовського.

## Вибір апроксимуючої функції

В теорії планування експерименту найчастіше використовується ряд Тейлора:

$$\eta = \varphi(x_1, x_2, \dots, x_k) = \beta_0 + \sum_1^k \beta_i x_i + \sum_1^k \beta_{ij} x_i x_j + \sum_1^k \beta_{ijk} x_i x_j x_k + \dots + \sum_1^k \beta_{ii} x_i^2 + \dots,$$

де  $\beta_0, \beta_i (i=\overline{1, k}), \beta_{ij} (i=\overline{1, k}; j=\overline{1, k}; i < j), \beta_{ijk} (i=\overline{1, k}; j=\overline{1, k}; k=\overline{1, k}; i < j < k), \dots, \beta_{ii} (i=\overline{1, k}) \dots$  - теоретичні коефіцієнти ряду Тейлора, значення яких обчислюються через значення частинних похідних функції. Але оскільки функція  $f(x_1, x_2, \dots, x_k)$  невідома, тому визначити значення теоретичних коефіцієнтів ряду Тейлора неможливо і тому зазначений теоретичний ряд  $\eta = \varphi(x_1, x_2, \dots, x_k)$  замінюється рівнянням регресії  $\check{y} = \varphi(x_1, x_2, \dots, x_k)$ :

$$\check{y} = \varphi(x_1, x_2, \dots, x_k) = b_0 + \sum_1^k b_i x_i + \sum_1^k b_{ij} x_i x_j + \sum_1^k b_{ijk} x_i x_j x_k + \dots + \sum_1^k b_{ii} x_i^2 + \dots,$$

де  $b_0, b_i (i=\overline{1, k}), b_{ij} (i=\overline{1, k}; j=\overline{1, k}; i < j), b_{ijk} (i=\overline{1, k}; j=\overline{1, k}; k=\overline{1, k}; i < j < k), \dots, b_{ii} (i=\overline{1, k}) \dots$  -- статистичні оцінки невідомих теоретичних коефіцієнтів ряду Тейлора, значення яких розраховуються по результатам експериментів, а  $\check{y}_i (i=\overline{1, N})$  – розраховані значення параметру оптимізації.

**P.S.:** Слід зазначити, що при деяких наукових дослідженнях можуть використовуватися і інші ряди.

Для мінімізації витрат спочатку обирається рівняння регресії, в якому присутні лише лінійні доданки, тобто  $\check{y} = \varphi(x_1, x_2, \dots, x_k) = b_0 + \sum_1^k b_i x_i$  (лінійна регресія); потім, після проведення необхідних для цього варіанту досліджень і розрахунку коефіцієнтів рівняння регресії перевіряється адекватність моделі (рівняння регресії) оригіналу (усім отриманим експериментальним значенням) з визначеною ймовірністю. Якщо модель адекватна оригіналу з визначеною ймовірністю, - це означає, що таке рівняння регресії є статистичним математичним описом об'єкту. Якщо ж модель не адекватна оригіналу з визначеною ймовірністю, - це означає, що необхідно змінити рівняння регресії, збільшивши кількість членів ряду, врахувавши усі доданки з добутками факторів:

$$\check{y} = \varphi(x_1, x_2, \dots, x_k) = b_0 + \sum_1^k b_i x_i + \sum_1^k b_{ij} x_i x_j + \sum_1^k b_{ijk} x_i x_j x_k + \dots$$

Такий варіант рівняння регресії має назву: «лінійна регресія з врахуванням взаємодії факторів».

Після проведення необхідних для цього варіанту додаткових досліджень і розрахунку коефіцієнтів рівняння регресії знову перевіряється адекватність моделі оригіналу з визначеною ймовірністю. Якщо модель адекватна оригіналу з визначеною ймовірністю, - це означає, що рівняння регресії є статистичним математич-

ним описом об'єкту. Якщо ж модель не адекватна оригіналу з визначеною ймовірністю,- це означає, що необхідно змінити рівняння регресії, збільшивши кількість членів ряду, врахувавши ще квадратичні доданки (**квадратична регресія**):

$$\hat{y} = \varphi(x_1, x_2, \dots, x_k) = b_0 + \sum_1^k b_i x_i + \sum_1^k b_{ij} x_i x_j + \sum_1^k b_{ijk} x_i x_j x_k + \dots + \sum_1^k b_{ii} x_i^2$$

Після проведення необхідних для цього варіанту додаткових досліджень і розрахунку коефіцієнтів рівняння регресії знову перевіряється адекватність моделі оригіналу (усім отриманим експериментальним значенням) з визначеною ймовірністю. Якщо модель адекватна оригіналу з визначеною ймовірністю,- це означає, що рівняння регресії є статистичним математичним описом об'єкту. Якщо ж модель не адекватна оригіналу з визначеною ймовірністю,- це означає, що необхідно змінити рівняння регресії, збільшивши кількість членів ряду. Збільшення членів ряду здійснюється, аж поки модель не стане адекватною оригіналу.

Приклади лінійної регресії та кількість невідомих коефіцієнтів **K** в залежності від кількості факторів **k**:

якщо **k=1**, тоді  $\hat{y} = b_0 + b_1 x_1$  або  $\hat{y} = b_0 + b_1 x$  і **K=2** ( $b_0$  та  $b_1$ );

якщо **k=2**, тоді  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2$  і **K=3** ( $b_0, b_1$  та  $b_2$ );

якщо **k=3**, тоді  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3$  і **K=4** ( $b_0, b_1, b_2$  та  $b_3$ );

якщо **k=4**, тоді  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4$  і **K=5** ( $b_0, b_1, b_2, b_3$  та  $b_4$ );

якщо **k=5**, тоді  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_5 x_5$  і **K=6** ( $b_0, b_1, b_2, b_3, b_4$  та  $b_5$ ) тощо.

Приклади лінійної регресії з врахуванням взаємодії факторів та кількість невідомих коефіцієнтів **K** в залежності від кількості факторів **k**:

для **k=1** маємо  $\hat{y} = b_0 + b_1 x_1$  або  $\hat{y} = b_0 + b_1 x$  і **K=2**;

для **k=2** маємо  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2$  і **K=4**;

для **k=3** маємо  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{123} x_1 x_2 x_3$  і **K=8**;

для **k=4** буде  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{14} x_1 x_4 + b_{23} x_2 x_3 + b_{24} x_2 x_4 + b_{34} x_3 x_4 + b_{123} x_1 x_2 x_3 + b_{124} x_1 x_2 x_4 + b_{134} x_1 x_3 x_4 + b_{234} x_2 x_3 x_4 + b_{1234} x_1 x_2 x_3 x_4$  і **K=16** тощо.

Приклади квадратичної регресії та кількість невідомих коефіцієнтів **K** в залежності від кількості факторів **k**:

якщо  $k=1$ , тоді  $\ddot{y}=b_0+b_1x_1+b_{11}x_1^2$  або  $\ddot{y}=b_0+b_1x_1+b_2x_2$  і  $K=3$ ;

якщо  $k=2$ , тоді  $\ddot{y}=b_0+b_1x_1+b_2x_2+b_{12}x_1x_2+b_{11}x_1^2+b_{22}x_2^2$  і  $K=6$ ;

якщо  $k=3$ , тоді  $\ddot{y}=b_0+b_1x_1+b_2x_2+b_3x_3+b_{12}x_1x_2+b_{13}x_1x_3+b_{23}x_2x_3+b_{123}x_1x_2x_3+b_{11}x_1^2+b_{22}x_2^2+b_{33}x_3^2$  і  $K=11$ ;

якщо  $k=4$ , тоді  $\ddot{y}=b_0+b_1x_1+b_2x_2+b_3x_3+b_4x_4+b_{12}x_1x_2+b_{13}x_1x_3+b_{14}x_1x_4+b_{23}x_2x_3+b_{24}x_2x_4+b_{34}x_3x_4+b_{123}x_1x_2x_3+b_{124}x_1x_2x_4+b_{134}x_1x_3x_4+b_{234}x_2x_3x_4+b_{1234}x_1x_2x_3x_4+b_{11}x_1^2+b_{22}x_2^2+b_{33}x_3^2+b_{44}x_4^2$  і  $K=20$  тощо.

### Плани факторних експериментів

Багатофакторний експеримент пов'язаний з одночасним варіюванням усіх факторів та перевіркою вірогідності результатів статистичними методами. Теоретично фактори можуть мати будь-яку кількість рівнів  $r$ , але в більшості досліджень достатньо використовувати два або п'ять рівнів. Слід також зазначити, що проміжні два варіанти ( $r=3$  та  $r=4$ ) практично не застосовуються. Проведення дослідів з багаторівневими факторами ( $r>5$ ) викликають значні складнощі і тому такі досліді знаходять обмежене застосування на практиці інженерних експериментів. Кількість необхідних рівнів  $r$  залежить від кількості невідомих коефіцієнтів  $K$  обраного варіанту рівняння регресії з урахуванням необхідності мінімізації кількості комбінацій  $N$ , що використовуються. При цьому слід пам'ятати, що кількість комбінацій  $N$  не може бути меншою, ніж кількість невідомих коефіцієнтів  $K$ , тобто  $N \geq K$ .

Плани факторних експериментів дуже зручно надавати у вигляді матриці планування **МП**, в яку заносяться кодовані значення рівнів факторів та значення отриманих експериментальних значень вихідної змінної  $y_{ij}(i=\overline{1,N};j=\overline{1,m})$ . **МП** повинна визначати розташування дослідних точок у факторному просторі. Кожен рядок відповідає умовам досвіду (комбінації кодованих значень), а стовпець – кодованим значенням незалежної змінної (або кодованим значенням добутоків факторів чи кодованим значенням квадратів факторів тощо). Окрім того, в кожен  $i$ -ий ( $i=\overline{1,N}$ ) рядок вносяться отримані в результаті проведення дослідження експериментальні значення вихідної змінної об'єкту  $y_{ij}(i=\overline{1,N};j=\overline{1,m})$ : в перший рядок – значення  $y_{11}, y_{12} \dots y_{1m}$ ; в другий рядок – значення  $y_{21}, y_{22} \dots y_{2m}$ ; ... в  $N$ -ий рядок – значення  $y_{N1}, y_{N2} \dots y_{Nm}$ .

Повний факторний експеримент **ПФЕ**,-- це такий факторний експеримент, коли використовуються усі можливі комбінації рівнів факторів. Для ПФЕ кількість

комбінацій  $N_n=r^k$ : якщо  $r=2$ ,-- тоді  $N_n=2^k$ ; якщо  $r=3$ ,-- тоді  $N_n=3^k$ ; якщо  $r=5$ ,-- тоді  $N_n=5^k$  тощо.

Матриця планування повного факторного експерименту **МП ПФЕ** для одного фактору ( $k=1$ ) використовує два рівня факторів ( $r=2$ ), кодовані значення яких  $-1$  та  $+1$ ; кількість комбінацій – дві ( $N_n=2^1=2$ ) (див. табл.1).

Таблиця 1 МП ПФЕ k=1

N комбінації	Фактори		Експериментальні значення вихідної змінної об'єкту	
	$\bar{x}_1$		$y_{ij}(i=\overline{1,2}; j=\overline{1,m})$	
1	-1		$y_{11}=$ ; $y_{12}=$ ; .....	$y_{1m}=$ ;
2	+1		$y_{21}=$ ; $y_{22}=$ ; .....	$y_{2m}=$ ;

Ця матриця використовується для лінійного рівняння регресії з одним фактором  $\hat{y}=b_0+b_1x_1$ , в якому два невідомих коефіцієнта  $b_0$  та  $b_1$  ( $K=2$ ) і  $N=K$ .

