До введення у дію Державної Фармакопеї України

Гризодуб О. І., Підпружников Ю. В., Леонтьєв Д. А., Петрус В. В., Іванов Л. В.
Державне підприємство «Державний науковий фармакопейний центр якості лікарських засобів», Харків, Україна
Національний фармацевтичний університет, Харків, Україна
ПАТ НВЦ «Борщагівський ХФЗ», Київ, Україна
НТУ «Харківський політехнічний інститут», Харків, Україна

ПОЯСНЮВАЛЬНА ЗАПИСКА
до проекту розділу 12 «Застосування методу найменших квадратів у регресійному аналізі» загальної статті ДФУ *5.3.N.1 «Статистичний аналіз результатів хімічного експерименту^N»*

Необхідність розробки цього розділу пов'язана з переходом вітчизняних підприємств на вимоги GMP. Водночас проводяться великі дослідження, пов'язані з валідацією аналітичних методик і технологічного процесу. У цих дослідженнях широко застосовується кореляційний метод, заснований на методі найменших квадратів (МНК).

Метою цього розділу є короткий теоретичний виклад застосування МНК з єдиного матричного погляду, який охоплює всі основні практично важливі для фармації аспекти застосування МНК: загальне формулювання, лінійний МНК, лінійний зважений МНК, нелінійний МНК і нелінійний зважений МНК.

Для ілюстрації викладеного теоретичного матеріалу наведено 4 приклади застосування МНК на реальних об'єктах:

— лінійний двопараметричний МНК (перевірка лінійності під час проведення валідації аналітичної методики),
— лінійний трипараметричний МНК (перевірка значущості впливу технологічних факторів під час відпрацювання технології),
— лінійний зважений двопараметричний МНК (перевірка лінійності в припущенні рівноточності відносної невизначеності ординат),
— нелінійний двопараметричний МНК (дослідження кінетики розчинення субстанції в середовищі розчинення).

Для зручності користувачів усі приклади наведені в такому вигляді, який дозволяє легко реалізувати їх в *Excel*®, що забезпечує простежуваність. Ці приклади також можуть бути використані для перевірки розроблених програм.

Ще однією важливою метою цього розділу є встановлення фармакопейного стандарту представлення результатів обробки за МНК: коефіцієнти рівняння, їх стандартні відхилення, залишкове стандартне відхилення, коефіцієнт кореляції.

Включення цього розділу в ДФУ дозволить полегшити користувачам ДФУ застосування кореляційного аналізу у своїх дослідженнях.

ПРОЕКТ

12. ЗАСТОСУВАННЯ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ У РЕГРЕСІЙНОМУ АНАЛІЗІ

12.1. Загальне формулювання методу найменших квадратів

12.2. Проблема нерівноточності значень ординат

12.2.1. Допущення рівноточності ординат

12.2.2. Загальний випадок

12.2.3. Допущення рівності відносних стандартних відхилень ординат

12.3. Пряме числове розв'язання співвідношень (12.3) або (12.7)

12.4. Використання часткових похідних регресійної функції за параметрами

12.5. Лінійний метод найменших квадратів (ЛМНК)

12.6. Лінійний зважений метод найменших квадратів (ЛЗМНК)

12.7. Нелінійний метод найменших квадратів (НМНК)

12.8. Нелінійний зважений метод найменших квадратів (НЗМНК)

12.9. Приклади

12.9.1. Валідація методики кількісного визначення таблеток амброксолу: розрахунок лінійної залежності методом найменших квадратів

12.9.2. Дослідження значимості впливу швидкості і сили пресування на однорідність маси і однорідність дозування одиниць ядер таблеток дезлоратадину

12.9.3. Лінійний зважений метод найменших квадратів: побудова калібрувальної прямої в припущенні рівності відносних невизначеностей ординат

12.9.4. Застосування нелінійного методу найменших квадратів для опису кінетики розчинення зразків субстанції фенсукциналу з різним розміром частинок

12. ЗАСТОСУВАННЯ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ У РЕГРЕСІЙНОМУ АНАЛІЗІ

Під час проведення фармацевтичної розробки лікарських засобів і розробки методик аналізу часто виникає необхідність математичного опису залежності певної залежної змінної y (яку називають «функцією відгуку» або «цільовою функцією») від k незалежних змінних (факторів) $x_1, x_2 \ldots x_k$, значення яких визначаються умовами проведення експерименту.

Якщо провести n дослідів, експериментальну залежність значень функції відгуку y_i (ординати) від значень незалежних змінних $x_{1j}, x_{2j} \ldots x_{kj}$ можна представити як Табл. 12.1.

Таблиця 12.1

Типове представлення експериментальних даних

Номер досліду i	Значення незалежних змінних x_j для різних дослідів $i = 1, 2\ldots n$				Значення залежної змінної y (ординати) у різних дослідах i
1	x_{11}	x_{21}	…	x_{k1}	y_1
2	x_{12}	x_{22}	…	x_{k2}	y_2
…					
n	x_{1n}	x_{2n}	…	x_{kn}	y_n

Як видно, за результатами експерименту ми отримуємо функцію y, значення якої відомі тільки в деяких дискретних точках (дослідах). Для її математичного опису необхідно розв'язати завдання вибору аналітичної функції, яка буде у найкращий спосіб описувати функцію, задану в Табл. 12.1. Загалом таке завдання не має вирішення, оскільки неможливо тільки на підставі експерименту визначити вигляд математичної функції. Однак можна розв'язати більш часткове завдання підбору параметрів, які описують деяку, наперед задану, функціональну залежність. Цим займається регресійний аналіз.

12.1. ЗАГАЛЬНЕ ФОРМУЛЮВАННЯ МЕТОДУ НАЙМЕНШИХ КВАДРАТІВ

Завдання регресійного аналізу можна сформулювати так. Для запропонованої математичної моделі загального вигляду

$$y^{calc} = f(x_1, x_2, \ldots x_k; b_1 \ldots b_m) \qquad (12.1)$$

необхідно підібрати такі оптимальні значення m постійних числових параметрів (коефіцієнтів регресії) $b_1 \ldots b_m$, щоб експериментально виміряні величини y_i і розраховані за рівнянням (12.1) значення функції y_i^{calc} для кожного рядка Табл. 12.1 відрізнялися якнайменше.

Як залежні змінні, незалежні змінні і числові параметри можуть бути різні величини. Наприклад, у разі багатохвильової спектрофотометрії залежними змінними є оптичні густини за аналітичних довжин хвиль, незалежними змінними — показники поглинання аналізованих компонентів за цих довжин хвиль, а числовими параметрами, які необхідно знайти, — концентрації аналізованих компонентів. Під час опису кривих вивільнення залежними змінними є концентрації аналізованих речовин у розчині (або ступінь вивільнення у відсотках до максимальної концентрації), незалежною змінною — час вивільнення, а числовими параметрами, які треба знайти — коефіцієнти математичної моделі тощо. Визначимо залишкову дисперсію в такий спосіб:

$$s_0^2 = \frac{1}{n-m} \times \sum_{i=1}^{i=n} \left(y_i - y_i^{calc}\right)^2 = \\ = \frac{1}{n-m} \times \sum_{i,j=1}^{n,k} \left[y_i - f(x_{ij}, b_1 \ldots b_m)\right]^2. \qquad (12.2)$$

Від звичайної дисперсії вона відрізняється тим, що використовує «плаваюче середнє»: як середнє значення для кожного y_i використовується y_i^{calc}, розраховане за рівнянням (12.1) для заданих значень параметрів $b_1 \ldots b_m$. Оптимальність коефіцієнтів регресії $b_1 \ldots b_m$ відповідає мінімуму залишкової дисперсії зі співвідношення (12.2), тобто:

$$s_0^2 = \frac{1}{n-m} \times \sum_{i=1}^{i=n} \left(y_i - y_i^{calc}\right)^2 \to \min. \qquad (12.3)$$

Загалом кількість параметрів m може відрізнятися від кількості незалежних змінних k.