Матриця планування повного факторного експерименту **МП ПФЕ** для двох факторів ( $k=2$ ) використовує також два рівня факторів ( $r=2$ ), кодовані значення яких  $+1$  та  $-1$ ; кількість комбінацій – чотири ( $N_n=2^2=4$ ) (див. табл.2).

Таблиця 2 МП ПФЕ k=2

N комбінації	Фактори		Експериментальні значення вихідної змінної об'єкту	
	$\bar{x}_1$	$\bar{x}_2$	$y_{ij}(i=\overline{1,4}; j=\overline{1,m})$	
1	-1	-1	$y_{11}=$ ; $y_{12}=$ ; .....	$y_{1m}=$ ;
2	-1	+1	$y_{21}=$ ; $y_{22}=$ ; .....	$y_{2m}=$ ;
3	+1	-1	$y_{31}=$ ; $y_{32}=$ ; .....	$y_{3m}=$ ;
4	+1	+1	$y_{41}=$ ; $y_{42}=$ ; .....	$y_{4m}=$ ;

Ця матриця використовується для лінійного рівняння регресії з двома факторами  $\hat{y}=b_0+b_1x_1+b_2x_2$ , в якому три невідомих коефіцієнта  $b_0$ ,  $b_1$  та  $b_2$  ( $K=3$ ) і  $N>K$ .

Також ця матриця використовується для лінійної регресії з врахуванням взаємодії факторів при двох факторах  $\hat{y}=b_0+b_1x_1+b_2x_2+b_{12}x_1x_2$ , в якій чотири невідомих коефіцієнта  $b_0$ ,  $b_1$ ,  $b_2$  та  $b_{11}$  ( $K=4$ ), тобто в цьому випадку  $N=K$ .

Послідовність комбінацій дуже легко отримати, скориставшись двійковим кодом для чисел 0,1,2 та 3 (див. таблицю 3), замінивши **0** на **-1**, а **1** на **+1**.

Таблиця 3

Число	Двійковий код числа
0	<b>00</b>
1	<b>01</b>
2	<b>10</b>
3	<b>11</b>

Факторний простір для двох факторів (**k=2**) -- це декартова система координат, осі якої – кодовані значення факторів  $x_2$  та  $x_1$ . Кожній з чотирьох комбінацій значень факторів відповідає одна точка факторного простору, а точка з «нульовими» координатами **[0,0]** (центр експерименту) відповідає основним рівням факторів  $x_{i0}$  ( $i=\overline{1,k}$ ). Геометрично ПФЕ для двох факторів (**k=2**) можемо уявити у вигляді вершин квадрату з центром в точці **[0,0]**. Ординати та абсциси вершин квадрата -- кодові значення факторів для усіх чотирьох комбінацій, тобто **[-1,-1]**, **[-1,+1]**, **[+1,-1]** та **[+1,+1]**. Геометрично ПФЕ для трьох факторів (**k=3**) можемо уявити у вигляді вершин кубу {центр в точці **[0,0,0]**}, а для **k** факторів (**k>3**) - у вигляді вершин **k**-мірного гіперкубу з центром в точці **[0,0,...0]**.

Властивості матриці планування будь-якого ПФЕ для **r=2** незалежно від кількості факторів **k**:

1.Симетричність плану відносно центру експерименту:  $\sum_1^N \bar{x}_{ij} = 0$  ( $i=\overline{1,k}$ ). Це означає, що сума **N** кодових значень будь-якого фактору дорівнює нулю.

2.Нормування плану відносно центру експерименту:  $\sum_1^N \bar{x}_{ij}^2 = 0$  ( $i=\overline{1,k}$ ). Це означає, що сума квадратів **N** кодових значень будь-якого фактору дорівнює **N**.

3.Ортогональність плану:  $\sum_1^N \bar{x}_{ij} \bar{x}_{sj} = 0$  ( $i,s=\overline{1,k}; s \neq i$ ). Це означає, що скалярний добуток вектор-стовпців матриці дорівнює нулю.

4.Ротатабельність плану (походить від англійського слова **rotate**):  $R = \sqrt{r}$ , де **R** – відстань від центру до будь-якої точки факторного простору. Це означає, що усі точки факторного простору розташовані на однаковій відстані від центру експеримента. Так для **k=2** усі точки факторного простору розташовані на колі раді-

усом  $R=\sqrt{2}$ , для  $k=3$  усі точки факторного простору розташовані на сфері радіусом  $R=\sqrt{3}$  тощо.

Усі ці властивості легко перевірити для  $k=2$ :

1. Симетричність плану:  $\bar{x}_{11}+\bar{x}_{12}+\bar{x}_{13}+\bar{x}_{14}=(-1)+(-1)+(+1)+(+1)=0$ ;

$$\bar{x}_{21}+\bar{x}_{22}+\bar{x}_{23}+\bar{x}_{24}=(-1)+(+1)+(-1)+(+1)=0$$

2. Нормування плану:  $\bar{x}_{11}^2+\bar{x}_{12}^2+\bar{x}_{13}^2+\bar{x}_{14}^2=(-1)^2+(-1)^2+(+1)^2+(+1)^2=4$ ;

$$\bar{x}_{21}^2+\bar{x}_{22}^2+\bar{x}_{23}^2+\bar{x}_{24}^2=(-1)^2+(+1)^2+(-1)^2+(+1)^2=4$$

3. Ортогональність плану:  $\bar{x}_{11}\bar{x}_{21}+\bar{x}_{12}\bar{x}_{22}+\bar{x}_{13}\bar{x}_{23}+\bar{x}_{14}\bar{x}_{24}=(-1)(-1)+(-1)(+1)+$

$$+(+1)(-1)+(+1)(+1)=0$$

4. Ротатабельність плану: усі четверо точок факторного простору розташовані на колі радіусом  $R=\sqrt{2}$ .

Для того, щоб не тільки МП ПФЕ для будь-якої кількості факторів  $k>2$ , які використовують два рівня факторів ( $r=2$ ), мала вищезазначені чотири властивості, а і будь-яка необхідна матриця планування дробового факторного експерименту МП ДФЕ також мала ці властивості, необхідно певним чином здійснювати побудову МП ПФЕ для наступної кількості факторів на основі двох попередніх МП ПФЕ.

Таблиця 4 Трафарет для МП ПФЕ  $k=3$

N комбінації	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,8}; j=\overline{1,m})$
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	
1				$y_{11}= \quad ; y_{12}= \quad ; \dots\dots\dots y_{1m}= \quad ;$
2				$y_{21}= \quad ; y_{22}= \quad ; \dots\dots\dots y_{2m}= \quad ;$
3				$y_{31}= \quad ; y_{32}= \quad ; \dots\dots\dots y_{3m}= \quad ;$
4				$y_{41}= \quad ; y_{42}= \quad ; \dots\dots\dots y_{4m}= \quad ;$
5				$y_{51}= \quad ; y_{52}= \quad ; \dots\dots\dots y_{5m}= \quad ;$
6				$y_{61}= \quad ; y_{62}= \quad ; \dots\dots\dots y_{6m}= \quad ;$
7				$y_{71}= \quad ; y_{72}= \quad ; \dots\dots\dots y_{7m}= \quad ;$
8				$y_{81}= \quad ; y_{82}= \quad ; \dots\dots\dots y_{8m}= \quad ;$

Розглянемо побудову **МП ПФЕ** для трьох факторів (**k=3**), яка також використовує два рівня факторів (**r=2**), кодовані значення яких **+1** та **-1**. Ця матриця має кількість комбінацій – **8** ( $N_n=2^3=8$ ) і для побудови її необхідно використати дві **МП ПФЕ** для двох факторів наступним чином. Трафарет для такої матриці наведено у табл.4.

Далі беремо із табл.3 чотири значення для  $\bar{X}_1$  ( $\bar{X}_{11}, \bar{X}_{12}, \bar{X}_{13}$  та  $\bar{X}_{14}$ ) та чотири значення для  $\bar{X}_2$  ( $\bar{X}_{21}, \bar{X}_{22}, \bar{X}_{23}$  та  $\bar{X}_{24}$ ) і вносимо цю інформацію у відповідні місця табл.4 (див.табл.5). Потім необхідно повторити таку операцію ще один раз, заповнивши наступні чотири ряди для  $\bar{X}_1$  та для  $\bar{X}_2$  ( $\bar{X}_{15}=\bar{X}_{11}, \bar{X}_{16}=\bar{X}_{12}, \bar{X}_{17}=\bar{X}_{13}, \bar{X}_{18}=\bar{X}_{14}, \bar{X}_{25}=\bar{X}_{21}, \bar{X}_{26}=\bar{X}_{22}, \bar{X}_{27}=\bar{X}_{23}, \bar{X}_{28}=\bar{X}_{24}$ ).

Таблиця 5

N комбінації	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,8}; j=\overline{1, m})$
	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_3$	
1	-1	-1		$y_{11}=\quad ; y_{12}=\quad ; \dots\dots\dots y_{1m}=\quad ;$
2	-1	+1		$y_{21}=\quad ; y_{22}=\quad ; \dots\dots\dots y_{2m}=\quad ;$
3	+1	-1		$y_{31}=\quad ; y_{32}=\quad ; \dots\dots\dots y_{3m}=\quad ;$
4	+1	+1		$y_{41}=\quad ; y_{42}=\quad ; \dots\dots\dots y_{4m}=\quad ;$
5	-1	-1		$y_{51}=\quad ; y_{52}=\quad ; \dots\dots\dots y_{5m}=\quad ;$
6	-1	+1		$y_{61}=\quad ; y_{62}=\quad ; \dots\dots\dots y_{6m}=\quad ;$
7	+1	-1		$y_{71}=\quad ; y_{72}=\quad ; \dots\dots\dots y_{7m}=\quad ;$
8	+1	+1		$y_{81}=\quad ; y_{82}=\quad ; \dots\dots\dots y_{8m}=\quad ;$

Після цього необхідно заповнити стовпець  $\bar{X}_3$  наступним чином: для визначення перших чотирьох значень  $\bar{X}_3$  використовується формула  $\bar{X}_3=\bar{X}_1\bar{X}_2$ , а для наступних чотирьох значень --  $-\bar{X}_3=\bar{X}_1\bar{X}_2$ , тобто  $\bar{X}_{31}=\bar{X}_{11}\bar{X}_{21}, \bar{X}_{32}=\bar{X}_{12}\bar{X}_{22}, \bar{X}_{33}=\bar{X}_{13}\bar{X}_{23}, \bar{X}_{34}=\bar{X}_{14}\bar{X}_{24}, \bar{X}_{35}=-\bar{X}_{11}\bar{X}_{21}, \bar{X}_{36}=-\bar{X}_{12}\bar{X}_{22}, \bar{X}_{37}=-\bar{X}_{13}\bar{X}_{23},$  та  $\bar{X}_{38}=-\bar{X}_{14}\bar{X}_{24}$  (див. табл.6).

Таблиця 6 МП ПФЕ k=3

N комбінації i	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,8}; j=\overline{1, m})$
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	
1	-1	-1	+1	$y_{11}=\quad ; y_{12}=\quad ; \dots\dots\dots y_{1m}=\quad ;$
2	-1	+1	-1	$y_{21}=\quad ; y_{22}=\quad ; \dots\dots\dots y_{2m}=\quad ;$
3	+1	-1	-1	$y_{31}=\quad ; y_{32}=\quad ; \dots\dots\dots y_{3m}=\quad ;$
4	+1	+1	+1	$y_{41}=\quad ; y_{42}=\quad ; \dots\dots\dots y_{4m}=\quad ;$
5	-1	-1	-1	$y_{51}=\quad ; y_{52}=\quad ; \dots\dots\dots y_{5m}=\quad ;$
6	-1	+1	+1	$y_{61}=\quad ; y_{62}=\quad ; \dots\dots\dots y_{6m}=\quad ;$
7	+1	-1	+1	$y_{71}=\quad ; y_{72}=\quad ; \dots\dots\dots y_{7m}=\quad ;$
8	+1	+1	-1	$y_{81}=\quad ; y_{82}=\quad ; \dots\dots\dots y_{8m}=\quad ;$

Усі чотири властивості матриці планування ПФЕ для  $k=3$  є в наявності:

1. Симетричність плану:

$$\bar{x}_{11}+\bar{x}_{12}+\bar{x}_{13}+\bar{x}_{14}+\bar{x}_{15}+\bar{x}_{16}+\bar{x}_{17}+\bar{x}_{18}=(-1)+(-1)+(1)+(1)+(-1)+(-1)+(1)+(1)=0;$$

$$\bar{x}_{21}+\bar{x}_{22}+\bar{x}_{23}+\bar{x}_{24}+\bar{x}_{25}+\bar{x}_{26}+\bar{x}_{27}+\bar{x}_{28}=(-1)+(1)+(-1)+(1)+(-1)+(1)+(-1)+(1)=0;$$

$$\bar{x}_{31}+\bar{x}_{32}+\bar{x}_{33}+\bar{x}_{34}+\bar{x}_{35}+\bar{x}_{36}+\bar{x}_{37}+\bar{x}_{38}=(1)+(-1)+(-1)+(1)+(-1)+(1)+(1)+(-1)=0;$$

2. Нормування плану:

$$\bar{x}_{11}^2+\bar{x}_{12}^2+\bar{x}_{13}^2+\bar{x}_{14}^2+\bar{x}_{15}^2+\bar{x}_{16}^2+\bar{x}_{17}^2+\bar{x}_{18}^2=(-1)^2+(-1)^2+(1)^2+(1)^2+(-1)^2+(-1)^2+(1)^2+(1)^2=8;$$

$$\bar{x}_{21}^2+\bar{x}_{22}^2+\bar{x}_{23}^2+\bar{x}_{24}^2+\bar{x}_{25}^2+\bar{x}_{26}^2+\bar{x}_{27}^2+\bar{x}_{28}^2=(-1)^2+(1)^2+(-1)^2+(1)^2+(-1)^2+(1)^2+(-1)^2+(1)^2=8;$$

$$\bar{x}_{31}^2+\bar{x}_{32}^2+\bar{x}_{33}^2+\bar{x}_{34}^2+\bar{x}_{35}^2+\bar{x}_{36}^2+\bar{x}_{37}^2+\bar{x}_{38}^2=(1)^2+(-1)^2+(-1)^2+(1)^2+(-1)^2+(1)^2+(1)^2+(-1)^2=8;$$

3. Ортогональність плану:

$$\bar{x}_{11}\bar{x}_{21}+\bar{x}_{11}\bar{x}_{31}+\bar{x}_{21}\bar{x}_{31}+\bar{x}_{12}\bar{x}_{22}+\bar{x}_{12}\bar{x}_{32}+\bar{x}_{22}\bar{x}_{32}+\bar{x}_{13}\bar{x}_{33}+\bar{x}_{23}\bar{x}_{33}+\bar{x}_{11}\bar{x}_{21}+\bar{x}_{14}\bar{x}_{24}+\bar{x}_{14}\bar{x}_{34}+\bar{x}_{24}\bar{x}_{34}+\bar{x}_{15}\bar{x}_{25}+\bar{x}_{15}\bar{x}_{35}+\bar{x}_{25}\bar{x}_{35}+\bar{x}_{16}\bar{x}_{26}+\bar{x}_{16}\bar{x}_{36}+\bar{x}_{26}\bar{x}_{36}+\bar{x}_{17}\bar{x}_{27}+\bar{x}_{17}\bar{x}_{37}+\bar{x}_{27}\bar{x}_{37}+\bar{x}_{18}\bar{x}_{28}+\bar{x}_{18}\bar{x}_{38}+$$

$$\bar{x}_1\bar{x}_3 = (-1)(-1) + (-1)(+1) + (-1)(+1) + (-1)(+1) + (-1)(-1) + (+1)(-1) + (+1)(-1) + (+1)(-1) +$$

$$(-1)(-1) + (+1)(+1) + (+1)(+1) + (+1)(+1) + (-1)(-1) + (-1)(-1) + (-1)(-1) + (-1)(+1) + (-1)(+1) +$$

$$(+1)(+1) + (+1)(-1) + (+1)(+1) + (-1)(+1) + (+1)(+1) + (+1)(-1) + (+1)(-1) = 0$$

4. Ротатабельність плану: усі вісім точок факторного простору розташовані на сфері радіусом  $R = \sqrt{3}$ .

З МП ПФЕ для трьох факторів ( $k=3$ ) можемо отримати МП ДФЕ  $2^{3-1}$  (напіврепліку), якщо візьмемо перші чотири комбінації (перші чотири ряди) цієї МП ПФЕ (див. табл. 7).

Таблиця 7 МП ДФЕ  $2^{3-1}$

N комбінації i	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=1,4; j=1, m)$
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	
1	-1	-1	+1	$y_{11} = \quad ; y_{12} = \quad ; \dots \dots \dots y_{1m} = \quad ;$
2	-1	+1	-1	$y_{21} = \quad ; y_{22} = \quad ; \dots \dots \dots y_{2m} = \quad ;$
3	+1	-1	-1	$y_{31} = \quad ; y_{32} = \quad ; \dots \dots \dots y_{3m} = \quad ;$
4	+1	+1	+1	$y_{41} = \quad ; y_{42} = \quad ; \dots \dots \dots y_{4m} = \quad ;$

Усі чотири властивості матриці планування ДФЕ  $2^{3-1}$  також є в наявності:

1. Симетричність плану:  $\bar{x}_{11} + \bar{x}_{12} + \bar{x}_{13} + \bar{x}_{14} = (-1) + (-1) + (+1) + (+1) = 0;$

$$\bar{x}_{21} + \bar{x}_{22} + \bar{x}_{23} + \bar{x}_{24} = (-1) + (+1) + (-1) + (+1) = 0;$$

$$\bar{x}_{31} + \bar{x}_{32} + \bar{x}_{33} + \bar{x}_{34} = (+1) + (-1) + (-1) + (+1) = 0$$

2. Нормування плану:  $\bar{x}_{11}^2 + \bar{x}_{12}^2 + \bar{x}_{13}^2 + \bar{x}_{14}^2 = (-1)^2 + (-1)^2 + (+1)^2 + (+1)^2 = 4;$

$$\bar{x}_{21}^2 + \bar{x}_{22}^2 + \bar{x}_{23}^2 + \bar{x}_{24}^2 = (-1)^2 + (+1)^2 + (-1)^2 + (+1)^2 = 4;$$

$$\bar{x}_{31}^2 + \bar{x}_{32}^2 + \bar{x}_{33}^2 + \bar{x}_{34}^2 = (+1)^2 + (-1)^2 + (-1)^2 + (+1)^2 = 4$$

3. Ортогональність плану:

$$\bar{x}_{11}\bar{x}_{21} + \bar{x}_{11}\bar{x}_{31} + \bar{x}_{21}\bar{x}_{31} + \bar{x}_{12}\bar{x}_{22} + \bar{x}_{12}\bar{x}_{32} + \bar{x}_{22}\bar{x}_{32} + \bar{x}_{13}\bar{x}_{23} + \bar{x}_{13}\bar{x}_{33} + \bar{x}_{23}\bar{x}_{33} + \bar{x}_{14}\bar{x}_{24} + \bar{x}_{14}\bar{x}_{34} + \bar{x}_{24}\bar{x}_{34}$$

$$= (-1)(-1) + (-1)(+1) + (-1)(+1) + (-1)(+1) + (-1)(-1) + (+1)(-1) + (+1)(-1) + (+1)(-1) +$$

$$+ (-1)(-1) + (+1)(+1) + (+1)(+1) + (+1)(+1) = 0$$

4. Ротатабельність плану: усі четверо точок факторного простору розташовані на сфері радіусом  $R=\sqrt{3}$ .

Слід зазначити, що з **МП ПФЕ** для трьох факторів ( $k=3$ ) можемо отримати ще одну (другу) **МП ДФЕ**  $2^{3-1}$  (напіврепліку), якщо візьмемо другі чотири комбінації (ряди 5,6,7 та 8) **МП ПФЕ** і лише позамінюємо нумерацію рядів (див. табл.8)

Таблиця 8 Друга **МП ДФЕ**  $2^{3-1}$

N комбінації i	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,4}; j=\overline{1, m})$
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	
1	-1	-1	-1	$y_{11}=\quad ; y_{12}=\quad ; \dots\dots\dots y_{1m}=\quad ;$
2	-1	+1	+1	$y_{21}=\quad ; y_{22}=\quad ; \dots\dots\dots y_{2m}=\quad ;$
3	+1	-1	+1	$y_{31}=\quad ; y_{32}=\quad ; \dots\dots\dots y_{3m}=\quad ;$
4	+1	+1	-1	$y_{41}=\quad ; y_{42}=\quad ; \dots\dots\dots y_{4m}=\quad ;$

Усі чотири властивості матриці планування другої **ДФЕ**  $2^{3-1}$  також є в наявності.

Аналогічно можливо побудувати **МП ПФЕ** для чотирьох факторів ( $k=4$ ), використавши дві **МП ПФЕ** для трьох факторів, **МП ПФЕ** для п'яти факторів ( $k=5$ ), використавши дві **МП ПФЕ** для чотирьох факторів, тощо. При використанні такої методики побудови матриць планування **ПФЕ** не тільки усі такі **МП ПФЕ** мають вищезазначені чотири властивості, а й усі **МП ДФЕ** потрібного степеню дробовості також будуть мати ці властивості. Треба мати на увазі, що необхідна степінь дробовості **ДФЕ** залежить від кількості факторів та від кількості членів рівняння регресії. З одного боку, кількість комбінацій рівнів факторів (кількість точок факторного простору)  $N$  не може бути меншою, ніж  $K$ , тобто  $N \geq K$ , а з іншого боку, з метою мінімізації витрат при дослідженнях, необхідно обирати мінімально можливе значення  $N$ .

Для двох факторів ( $k=2$ ) при лінійній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює трьом ( $K=3$ ). **ПФЕ** при використанні двох рівнів факторів має 4 комбінації рівнів факторів ( $N_n=2^2=4$ ) і в цьому випадку забезпечується виконання нерівності  $N \geq K$ . Обрати варіант **ДФЕ**  $2^{2-1}$  неможливо, тому що тоді  $N_d=2 < K$ .

Для трьох факторів ( $k=3$ ) при лінійній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює чотирьом ( $K=4$ ). **ПФЕ** при використанні двох рів-

нів факторів має **8** комбінації рівнів факторів ( $N_n=2^3=8$ ), а ДФЕ  $2^{3-1}$  має **4** комбінації рівнів факторів і цього достатньо ( $N_d=K$ ). Отже доцільно для  $k=3$  обирати не ПФЕ, а ДФЕ  $2^{3-1}$ .