Співвідношення (12.2-12.3) є загальним формулюванням *методу найменших квадратів* (МНК).

Враховуючи, що коли здійснюється варіювання параметрів $b_1 \ldots b_m$, число ступенів свободи $(n - m)$ є постійною величиною, його можна виключити з розгляду. Тоді оптимальність параметрів $b_1 \ldots b_m$ відповідає мінімуму залишкової суми квадратів (*RSS*):

$$RSS = \sum_{i=1}^{i=n} \left(y_i - y_i^{calc} \right)^2 = \\ = \sum_{i,j=1}^{n,k} \left[y_i - f(x_{ij}, b_1 \ldots b_m) \right]^2 \to \min. \quad (12.4)$$

Вираз (12.4) дав назву *методу найменших квадратів*, оскільки мінімізується сума квадратів відхилень експериментальних точок від запропонованої математичної моделі (12.1).

Слід зазначити, що під час опису регресії за математичною моделлю (12.1) деякі параметри $b_1 \ldots b_m$, отримані за допомогою МНК, можуть бути статистично незначущими і тому мають бути відкинуті. Розрахунки за допомогою МНК мають бути повторені зі зменшеною кількістю параметрів. Водночас *RSS* може зменшитися не внаслідок покращення опису експериментальних даних, а просто через зменшення кількості параметрів. Для уникнення цього краще відразу використовувати більш загальні співвідношення (12.2-12.3), тим більше, що використання (12.4) не має ніяких переваг перед (12.2-12.3).

Для оптимізованих значень параметрів $b_1 \ldots b_m$ залишкове стандартне відхилення s_0 розраховують зі співвідношення (12.2), а загальний індекс (коефіцієнт) кореляції R_c (див. рівняння (7.3а)) зі співвідношення:

$$R_c = \sqrt{1 - \frac{s_0^2}{s_y^2}}. \quad (12.5)$$

Тут s_y — стандартне відхилення величин y навколо свого середнього значення.

12.2. ПРОБЛЕМА НЕРІВНОТОЧНОСТІ ЗНАЧЕНЬ ОРДИНАТ

12.2.1. Допущення рівноточності ординат

Необхідно зазначити одну важливу обставину (що має загальний характер для всіх описів експериментальних даних за теоретичними рівняннями), яка стає актуальною в тому разі, коли нас цікавить не тільки сам факт відповідності експериментальних даних теоретичним рівнянням, але й параметри цих рівнянь.

Йдеться про те, що розв'язання співвідношення (12.3) передбачає *рівноточність значень ординат* y_i (і, відповідно, відхилень від теоретичної кривої y_i - y_i^{calc}), тобто рівності абсолютних стандартних відхилень для значень ординат y_i:

$$s(y_i) = const, \; i = 1 \ldots n. \quad (12.6)$$

Це допущення є коректним для достатньо вузького інтервалу величин y_i (зокрема, під час валідації методик кількісного визначення Фармакопея вимагає доказу лінійності аналітичного сигналу від концентрації в інтервалі 80-120 % від номінального значення), однак з розширенням цього інтервалу воно стає все менш і менш коректним.

Зазначимо, що «лінійність» і «рівноточність» — поняття різні. Аналітичний сигнал (y_i) може лінійно, але нерівноточно залежати від концентрації в межах аналітичного діапазону (можливий і зворотній випадок: залежність рівноточна, але нелінійна — див. нижче *нелінійний МНК*). Саме ця обставина ускладнює, зокрема, валідацію методик кількісного визначення в широкому інтервалі концентрацій.

Проблема нерівноточності стає ще більш актуальною в тому разі, коли як ординати використовують не значення самої функції відгуку (для яких часто можна припустити рівноточність), а якісь їх трансформовані величини (логарифми, зворотні величини тощо), які дозволяють отримати лінійні (за параметрами) залежності і застосувати ЛМНК. У такому разі для трансформованих значень функції відгуку допущення рівноточності (12.6) часто є некоректним. Тут доцільно використати *нелінійний МНК* для опису вихідної функції відгуку. Альтернативою може бути використання статистики функції декількох випадкових змінних (див. розділ 9) для знаходження взаємозв'язку вихідних і перетворених стандартних відхилень $s(y_i)$. Однак такий підхід за складністю мало чим відрізняється від *нелінійного МНК*, поступаючись останньому в коректності.