Для чотирьох факторів ( $k=4$ ) при лінійній регресії кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює п'ятьом ( $K=5$ ). ПФЕ при використанні двох рівнів факторів має **16** комбінацій рівнів факторів ( $N_n=2^4=16$ ), а ДФЕ  $2^{4-1}$  має **8** комбінацій рівнів факторів. ДФЕ  $2^{4-2}$  має лише **4** комбінації рівнів факторів і цього недостатньо ( $N_d < K$ ). Отже доцільно для  $k=4$  обрати не ПФЕ, а ДФЕ  $2^{4-1}$ .

Для п'яти факторів ( $k=5$ ) при лінійній регресії кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює шести ( $K=6$ ). ПФЕ при використанні двох рівнів факторів має **32** комбінації рівнів факторів ( $N_n=2^5=32$ ); ДФЕ  $2^{5-1}$  має **16** комбінацій рівнів факторів, ДФЕ  $2^{5-2}$  має **8** комбінацій рівнів факторів. ДФЕ  $2^{5-3}$  має лише **4** комбінації рівнів факторів і цього недостатньо ( $N_d < K$ ). Таким чином доцільно обрати не ПФЕ, а ДФЕ  $2^{5-2}$ .

Аналогічно можемо знайти оптимальний варіант ДФЕ для будь-якої кількості факторів. Слід зазначити, що для лінійної регресії при  $k > 2$  ніколи не потрібний ПФЕ, а достатньо використовувати відповідний ДФЕ.

Розглянемо варіанти, якщо є лінійна регресія з врахуванням взаємодії факторів, починаючи з  $k=2$ , тому що для  $k=1$  цих доданків в рівнянні регресії не існує.

Для двох факторів ( $k=2$ ) при лінійній регресії з врахуванням взаємодії факторів кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює чотирьом ( $K=4$ ). ПФЕ при використанні двох рівнів факторів також має **4** комбінації рівнів факторів ( $N_n=2^2=4$ ) і при цьому теж забезпечується виконання нерівності  $N \geq K$  ( $N=K$ ).

Для трьох факторів ( $k=3$ ) при лінійній регресії з врахуванням взаємодії факторів кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює восьми ( $K=8$ ). ПФЕ при двох рівнях факторів теж має **8** комбінацій рівнів факторів ( $N_n=2^3=8$ ), і в цьому випадку також забезпечується виконання нерівності  $N \geq K$  ( $N=K$ ).

Для чотирьох факторів ( $k=4$ ) при лінійній регресії з врахуванням взаємодії факторів кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює шістнадцяти ( $K=16$ ). ПФЕ при використанні двох рівнів факторів теж має **16** комбінацій

рівнів факторів ( $N_{\Pi}=2^4=16$ ), і в цьому випадку також забезпечується виконання нерівності  $N \geq K$  ( $N=K$ ).

Для п'яти факторів ( $k=5$ ) при лінійній регресії з врахуванням взаємодії факторів кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює тридцяти двом ( $K=32$ ). ПФЕ при використанні двох рівнів факторів також має 32 комбінації рівнів факторів ( $N_{\Pi}=2^5=32$ ), і в цьому випадку теж забезпечується виконання нерівності  $N \geq K$  ( $N=K$ ).

Не тільки при кількості факторів 3,4 або 5, а і для будь-якої іншої кількості факторів ( $k > 5$ ) при лінійній регресії з врахуванням взаємодії факторів завжди необхідно використовувати ПФЕ, оскільки завжди кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює кількості комбінацій  $N_{\Pi}=2^k$  ( $K=N_{\Pi}$ ).

Слід зазначити, що для лінійної регресії з врахуванням взаємодії факторів є випадок (для  $k > 2$ ), коли необхідно обирати не дворівневий ПФЕ, а збільшити кількість рівнів факторів та збільшити кількість комбінацій. Це трапляється, якщо з'ясується, що усі коефіцієнти рівняння регресії значущі і тоді буде неможливо зробити перевірку адекватності моделі оригіналу по критерію Фішера.

Якщо після перевірки адекватності рівняння лінійної регресії (моделі) оригіналу (усім експериментальним значенням вихідної змінної) і отриманні негативного результату (адекватність не підтверджується з обраною ймовірністю) необхідно збільшити кількість членів рівняння регресії і перейти до рівняння лінійної регресії з врахуванням взаємодії факторів. Для цього потрібно здійснити перехід від ДФЕ до ПФЕ, збільшивши кількість комбінацій до  $2^k$ . Це дозволяє повністю використовувати усі експериментальні значення вихідної змінної, які отримані згідно МП ДФЕ.

Якщо після перевірки адекватності рівняння лінійної регресії з врахуванням взаємодії факторів (моделі) оригіналу і отриманні негативного результату необхідно збільшити кількість членів рівняння регресії і перейти до рівняння квадратичної регресії. Для цього потрібно збільшити кількість рівнів факторів і здійснити перехід від МП ПФЕ до матриці планування одного з композиційних планів, доповнивши МП ПФЕ необхідною кількістю комбінацій, тобто збільшити кількість точок факторного простору.

Це дозволяє повністю використовувати усі експериментальні значення вихідної змінної, які отримані згідно МП ПФЕ.

## Композиційні плани факторних експериментів

Для квадратичної регресії кількість невідомих коефіцієнтів рівняння регресії **K** завжди перевищує кількість комбінацій **ПФЕ**, в якому використовуються тільки два рівня факторів, оскільки значення **K** для квадратичної регресії перевищує значення **K** для відповідного рівняння лінійної регресії з врахуванням взаємодії факторів на кількість квадратичних доданків **k**. Таким чином, для квадратичної регресії недостатньо використовувати два рівня факторів і необхідно збільшити кількість рівнів факторів. Плани експериментів з кількістю рівнів більше двох (**r>2**) мають назву «**композиційні плани**» (**КП**).

Для одного фактора (**k=1**) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює трьом ( $\hat{y}=b_0+b_1x+b_2x^2$ ) і тому достатньо використати композиційний план з трьома рівнями факторів (**r=3**), кодові значення яких **-1, 0** та **+1**. Матрицю планування такого композиційного плану (централь-но-композиційного плану першого порядку) **МП КП** представлено в табл.9

Таблиця 9 МП КП k=1

N комбінації	Фактори	Експериментальні значення вихідної змінної об'єкту
	$\bar{x}_1$	$y_{ij}(i=\overline{1,3}; j=\overline{1,m})$
1	<b>-1</b>	$y_{11}= \quad ; y_{12}= \quad ; \dots\dots\dots y_{1m}= \quad ;$
2	<b>0</b>	$y_{21}= \quad ; y_{22}= \quad ; \dots\dots\dots y_{2m}= \quad ;$
3	<b>+1</b>	$y_{31}= \quad ; y_{32}= \quad ; \dots\dots\dots y_{3m}= \quad ;$

Для іншої кількості факторів (**k>1**) при квадратичній регресії в теорії планування експерименту використовуються лише композиційний план з п'ятьма рівнями факторів (**r=5**), кодові значення яких **-1, 0, +1, -1** та **+1**, де **l** – величина зоряного плеча. Зоряного плече – це відстань від нульової (центральної) точки факторного простору (початку кодованої системи координат  $\bar{x}_k \bar{x}_{k-1} \dots \bar{x}_2 \bar{x}_1$ ) до будь-якої зоряної точки, розташованої на одній з осей. Це дозволяє побудувати **МП КП**, використавши за основу дворівневу **МП ПФЕ**, додавши до неї ще деякі п'ятирівневі комбінації для того, щоб забезпечити виконання нерівності **N≥K**, тобто додати до **2<sup>k</sup>** точок факторного простору зоряні точки в кількості **2k** та, можливо, деяку кількість нульових точок. Нульова (центральна) точка факторного простору (початок кодованої системи координат  $\bar{x}_k \bar{x}_{k-1} \dots \bar{x}_2 \bar{x}_1$ ) – це точка з координатами **[0,0,...0,0]**. Зоряні точки розташовані по осях кодованої системи координат на відстані **l** від центральної точки, де **l** – величина зоряного плеча, значення яко-

го залежить від обраного варіанту композиційного плану. Такий метод побудови композиційних планів дозволяє повністю використовувати усі експериментальні значення вихідної змінної, які отримані згідно дворівневій МП ПФЕ.

Слід зазначити, що такі властивості матриці будь-якого композиційному плані, як симетричність плану забезпечуються завжди, а забезпечити при цьому і ортогональність і ротатабельність плану одночасно неможливо; а нормування плану забезпечується виключно разом з ротатабельністю.

Якщо забезпечується ортогональність композиційного плану, -- такий композиційний план має назву: **центральний ортогональний композиційний план (ортогональний центрально-композиційний план другого порядку) ЦОКП**.

**P.S.:** Центральний тому, що окрім зоряних точок додається декілька нульових точок.

Якщо забезпечується ротатабельність та нормування композиційного плану, -- такий композиційний план має назву: **ротатабельний композиційний план (ротатабельний центрально-композиційний план другого порядку) РКП**.

Для забезпечення ортогональності, тобто для ЦОКП, значення зоряного плеча  $l$  розраховується з біквдратного рівняння:  $4l^4+4-2^k l^2-2^k(2k+1)=0$ . Так для  $k=2$  значення  $l=1$ ; для  $k=3$  значення  $l=\sqrt{(\sqrt{30}-4)}$ ; для  $k=4$  значення  $l=\sqrt{2}$  тощо.

Таблиця 10 МП ЦОКП  $k=2$

N комбінації	Фактори		Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,12}; j=\overline{1, m})$
	$\bar{x}_1$	$\bar{x}_2$	
1	-1	-1	$y_{11}=$ ; $y_{12}=$ ; ..... $y_{1m}=$ ;
2	-1	+1	$y_{21}=$ ; $y_{22}=$ ; ..... $y_{2m}=$ ;
3	+1	-1	$y_{31}=$ ; $y_{32}=$ ; ..... $y_{3m}=$ ;
4	+1	+1	$y_{41}=$ ; $y_{42}=$ ; ..... $y_{4m}=$ ;
5	-1	0	$y_{51}=$ ; $y_{52}=$ ; ..... $y_{5m}=$ ;
6	+1	0	$y_{61}=$ ; $y_{62}=$ ; ..... $y_{6m}=$ ;
7	0	-1	$y_{71}=$ ; $y_{72}=$ ; ..... $y_{7m}=$ ;
8	0	+1	$y_{81}=$ ; $y_{82}=$ ; ..... $y_{8m}=$ ;
9	0	0	$y_{91}=$ ; $y_{92}=$ ; ..... $y_{9m}=$ ;
10	0	0	$y_{101}=$ ; $y_{102}=$ ; ..... $y_{10m}=$ ;
11	0	0	$y_{111}=$ ; $y_{112}=$ ; ..... $y_{11m}=$ ;
12	0	0	$y_{121}=$ ; $y_{122}=$ ; ..... $y_{12m}=$ ;

Матрицю планування такого центрального ортогонального композиційного плану для  $k=2$  представлено в табл.10, де значення  $l$  дорівнює  $1$ . Точки факторного простору цього ЦОКП розташовані на трьох колах з радіусами  $1, \sqrt{2}$  і  $0$ .

Перші чотири комбінації – це комбінації ПФЕ, тобто перша частина МП ЦОКП співпадає з МП ПФЕ, другі чотири комбінації – це зоряні точки, а треті чотири комбінації - це центральні точки  $[0,0]$ . Центральних точок  $[0,0]$  може бути не чотири, а інша кількість. Рекомендації по обранню їх кількості відсутні в ТПЕ. Легко перевірити, що МП ЦОКП симетрична, незалежно від кількості нульових точок.

Таблиця 11 МП ЦОКП  $k=3$

N комбінації	Фактори			Експериментальні значення вихідної змінної об'єкту	
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$y_{ij}(i=\overline{1,18}; j=\overline{1, m})$	
1	-1	-1	+1	$y_{11} =$	$; y_{12} =$ ; ..... $y_{1m} =$ ;
2	-1	+1	-1	$y_{21} =$	$; y_{22} =$ ; ..... $y_{2m} =$ ;
3	+1	-1	-1	$y_{31} =$	$; y_{32} =$ ; ..... $y_{3m} =$ ;
4	+1	+1	+1	$y_{41} =$	$; y_{42} =$ ; ..... $y_{4m} =$ ;
5	-1	-1	-1	$y_{51} =$	$; y_{52} =$ ; ..... $y_{5m} =$ ;
6	-1	+1	+1	$y_{61} =$	$; y_{62} =$ ; ..... $y_{6m} =$ ;
7	+1	-1	+1	$y_{71} =$	$; y_{72} =$ ; ..... $y_{7m} =$ ;
8	+1	+1	-1	$y_{81} =$	$; y_{82} =$ ; ..... $y_{8m} =$ ;
9	-1	0	0	$y_{91} =$	$; y_{92} =$ ; ..... $y_{9m} =$ ;
10	+1	0	0	$y_{101} =$	$; y_{102} =$ ; ..... $y_{10m} =$ ;
11	0	-1	0	$y_{111} =$	$; y_{112} =$ ; ..... $y_{11m} =$ ;
12	0	+1	0	$y_{121} =$	$; y_{122} =$ ; ..... $y_{12m} =$ ;
13	0	0	-1	$y_{131} =$	$; y_{132} =$ ; ..... $y_{13m} =$ ;
14	0	0	+1	$y_{141} =$	$; y_{142} =$ ; ..... $y_{14m} =$ ;
15	0	0	0	$y_{151} =$	$; y_{152} =$ ; ..... $y_{15m} =$ ;
16	0	0	0	$y_{161} =$	$; y_{162} =$ ; ..... $y_{16m} =$ ;
17	0	0	0	$y_{171} =$	$; y_{172} =$ ; ..... $y_{17m} =$ ;
18	0	0	0	$y_{181} =$	$; y_{182} =$ ; ..... $y_{18m} =$ ;

Для двох факторів ( $k=2$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює шести ( $K=6$ ), а кількість комбінацій для

**ЦОКП** завжди більше ніж **8** ( $N_{\text{цокп}} > 8$ ), і в цьому випадку забезпечується виконання нерівності  $N > K$ .

Матрицю планування такого центрального ортогонального композиційного плану для  $k=3$  представлено в табл.11, де значення  $l$  дорівнює  $\sqrt{(\sqrt{30}-4)} \approx 1.215$ .

Перші вісім комбінацій – це комбінації **ПФЕ**, тобто перша частина **МП ЦОКП** співпадає з **МП ПФЕ**, другі шість комбінацій – це зоряні точки, а треті чотири комбінації - це центральні точки **[0,0]**. Центральних точок **[0,0]** може бути не чотири, а інша кількість. Рекомендації по обранню їх кількості відсутні в ТПЕ. Ця **МП ЦОКП** також симетрична, незалежно від кількості нульових точок. Точки факторного простору цього **ЦОКП** розташовані на трьох сферах з радіусами  $l$ ,  $\sqrt{2}$  і  $0$ .

Для трьох факторів ( $k=3$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює одинадцяти ( $K=11$ ), а кількість комбінацій для **ЦОКП** завжди більше ніж чотирнадцять ( $N_{\text{цокп}} > 14$ ), і в цьому випадку також забезпечується виконання нерівності  $N > K$ .

При побудові **МП ЦОКП** для  $k=4$ , додають до  $2^4$  точок **ПФЕ** вісім зоряних точок та слід обрати **вісім** нульових точок, тому що при цьому **буде забезпечено також і нормування ЦОКП**. До речі, це єдиний випадок, коли забезпечується і нормування **ЦОКП**.

Аналогічно можемо побудувати **МП ЦОКП** для  $k=5$ , додавши до  $2^5$  точок **ПФЕ** десять зоряних точок та нульові точки тощо.

Для чотирьох факторів ( $k=4$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії **K** дорівнює двадцяти ( $K=20$ ), а кількість комбінацій для **ЦОКП** завжди більше ніж двадцять чотири ( $N_{\text{цокп}} > 24$ ), і в цьому випадку також забезпечується виконання нерівності  $N > K$ .

Взагалі, для будь-якої кількості факторів **k** (в тому числі і для  $k > 4$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії можемо визначити по формулі  $K=2^k+k$ , а кількість комбінацій рівнів факторів для **ЦОКП** є більшою, ніж  $2^k+2k$ , на кількість нульових точок факторного простору і тому завжди забезпечується виконання нерівності  $N > K$ , тому що  $N-K > k$ .

Для забезпечення ротатабельності, тобто для РКП, значення зоряного плеча **l** розраховується за формулою:  $l=\sqrt{k}$ . Так для  $k=2$  значення  $l=\sqrt{2}$ ; для  $k=3$  значення

$l=\sqrt{3}$ ; для  $k=4$  значення  $l=\sqrt{4}=2$  тощо. Матрицю планування ротатабельного композиційного плану для  $k=2$  представлено в табл.12, де значення  $l$  дорівнює  $\sqrt{2}$ .

Таблиця 12 **МП РКП  $k=2$**

N комбінації	Фактори		Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,8}; j=\overline{1, m})$
	$\bar{x}_1$	$\bar{x}_2$	
1	-1	-1	$y_{11}=$ ; $y_{12}=$ ; ..... $y_{1m}=$ ;
2	-1	+1	$y_{21}=$ ; $y_{22}=$ ; ..... $y_{2m}=$ ;
3	+1	-1	$y_{31}=$ ; $y_{32}=$ ; ..... $y_{3m}=$ ;
4	+1	+1	$y_{41}=$ ; $y_{42}=$ ; ..... $y_{4m}=$ ;
5	-1	0	$y_{51}=$ ; $y_{52}=$ ; ..... $y_{5m}=$ ;
6	+1	0	$y_{61}=$ ; $y_{62}=$ ; ..... $y_{6m}=$ ;
7	0	-1	$y_{71}=$ ; $y_{72}=$ ; ..... $y_{7m}=$ ;
8	0	+1	$y_{81}=$ ; $y_{82}=$ ; ..... $y_{8m}=$ ;

Перші чотири комбінації – це комбінації ПФЕ, тобто перша частина МП РКП співпадає з МП ПФЕ а другі чотири комбінації – це зоряні точки. Легко перевірити, що МП РКП для  $k=2$  симетрична і нормована, оскільки  $l=\sqrt{2}$ . Усі точки факторного простору РКП для  $k=2$  розташовані на колі радіусом  $\sqrt{2}$ .

Для двох факторів ( $k=2$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює шести ( $K=6$ ), а МП РКП має вісім комбінацій рівнів факторів (точок факторного простору) ( $N_{\text{ркп}}=8$ ) і в цьому випадку забезпечується виконання нерівності  $N>K$ .

Матрицю планування такого центрального ортогонального композиційного плану для  $k=3$  представлено в табл.13, де значення  $l$  дорівнює  $\sqrt{3}$ .

Легко перевірити, що МП РКП для  $k=3$  є симетричною і нормованою, оскільки  $l=\sqrt{3}$ . Усі точки факторного простору РКП для  $k=3$  розташовані на сфері радіусом  $\sqrt{3}$ .

Для трьох факторів ( $k=3$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює шести ( $K=11$ ), а МП РКП має чотирнадцять

п'ять комбінацій рівнів факторів (точок факторного простору) ( $N_{\text{РКП}}=14$ ) і в цьому випадку також забезпечується виконання нерівності  $N > K$ .

Таблиця 13 МП РКП  $k=3$

N комбінації i	Фактори			Експериментальні значення вихідної змінної об'єкту $y_{ij}(i=\overline{1,14}; j=\overline{1,m})$
	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	
1	-1	-1	+1	$y_{11} = \quad ; y_{12} = \quad ; \dots \quad y_{1m} = \quad ;$
2	-1	+1	-1	$y_{21} = \quad ; y_{22} = \quad ; \dots \quad y_{2m} = \quad ;$
3	+1	-1	-1	$y_{31} = \quad ; y_{32} = \quad ; \dots \quad y_{3m} = \quad ;$
4	+1	+1	+1	$y_{41} = \quad ; y_{42} = \quad ; \dots \quad y_{4m} = \quad ;$
5	-1	-1	-1	$y_{51} = \quad ; y_{52} = \quad ; \dots \quad y_{5m} = \quad ;$
6	-1	+1	+1	$y_{61} = \quad ; y_{62} = \quad ; \dots \quad y_{6m} = \quad ;$
7	+1	-1	+1	$y_{71} = \quad ; y_{72} = \quad ; \dots \quad y_{7m} = \quad ;$
8	+1	+1	-1	$y_{81} = \quad ; y_{82} = \quad ; \dots \quad y_{8m} = \quad ;$
9	-1	0	0	$y_{91} = \quad ; y_{92} = \quad ; \dots \quad y_{9m} = \quad ;$
10	+1	0	0	$y_{101} = \quad ; y_{102} = \quad ; \dots \quad y_{10m} = \quad ;$
11	0	-1	0	$y_{111} = \quad ; y_{112} = \quad ; \dots \quad y_{11m} = \quad ;$
12	0	+1	0	$y_{121} = \quad ; y_{122} = \quad ; \dots \quad y_{12m} = \quad ;$
13	0	0	-1	$y_{131} = \quad ; y_{132} = \quad ; \dots \quad y_{13m} = \quad ;$
14	0	0	+1	$y_{141} = \quad ; y_{142} = \quad ; \dots \quad y_{14m} = \quad ;$

Аналогічно можемо побудувати МП РКП для  $k=4$ , додавши до  $2^4$  точок ПФЕ вісім зоряних точок та нульові точки; побудувати МП РКП для  $k=5$ , додавши до  $2^5$  точок ПФЕ десять зоряних точок та нульові точки тощо.

Для чотирьох факторів ( $k=4$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії  $K$  дорівнює двадцяти ( $K=20$ ), а МП РКП має двадцять чотири комбінації рівнів факторів (точок факторного простору) ( $N_{\text{РКП}}=24$ ) і в цьому випадку також забезпечується виконання нерівності  $N > K$ .

Взагалі для будь-якої кількості факторів  $k$  (в тому числі і для  $k > 4$ ) при квадратичній регресії кількість невідомих коефіцієнтів рівняння регресії можемо визначити по формулі  $K=2^k+k$ , а кількість комбінацій рівнів факторів для РКП дорівнює  $2^k+2k$ , і завжди забезпечується виконання нерівності  $N > K$ , тому що  $N-K > k$ .

Якщо після перевірки адекватності рівняння квадратичної регресії (моделі) оригіналу (усім експериментальним значенням вихідної змінної) і отриманні негативного результату (адекватність знову не підтверджується з обраною ймовірністю) необхідно збільшити кількість членів рівняння регресії. Для цього потрібно здійснити перехід від центрально-композиційного плану другого порядку до центрально-композиційного плану третього порядку доповнивши центрально-композиційний план другого порядку необхідною кількістю комбінацій і так далі до тих пір, поки не отримаємо підтвердження адекватності моделі.