12.2.2. Загальний випадок

Загальним підходом до розв'язання проблеми нерівноточності значень ординат є введення в співвідношення (12.3) вагових множників, які ураховують цю нерівноточність. Як ваги w_i

логічно використовувати величини, зворотні дисперсіям $s^2(y_i)$ величин y_i. У цьому разі, враховуючи (12.2-12.3), загальне формулювання *зваженого методу найменших квадратів* (ЗМНК) має вигляд:

$$s_0^2(w) = \frac{1}{n-m} \times \sum_{i=1}^{i=n} [y_i - y_i^{calc}]^2 \times w_i = \\ = \frac{1}{n-m} \times \sum_{i,j=1}^{n,k} [y_i - f(x_{ij}, b_1 ... b_m)]^2 \times w_i \to \min, \quad (12.7)$$

$$w_i = \frac{1}{s^2(y_i)}, \quad i = 1...n. \quad (12.8)$$

Для оптимізованих значень параметрів $b_1 ... b_m$ залишкове стандартне відхилення $s_0(w)$ розраховують за співвідношеннями (12.7-12.8), а загальний коефіцієнт кореляції $R_c(w)$ — за співвідношенням:

$$R_c(w) = \sqrt{1 - \frac{s_0^2(w)}{s_y^2}}, \quad (12.9)$$

які є узагальненнями співвідношень (12.2) і (12.5).

Слід зазначити, що знайдені за допомогою ЗМНК (12.7-12.9) параметри $b_1 ... b_m$ і метрологічні характеристики регресії (12.1) значно складніше інтерпретувати, ніж у разі звичайного незваженого методу найменших квадратів (12.3). Із співвідношень (12.7-12.9) неважко бачити, що $s_0(w)$ являє собою середньоквадратичний критерій Стьюдента, розрахований із критеріїв Стьюдента для кожної точки, що може бути використано під час оцінки рівняння регресії (12.1). Водночас коефіцієнт кореляції $R_c(w)$ і розраховані за інформаційно-коваріаційною матрицею (див. нижче) стандартні відхилення параметрів $b_1 ... b_m$ значно складніше інтерпретувати. Це обмежує застосування співвідношень (12.7-12.9) на практиці.

Також необхідно підкреслити, що у разі використання не генеральних (які зазвичай невідомі), а вибіркових дисперсій $s^2(y_i)$, для коректного застосування зваженого методу найменших квадратів необхідно перевірити гіпотезу нерівності дисперсій за критерієм Бартлета (див. розділ 1.3.2) або Кокрена (див. розділ 1.3.3 — у разі якщо всі дисперсії мають однакове число ступенів свободи). Якщо критерії нерівності дисперсій не виконуються (а для вибіркових дисперсій із малим числом ступенів свободи гіпотезу нерівності дисперсій дуже важко обґрунтувати), то застосування ЗМНК не є статистично коректним, і слід використовувати звичайний (незважений) МНК.

12.2.3. Допущення рівності відносних стандартних відхилень ординат

Важливим окремим випадком ЗМНК є використання допущення рівності відносних стандартних відхилень ординат, тобто:

$$s_r(y_i) = s(y_i) / y_i = const, \; i = 1...n. \quad (12.10)$$

Допущення (12.10) відповідає значенням ваг

$$w_i = \frac{1}{y_i^2}, \quad i = 1...n. \quad (12.11)$$

у співвідношенні (12.8).

Допущення (12.10) зазвичай виконується в ширшому інтервалі, ніж допущення (12.4) рівності абсолютних стандартних відхилень, і його застосовують під час побудови калібрувальних прямих у ЛВМНК (див. розділ 12.7.3).

Для знаходження мінімуму s_0 (12.3) або $s_0(w)$ (12.7) можуть застосовуватися різні підходи, які можна поділити на числові й аналітичні.