Таким чином отримуємо статистичну математичну модель – рівняння регресії, яке адекватно оригіналу (об'єкту, що досліджувався), з обраною ймовірністю.

Згідно з цим планами проводяться експерименти в кількості  $n$   $\{n=Nm\}$ ; при проведенні кожного з них на входи подаються абсолютні значення вхідних змінних, які відповідають кодованим значенням і вимірюються значення вихідної змінної  $y_{ij}$  ( $i=\overline{1,N};j=\overline{1,m}$ ), які вносяться до МП. Нагадуємо, що кожна з  $N$  комбінацій повторюється  $m$  разів.

**P.S.:** Слід зазначити, що для кожної комбінації може обиратись і різна кількість повторів  $m_i$  ( $i=\overline{1,N}$ ).

Отримані експериментальні значення вихідної змінної  $y_{ij}$  ( $i=\overline{1,N};j=\overline{1,m}$ ) заносяться до матриці планування.

Після отримання усіх експериментальних значень вихідної змінної  $y_{ij}$  ( $i=\overline{1,N};j=\overline{1,m}$ ) доцільно (так буде зручніше при розрахунку коефіцієнтів рівняння регресії) зробити заміну індексів при цих змінних, застосувавши послідовну індексацію від  $1$  до  $n$ , тобто отримати інформацію у наступному вигляді:  $y_i$  ( $i=\overline{1,n}$ ), тобто  $y_{11}=y_1, \dots, y_{1m}=y_m, y_{21}=y_{m+1}, \dots, y_{km}=y_n$ .

### **Розрахунок коефіцієнтів рівняння регресії**

Розрахунок коефіцієнтів рівняння регресії здійснюється після підтвердження однорідності дисперсії з використанням методу найменших квадратів при апроксимації функцій.

**Апроксимація функції** – це заміна початкової функції (апроксимовуваної або апроксумованої) іншою функцією (апроксимуючою), яка суттєво наближена до початкової, тобто апроксимує її з потрібною, наперед заданою точністю. Початкова функція може бути або функцією однієї змінної  $y=f(x)$ , або функцією  $k$  змінних  $y=f(x_1, x_2, \dots, x_k)$ . Апроксимуючу функцію будемо позначати, як  $y=\varphi(x)$  або

$y = \varphi(x_1, x_2, \dots, x_k)$ . Критерієм точності наближення апроксимуючої функції одного аргументу  $y = \varphi(x)$  до апроксимованої функції одного аргументу  $y = f(x_1, x_2, \dots, x_k)$  є сума квадратів відхилень однієї функції від іншої для всього діапазону зміни аргументу, тобто для усіх значень  $x_i$  ( $i = \overline{1, n}$ ), якщо функції задано дискретно, або для  $x_{\min} \leq x \leq x_{\max}$ , якщо ці обидві функції задано аналітично. Аналогічно, для апроксимуючої функції  $k$  аргументів  $y = \varphi(x_1, x_2, \dots, x_k)$  критерієм точності її наближення до апроксимованої функції  $k$  аргументів  $y = f(x_1, x_2, \dots, x_k)$  є сума квадратів відхилень однієї функції від іншої для усіх діапазонів зміни аргументів, тобто для  $[x_{j\min} \leq x_j \leq x_{j\max} \text{ (} j = \overline{1, k}\text{)}]$ , якщо функції задано аналітично, або для усіх значень  $x_i$  ( $i = \overline{1, n}$ ), якщо функції задано дискретно.

Якщо критерієм точності є сума квадратів відхилень, -- це має назву «метод найменших квадратів». Це метод знаходження найближчої до апроксимованої функції апроксимуючої функції з декількох обраних, тобто та з  $m$  апроксимуючих функцій є найближчою, для якої значення  $\Phi_s$  ( $s = \overline{1, m}$ ) є мінімальним.

Для функцій з  $k$  аргументами  $x_j$  ( $j = \overline{1, k}$ ), які задано дискретно для  $n$  значень аргументу, -- значення  $\Phi_s$  обчислюються таким чином:

$$\Phi_s = \sum_1^n [\varphi_s(x_{1i}, x_{2i}, \dots, x_{ki}) - f(x_{1i}, x_{2i}, \dots, x_{ki})]^2 \text{ (} s = \overline{1, m}\text{)}.$$

**P.S.:** Якщо одну з функцій задано аналітично (частіше це апроксимуюча функція), а іншу – дискретно ( $n$  значеннями), необхідно зйти  $n$  значень першої функції.

В теорії планування експерименту апроксимуюча функція – це рівняння регресії з  $K$  невідомими коефіцієнтами  $\check{y} = \varphi(b_0, b_1, \dots, b_{K-1}, x_1, x_2, \dots, x_k)$ , а значення апроксимованої функції – це експериментальні значення вихідної змінної об'єкту  $y_i$  ( $i = \overline{1, n}$ ). При цьому критерій найменших квадратів для рівняння регресії може бути представлений таким чином:

$$\Phi(b_0, b_1, \dots, b_{K-1}) = \sum_1^n [\varphi(b_0, b_1, \dots, b_{K-1}, x_{1i}, x_{2i}, \dots, x_{ki}) - y_i]^2.$$

Розрахунок коефіцієнтів рівняння регресії необхідно виконати таким чином, щоб отримати мінімальне значення функції  $\Phi$ , тобто знайти такі значення коефіцієнтів рівняння регресії, при яких отримаємо мінімальне значення суми квадратів відхилень усіх експериментальних значень вихідної змінної  $y_i$  ( $i = \overline{1, n}$ ) від відповідних розрахункових значень  $\check{y}_i$  ( $i = \overline{1, n}$ ), отриманих з рівняння регресії.

Необхідною умовою мінімуму функції  $\Phi(b_0, b_1, \dots, b_{K-1})$  є виконання наступного: частинні похідні функції по змінним коефіцієнтам рівняння регресії  $\partial\Phi/\partial b_s$

$(s=\overline{0, K-1})$  повинні дорівнювати нулю, тобто  $\{\partial\Phi/\partial b_s=0 (s=\overline{0, K-1})\}$ . Для скорочення запису введемо наступне позначення  $\varphi(b_0, b_1, \dots, b_{K-1}, x_{1i}, x_{2i}, \dots, x_{ki}) = \varphi_i$ .

Виконавши підстановку значення  $\Phi(b_0, b_1, \dots, b_{K-1})$  в систему частинних похідних отримаємо наступну систему  $K$  рівнянь з  $K$  невідомими:

$$\left\{ \sum_1^n 2[\varphi_i - y_i] \right\} (\partial\varphi_i / \partial b_s) = 0 \quad (s=\overline{0, K-1}). \quad [1]$$

Розглянемо лінійну регресію для  $k=1$ :  $\hat{y} = b_0 + b_1 x = \varphi(b_0, b_1, x)$ . Для цього рівняння регресії  $\varphi_i = b_0 + b_1 x_i (i=\overline{1, n})$ . Тоді маємо:  $\partial\varphi_i / \partial b_0 = 1$  та  $\partial\varphi_i / \partial b_1 = x_i (i=\overline{1, n})$ . Виконавши підстановку значень  $\varphi_i$ ,  $\partial\varphi_i / \partial b_0$  і  $\partial\varphi_i / \partial b_1$  в систему рівнянь [1] для двох невідомих  $K=2$ , отримаємо систему двох лінійних рівнянь з двома невідомими:

$$\sum_1^n [(b_0 + b_1 x_i) - y_i] (1) = 0;$$

$$\sum_1^n [(b_0 + b_1 x_i) - y_i] (x_i) = 0.$$

Перетворимо цю систему в традиційну форму двох лінійних рівнянь з двома невідомими  $b_0$  та  $b_1$  і отримаємо наступну систему рівнянь:

$$(\sum_1^n 1) b_0 + (\sum_1^n x_i) b_1 = (\sum_1^n y_i)$$

$$(\sum_1^n x_i) b_0 + (\sum_1^n x_i^2) b_1 = (\sum_1^n y_i x_i)$$

Якщо поділити обидва рівняння на  $n$ , врахувавши, що  $(\sum_1^n 1) = n$ , і ввести заміни, використовуючи наступні формули для статистичних моментів та статистичних коефіцієнтів:  $m_x = (\sum_1^n x_i) / n$ ,  $m_y = (\sum_1^n y_i) / n$ ,  $a_2 = (\sum_1^n x_i^2) / n$  та  $a_{11} = (\sum_1^n y_i x_i) / n$ , отримаємо систему у такому вигляді:

$$b_0 + m_x b_1 = m_y;$$

$$m_x b_0 + a_2 b_1 = a_{11}.$$

Використавши теорію визначників, вирахуємо значення невідомих коефіцієнтів  $b_0$  та  $b_1$  по наступним формулам:

$$b_0 = (m_y a_2 - m_x a_{11}) / (a_2 - m_x^2); \quad b_1 = (a_{11} - m_x m_y) / (a_2 - m_x^2).$$

Розглянемо лінійну регресію для  $k=2$ :  $\hat{y} = b_0 + b_1 x_1 + b_2 x_2 = \varphi(b_0, b_1, b_2, x_1, x_2)$ . Для цього рівняння:  $\varphi_i = b_0 + b_1 x_{1i} + b_2 x_{2i} (i=\overline{1, n})$ . Тоді маємо:  $\partial\varphi_i / \partial b_0 = 1$ ;  $\partial\varphi_i / \partial b_1 = x_{1i}$  та  $\partial\varphi_i / \partial b_2 = x_{2i} (i=\overline{1, n})$ . Виконавши підстановку значень  $\varphi_i$ ,  $\partial\varphi_i / \partial b_0$ ,  $\partial\varphi_i / \partial b_1$  та

$\partial\varphi_i/\partial\mathbf{b}_2$  в систему рівнянь [1] для  $\mathbf{K}=3$ , отримаємо систему трьох лінійних рівнянь з трьома невідомими:

$$\sum_1^n[(\mathbf{b}_0+\mathbf{b}_1\mathbf{x}_{1i}+\mathbf{b}_2\mathbf{x}_{2i})-\mathbf{y}_i](1)=0;$$

$$\sum_1^n[(\mathbf{b}_0+\mathbf{b}_1\mathbf{x}_{1i}+\mathbf{b}_2\mathbf{x}_{2i})-\mathbf{y}_i](\mathbf{x}_{1i})=0;$$

$$\sum_1^n[(\mathbf{b}_0+\mathbf{b}_1\mathbf{x}_{1i}+\mathbf{b}_2\mathbf{x}_{2i})-\mathbf{y}_i](\mathbf{x}_{2i})=0.$$

Перетворимо і цю систему в традиційну форму трьох лінійних рівнянь з трьома невідомими  $\mathbf{b}_0$ ,  $\mathbf{b}_1$  та  $\mathbf{b}_2$ , та отримаємо наступну систему рівнянь:

$$(\sum_1^n 1)\mathbf{b}_0+(\sum_1^n \mathbf{x}_{1i})\mathbf{b}_1+(\sum_1^n \mathbf{x}_{2i})\mathbf{b}_2=(\sum_1^n \mathbf{y}_i);$$

$$(\sum_1^n \mathbf{x}_{1i})\mathbf{b}_0+(\sum_1^n \mathbf{x}_{1i}^2)\mathbf{b}_1+(\sum_1^n \mathbf{x}_{2i}\mathbf{x}_{1i})\mathbf{b}_2=(\sum_1^n \mathbf{y}_i\mathbf{x}_{1i});$$

$$(\sum_1^n \mathbf{x}_{2i})\mathbf{b}_0+(\sum_1^n \mathbf{x}_{1i}\mathbf{x}_{2i})\mathbf{b}_1+(\sum_1^n \mathbf{x}_{2i}^2)\mathbf{b}_2=(\sum_1^n \mathbf{y}_i\mathbf{x}_{2i}).$$

Якщо поділити усі рівняння на  $n$  і ввести наступні заміни, використовуючи наступні формули для статистичних моментів та статистичних коефіцієнтів:  $\mathbf{m}_{x1}=(\sum_1^n \mathbf{x}_{1i})/n$ ,  $\mathbf{m}_{x2}=(\sum_1^n \mathbf{x}_{2i})/n$ ,  $\mathbf{m}_y=(\sum_1^n \mathbf{y}_i)/n$ ,  $\mathbf{a}_{21}=(\sum_1^n \mathbf{x}_{1i}^2)/n$ ,  $\mathbf{a}_{11}=(\sum_1^n \mathbf{y}_i\mathbf{x}_{1i})/n$ ,  $\mathbf{a}_{22}=(\sum_1^n \mathbf{x}_{2i}^2)/n$ ,  $\mathbf{a}_{12}=(\sum_1^n \mathbf{y}_i\mathbf{x}_{2i})/n$  та  $\mathbf{a}_1=(\sum_1^n \mathbf{x}_{1i}\mathbf{x}_{2i})/n$ ,-- отримаємо систему рівнянь у такому вигляді:

$$\mathbf{b}_0+\mathbf{m}_{x1}\mathbf{b}_1+\mathbf{m}_{x2}\mathbf{b}_2=\mathbf{m}_y;$$

$$\mathbf{m}_{x1}\mathbf{b}_0+\mathbf{a}_{21}\mathbf{b}_1+\mathbf{a}_1\mathbf{b}_2=\mathbf{a}_{11}; \tag{2}$$

$$\mathbf{m}_{x2}\mathbf{b}_0+\mathbf{a}_1\mathbf{b}_1+\mathbf{a}_{22}\mathbf{b}_2=\mathbf{a}_{12}.$$

Використавши теорію визначників можемо розрахувати значення невідомих коефіцієнтів  $\mathbf{b}_0$ ,  $\mathbf{b}_1$  та  $\mathbf{b}_2$ . Аналогічно виконуються розрахунки коефіцієнтів лінійної регресії для будь-якого значення  $\mathbf{k}$ .

Розглянемо квадратичну регресію для  $\mathbf{k}=1$ :  $\ddot{\mathbf{y}}=\mathbf{b}_0+\mathbf{b}_1\mathbf{x}+\mathbf{b}_2\mathbf{x}^2=\varphi(\mathbf{b}_0,\mathbf{b}_1,\mathbf{b}_2,\mathbf{x})$ . Для цього рівняння:  $\varphi_i=\mathbf{b}_0+\mathbf{b}_1\mathbf{x}_i+\mathbf{b}_2\mathbf{x}_i^2$  ( $i=\overline{1,n}$ ). Тоді маємо:  $\partial\varphi_i/\partial\mathbf{b}_0=1$ ;  $\partial\varphi_i/\partial\mathbf{b}_1=\mathbf{x}_i$  та  $\partial\varphi_i/\partial\mathbf{b}_2=\mathbf{x}_i^2$  ( $i=\overline{1,n}$ ). Виконавши підстановку значень  $\varphi_i$ ,  $\partial\varphi_i/\partial\mathbf{b}_0$ ,  $\partial\varphi_i/\partial\mathbf{b}_1$  і  $\partial\varphi_i/\partial\mathbf{b}_2$  в систему рівнянь [1] для  $\mathbf{K}=3$ , отримаємо систему трьох лінійних рівнянь з трьома невідомими:

$$\sum_1^n[(\mathbf{b}_0+\mathbf{b}_1\mathbf{x}_i+\mathbf{b}_2\mathbf{x}_i^2)-\mathbf{y}_i](1)=0;$$

$$\sum_1^n [(b_0 + b_1 x_i + b_2 x_i^2) - y_i] (x_i) = 0;$$

$$\sum_1^n [(b_0 + b_1 x_i + b_2 x_i^2) - y_i] (x_i^2) = 0.$$

Перетворимо цю систему в традиційну форму трьох лінійних рівнянь з трьома невідомими  $b_0$ ,  $b_1$  та  $b_2$  і отримаємо наступну систему рівнянь:

$$(\sum_1^n 1) b_0 + (\sum_1^n x_i) b_1 + (\sum_1^n x_i^2) b_2 = (\sum_1^n y_i);$$

$$(\sum_1^n x_i) b_0 + (\sum_1^n x_i x_i) b_1 + (\sum_1^n x_i^2 x_i) b_2 = (\sum_1^n y_i x_i);$$

$$(\sum_1^n x_i^2) b_0 + (\sum_1^n x_i x_i^2) b_1 + (\sum_1^n x_i^4) b_2 = (\sum_1^n y_i x_i^2).$$

Якщо поділити усі рівняння на  $n$  і ввести наступні заміни:  $m_{x1} = (\sum_1^n x_i) / n$ ,  $m_{x2} = (\sum x_i^2) / n$ ,  $m_y = (\sum_1^n y_i) / n$ ,  $a_{21} = (\sum_1^n x_i x_i) / n = (\sum_1^n x_i^2) / n = m_{x2}$ ,  $a_{11} = (\sum_1^n y_i x_i) / n$ ,  $a_{11} = (\sum_1^n y_i x_i) / n$ ,  $a_{22} = (\sum_1^n x_i^2 x_i^2) / n = (\sum_1^n x_i^4) / n$ ,  $a_{12} = (\sum_1^n y_i x_i^2) / n$  та  $a_1 = (\sum_1^n x_i^2 x_i) / n = (\sum_1^n x_i x_i^2) / n = (\sum_1^n x_i^3) / n$ , -- отримаємо систему рівнянь у такому вигляді:

$$b_0 + m_{x1} b_1 + m_{x2} b_2 = m_y;$$

$$m_{x1} b_0 + a_{21} b_1 + a_1 b_2 = a_{11}; \quad [3]$$

$$m_{x2} b_0 + a_1 b_1 + a_{22} b_2 = a_{12}.$$

Використавши теорію визначників можемо розрахувати значення невідомих коефіцієнтів  $b_0$ ,  $b_1$  та  $b_2$ .

Система рівнянь [3], яку отримано для квадратичної регресії при  $k=1$ , тотожна системі рівнянь [2] для лінійної регресії при  $k=2$ , і обидві системи мають по три коефіцієнти. Проте маємо можливість одразу знайти рівняння лінійної регресії, яке буде тотожне рівнянню квадратичної регресії (лише для розрахунку коефіцієнтів рівняння регресії !!!), якщо одразу в цьому рівнянні ( $\hat{y} = b_0 + b_1 x + b_2 x^2$ ) виконати наступну заміну:  $x = x_1$  та  $x^2 = x_2$ .

Аналогічно, для будь-якого нелінійного рівняння регресії (для лінійної регресії з врахуванням взаємодії факторів, квадратичній регресії тощо), яке має  $K$  невідомих коефіцієнтів рівняння регресії, можемо отримати тотожне рівняння множинної лінійної регресії з  $k$  факторами ( $k=K-1$ ), яке також буде мати  $K$  невідомих коефіцієнтів, використавши відповідні заміни.

Наприклад, для квадратичної регресії з двома факторами, яка має шість коефіцієнтів рівняння ( $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_{12}x_1x_2 + b_{11}x_1^2 + b_{22}x_2^2$ ), введенням заміни  $b_{12} = b_3$ ,  $x_1x_2 = x_3$ ,  $b_{11} = b_4$ ,  $x_1^2 = x_4$ ,  $b_{22} = b_5$  та  $x_2^2 = x_5$ , отримуємо наступне тотожне рівняння лінійної регресії:  $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5$ .

Для квадратичної регресії з двома факторами, яка має одинадцять коефіцієнтів рівняння ( $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_{12}x_1x_2 + b_{13}x_1x_3 + b_{23}x_2x_3 + b_{123}x_1x_2x_3 + b_{11}x_1^2 + b_{22}x_2^2 + b_{33}x_3^2$ ), введенням заміни  $b_{12} = b_4$ ,  $x_1x_2 = x_4$ ,  $b_{13} = b_5$ ,  $x_1x_3 = x_5$ ,  $b_{23} = b_6$ ,  $x_2x_3 = x_6$ ,  $b_{123} = b_7$ ,  $x_1x_2x_3 = x_7$ ,  $b_{11} = b_8$ ,  $x_1^2 = x_8$ ,  $b_{22} = b_9$ ,  $x_2^2 = x_9$ ,  $b_{33} = b_{10}$  та  $x_3^2 = x_{10}$ , отримуємо наступне тотожне рівняння:  $\hat{y} = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + b_5x_5 + b_6x_6 + b_7x_7 + b_8x_8 + b_9x_9 + b_{10}x_{10}$ .

Отже необхідно розглядати розрахунок коефіцієнтів лише для лінійних регресій.

Розглянемо множинну лінійну регресію для  $k$  факторів:

$$\hat{y} = b_0 + \sum_1^k b_s x_s = \varphi(b_0, b_1, \dots, b_{k-1}, x_1, x_2, \dots, x_k).$$

Для цього рівняння:  $\varphi_i = b_0 + \sum_1^k b_s x_{si}$  ( $i = \overline{1, n}$ ) і  $\partial \varphi_i / \partial b_0 = 1$  та  $\partial \varphi_i / \partial b_k = x_{ki}$  ( $k = \overline{1, k}$ ).

Виконавши підстановку значень  $\varphi_i$ ,  $\partial \varphi_i / \partial b_0$  та  $\partial \varphi_i / \partial b_k$  ( $k = \overline{1, k}$ ) в систему рівнянь [1] для  $K$  невідомих ( $K = k + 1$ ), отримаємо систему  $K$  лінійних рівнянь з  $K$  невідомими:

$$\begin{aligned} \sum_1^n [(b_0 + \sum_1^k b_s x_{si}) - y_i] &= 0; \\ \sum_1^n [(b_0 + \sum_1^k b_s x_{si}) - y_i] (x_{1i}) &= 0; \\ \sum_1^n [(b_0 + \sum_1^k b_s x_{si}) - y_i] (x_{2i}) &= 0 \\ \dots & \\ \sum_1^n [(b_0 + \sum_1^k b_s x_{si}) - y_i] (x_{ki}) &= 0 \end{aligned}$$

Перетворивши цю систему в традиційну форму  $K$  лінійних рівнянь з  $K$  невідомими  $b_i$  ( $i = \overline{0, k}$ ),-- отримаємо наступну систему рівнянь:

$$\begin{aligned} (\sum_1^n 1) b_0 + (\sum_1^n x_{1i}) b_1 + (\sum_1^n x_{2i}) b_2 + \dots + (\sum_1^n x_{ki}) b_k &= (\sum_1^n y_i); \\ (\sum_1^n x_{1i}) b_0 + (\sum_1^n x_{1i} x_{1i}) b_1 + (\sum_1^n x_{2i} x_{1i}) b_2 + \dots + (\sum_1^n x_{ki} x_{1i}) b_k &= (\sum_1^n y_i x_{1i}); \\ (\sum_1^n x_{2i}) b_0 + (\sum_1^n x_{1i} x_{2i}) b_1 + (\sum_1^n x_{2i} x_{2i}) b_2 + \dots + (\sum_1^n x_{ki} x_{2i}) b_k &= (\sum_1^n y_i x_{2i}); \\ \dots & \end{aligned}$$

$$(\sum_1^n x_{ki})b_0 + (\sum_1^n x_{1i}x_{ki})b_1 + (\sum_1^n x_{2i}x_{ki})b_2 + \dots + (\sum_1^n x_{ki}x_{ki})b_k = (\sum_1^n y_i x_{ki}).$$

Якщо поділити усі рівняння на  $n$  та використати наступні заміни:

$$1 = a_{00}; \{(\sum_1^n x_{ji})/n\} = a_{j0} = a_{0i} \ (j = \overline{1, k}); \{(\sum_1^n y_i)/n\} = c_0; \{(\sum_1^n y_i x_{ji})/n\} = c_j \ (j = \overline{1, k});$$

$\{(\sum_1^n x_{1i}x_{ji})/n\} = a_{j1} \ (j = \overline{1, k}); \{(\sum_1^n x_{2i}x_{ji})/n\} = a_{j2} \ (j = \overline{1, k}); \dots \{(\sum_1^n x_{ki}x_{ji})/n\} = a_{jk} \ (j = \overline{1, k}),$  --  
можемо отримати систему рівнянь у такому, дуже зручному вигляді:

$$a_{00}b_0 + a_{01}b_1 + a_{02}b_2 + \dots + a_{0k}b_k = c_0;$$

$$a_{10}b_0 + a_{11}b_1 + a_{12}b_2 + \dots + a_{1k}b_k = c_1;$$

$$a_{20}b_0 + a_{21}b_1 + a_{22}b_2 + \dots + a_{2k}b_k = c_2;$$

. . . . .

$$a_{k0}b_0 + a_{k1}b_1 + a_{k2}b_2 + \dots + a_{kk}b_k = c_k.$$

### Перевірка значущості коефіцієнтів рівняння регресії (нуль-гіпотеза)

Після розрахунків значень коефіцієнтів здійснюється ще одна статистична перевірка – перевірка значущості коефіцієнтів рівняння регресії. Якщо буде встановлено, що якийсь коефіцієнт рівняння регресії незначущий (з обраною ймовірністю), це означає, що відповідний теоретичний коефіцієнт ряду Тейлора **дорівнює нулю** і необхідно вилучити з рівняння регресії відповідний доданок. Саме тому ця перевірка має і іншу назву – **нуль-гіпотеза**.

Можливість встановлення цієї обставини (те, що якийсь з теоретичних коефіцієнтів ряду Тейлора дорівнює нулю) в теорії планування експерименту пропонується робити з допомогою оцінок значень коефіцієнтів рівняння регресії, перевіряючи гіпотезу стосовно їх значущості. Перевірка значущості коефіцієнтів лінійної регресії виконується з допомогою **t**-критерія Стьюдента окремо для кожного коефіцієнта  $b_k \ (k = \overline{0, k})$ , де  $k$  - кількість факторів, наступним чином:

1. Знаходять значення статистичної оцінки дисперсії помилки при визначенні будь-якого коефіцієнту рівняння регресії  $S^2\{b_k\} \ (k = \overline{0, k})$ . При цьому вважається, що статистичні оцінки дисперсії помилки однакові для усіх коефіцієнтів і розраховуються по наступній формулі:  $S^2\{b_k\} = (1/Nm)S^2 \ (k = \overline{0, k})$ , де  $S^2$  – статистична оцінка дисперсії функції відгуку (див. статистичну перевірку однорідності дисперсії з використанням критерія Кохрена).

2. Розраховують значення  $t$ -критеріїв Стьюдента  $t_k$  ( $k=\overline{0,k}$ ) для кожного коефіцієнта рівняння регресії  $b_k$  ( $k=\overline{0,k}$ ), використовуючи модулі абсолютних значень коефіцієнтів  $|b_k|$  ( $k=\overline{0,k}$ ) та значення статистичної оцінки дисперсії помилки  $S^2\{b_k\}$  ( $k=\overline{0,k}$ ) по формулі:  $t_k = |b_k| / S^2\{b_k\}$  ( $k=\overline{0,k}$ ).

3. Визначають число ступенів свободи  $f_3 = f_1 f_2 = (m-1)N$ .

4. Обирають рівень значущості  $q$ .

5. По спеціальним таблицям Стьюдента знаходять критичне (табличне) значення  $t$ -критерія Стьюдента  $t_{кр}$ , яке відповідає значенням  $q$  та  $f_3$ .

6. Порівнюють по черзі значення  $t_k$  ( $k=\overline{0,k}$ ) із значенням  $t_{кр}$ .

Якщо значення  $t_k$  менше, ніж значення  $t_{кр}$ , тобто якщо  $t_k < t_{кр}$ , тоді відповідний коефіцієнт рівняння регресії  $b_k$  визнається незначущим з ймовірністю  $p$  ( $p=1-q$ ), і теоретичний коефіцієнт ряду Тейлора  $\beta_k$  визнається рівним нулю ( $\beta_k=0$ ) з ймовірністю  $p$  і відповідний доданок вилучається з рівняння регресії.

Якщо значення  $t_k$  не є меншим, ніж значення  $t_{кр}$ , тобто якщо  $t_k \geq t_{кр}$ , тоді відповідний коефіцієнт рівняння регресії  $b_k$  визнається значущим з ймовірністю  $p$ , і теоретичний коефіцієнт ряду Тейлора  $\beta_k$  визнається нерівним нулю ( $\beta_k \neq 0$ ) з ймовірністю  $p$ .

Кількість значущих доданків рівняння регресії, що залишаються після нуль-гіпотези, позначається як  $d$  і її значення далі використовується при перевірці адекватності моделі оригіналу по критерію Фішера.

Перевірка значущості коефіцієнтів нелінійних регресій також виконується з допомогою того ж  $t$ -критерія Стьюдента окремо для кожного коефіцієнта, але статистична оцінка дисперсії помилки для кожного коефіцієнта рівняння регресії (пункт 1 виконання  $t$ -критерія) визначається окремою формулою і залежить від композиційного плану, який використовується.

### Перевірка адекватності моделі оригіналу

Після розрахунків коефіцієнтів рівняння регресії і корегування рівняння регресії у відповідності з нуль-гіпотезою необхідно переконатися в її адекватності об'єкту, що досліджувався. Рівняння регресії – це побудована статистична математична модель (точніше кажучи – це статистичний математичний опис об'єкту) і

необхідно переконатись в її адекватності оригіналу, тобто усім отриманим експериментальним даним  $y_i$  ( $i=\overline{1,n}$ ).

Гіпотеза адекватності базується на оцінюванні степені наближення розрахункових значень вихідної змінної  $\check{y}_i$  ( $i=\overline{1,n}$ ) до отриманих експериментальних значень вихідної змінної  $y_i$  ( $i=\overline{1,n}$ ). Але практично аналізується сума квадратів відхилень експериментальних значень від розрахункових значень в усіх точках факторного простору, а не по усьому об'єму вибірки  $n$ , тому що кількість розрахункових значень вихідної змінної –  $N$ , а не  $n$ , тобто  $\check{y}_j$  ( $j=\overline{1,N}$ ).

В зв'язку з тим, що необхідно використовувати також  $N$  експериментальних значень вихідної змінної, для розрахунків обираються середні значення (математичні сподівання) для кожної точки факторного простору  $\bar{y}_j$  ( $j=\overline{1,N}$ ).

Для перевірки адекватності моделі оригіналу (для оцінки відхилень) в теорії планування експерименту використовується критерій Фішера:

1. Знаходять значення статистичну оцінку дисперсії адекватності  $S^2_{ад}$ , використовуючи формулу  $S^2_{ад} = \{m/(N-d)\} \sum_1^N (\bar{y}_j - y_j)^2$ .

2. Розраховують значення критерія Фішера  $F$  («дисперсійне відношення»):  $F = S^2_{ад} / S^2$ .

3. Визначають число ступенів свободи  $f_3$  та  $f_4$ :  $f_3 = f_1 f_2 = (m-1)N$ ;  $f_4 = N-d$ .

4. Обирають рівень значущості  $q$ .

5. По спеціальним таблицям Фішера знаходять критичне (табличне) значення критерія Фішера  $F_{кр}$ , яке відповідає значенням  $q$ ,  $f_3$  та  $f_4$ .

6. Порівнюють значення  $F$  із значенням  $F_{кр}$ .

Якщо значення  $F$  більше, ніж значення  $F_{кр}$ , тобто якщо  $F > F_{кр}$ , тоді вважається, що статистична математична модель, яка пропонується (отримане рівняння регресії), не адекватна оригіналу (усім масиву експериментальних даних) з ймовірністю  $p$ .

В такому випадку необхідно:

- 1) збільшити кількість доданків рівняння регресії;
- 2) змінити план експерименту;
- 3) збільшити кількість точок факторного простору;
- 4) провести додаткові експерименти в нових точках факторного простору;

- 5) перевірити однорідність дисперсії, збільшуючи кількість повторів  $m$  до того часу, доти не отримаємо підтвердження однорідності дисперсії з ймовірністю  $p$ ;
- 6) здійснити розрахунок коефіцієнтів рівняння регресії;
- 7) перевірити значущість коефіцієнтів і здійснити корегування рівняння регресії;
- 8) знову перевірити адекватність оригіналу моделі ... і так далі до того часу, доти не буде підтверджена адекватність моделі оригіналу.

**Якщо  $F \leq F_{кр}$ , тоді вважається, що ця статистична математична модель (отримане рівняння регресії), адекватна оригіналу (усім масиву експериментальних даних) з ймовірністю  $p$ . Таким чином, мета досягнута – з мінімальними витратами часу і коштів отримали статистичну математичну модель!!!**

## РЕКОМЕНДОВАНА ЛІТЕРАТУРА

### Базова

1. Теорія планування експерименту (20402040) Селіванов В.Л. Електронний конспект лекцій. Порев В.М. Електронний конспект практик. КПІ WiKi, 2016.
2. Л.А.Назаренко Планування і обробка результатів експерименту. Конспект лекцій. Харків. ХНУНГ ім.О.М.Бекетова,2018
3. Ю.С.Гришук Основи наукових досліджень Навчальний посібник.Харків.НТУ «ХП», 2008.
4. Крушельницька О.В. Методологія та організація наукових досліджень. Навч. Посібник. К. Кондор, 2003.
5. Основи наукових досліджень. Навчальний посібник. Марцин В.С., Міщенко Н.Г., Дані-ленко О.А., Львів, Рамус-Поліграф,2002.

### Допоміжна

6. Григор'єв Ю.Д. Методи оптимального планування експерименту: Лінійні моделі. Навчальний посібник. СПБ. Видавництво «Лань»,2015.
7. К.В.Губкін, А.І.Яковлев ТПЕ в енергетиці. Курс лекцій для вищих навчальних закладів. Київ. "Міленіум", 2009.
8. Волков В.В. Планування експерименту та обробка результатів за допомогою програмного засобу EOSupport. Вісник ДДТУ, Т.8., с.120-126, 2008.
9. С.М.Лапич Теорія планування експерименту. Виконання розрахунково-графічної роботи. Київ. КПІ ім. Ігоря Сікорського.2020

### Інформаційні ресурси

10. [http://wiki.kpi.ua/index.php/Теорія\\_планування\\_експерименту\\_\(20402040\)](http://wiki.kpi.ua/index.php/Теорія_планування_експерименту_(20402040))