Аналітичні підходи — це знаходження параметрів співвідношення (12.3) або (12.7) як функції, тобто точне розв'язання. Таке точне розв'язання можна отримати для *лінійного* (за параметрами) МНК (ЛМНК).

Загалом визначення параметрів співвідношень (12.3) або (12.7) може проводитися числовими методами оптимізації (мінімізації). У такому разі говорять про *нелінійний* МНК (НМНК).

12.3. ПРЯМЕ ЧИСЛОВЕ ВИЗНАЧЕННЯ ПАРАМЕТРІВ СПІВВІДНОШЕНЬ (12.3) АБО (12.7)

Одним з найпростіших варіантів НМНК є *метод самоузгодженого перебору* (МСП). МСП може бути застосований для будь-якого виду функції (12.1) — як лінійних, так і нелінійних за параметрами $b_1 ... b_m$. Це дозволяє розглядати його окремо від інших варіантів НМНК.

Для реалізації МСП можна використовувати, наприклад, такий алгоритм:

1. Для кожного з m параметрів $b_1 ... b_m$ задають вихідні значення $b_1(0) ... b_m(0)$.
2. Для кожного з m параметрів $b_1 ... b_m$ задають вихідні кроки ітерацій $\Delta b_1(0) ... \Delta b_m(0)$.
3. Для кожного з m параметрів $b_1 ... b_m$ задають діапазон — три вихідні точки: $[b_1(0) - \Delta b_1(0)]$, $b_1(0)$, $[b_1(0) + \Delta b_1(0)]$; ...$[b_m(0) - \Delta b_m(0)]$, $b_m(0)$, $[b_m(0) + \Delta b_m(0)]$. Загальна кількість комбінацій точок за всіма параметрами становить 3^m.

4. Для всіх 3^m комбінацій заданих значень параметрів b_1 ... b_m розраховують значення функції (12.1) і розраховують значення функціоналів (12.3) або (12.7).
5. Вибирають комбінацію значень параметрів b_1 ... b_m, яка відповідає мінімуму функціоналів (12.3) або (12.7). Вибрані в такий спосіб значення беруть за нові вихідні значення й ітераційний процес повторюють.
6. Якщо знайдене значення конкретного параметра b_i протягом трьох ітерацій змінюється (росте або зменшується) в одному напрямку, то його крок ітерації Δb_i збільшують удвічі.
7. Якщо знайдене значення конкретного параметра b_i протягом трьох ітерацій не змінюється, то його крок ітерації Δb_i зменшують удвічі.
8. Ітераційний процес завершують, коли значення функціоналів (12.3) або (12.7) на останній ітерації відрізняється від попереднього значення менше наперед заданої величини.
9. Залишкове стандартне відхилення s_0 і загальний коефіцієнт кореляції R_c розраховують для оптимізованих значень параметрів b_1 ...b_m за співвідношеннями (12.2), (12.5) або (12.7-12.9).
10. Стандартні відхилення для оптимізованих параметрів b_1 ... b_m розраховують на підставі інформаційно-коваріаційної матриці (див. нижче).

Переваги МСП

1. Він може бути застосований до будь-яких функцій (12.1) — як лінійних, так і нелінійних відносно параметрів b_1 ... b_m.
2. Для знаходження параметрів b_1 ...b_m не вимагається розрахунку похідних функції (12.1) за параметрами b_1 ... b_m на кожному кроці ітерацій. Не потрібно також і обернення матриць. Цей аспект може бути важливим, наприклад, якщо функція (12.1) являє собою лінійну комбінацію експонент. Мінімізація функціоналів (12.3) або (12.7) у цьому разі належить до некоректних завдань.

Недоліки МСП

1. МСП є приблизним методом, на відміну від лінійного методу найменших квадратів, який дає аналітичне розв'язання.
2. Оптимізація параметрів b_1 ... b_m залежить від вихідних заданих значень $b_1(0)$ ... $b_m(0)$. Немає впевненості, що знайдений мінімум функціоналів (12.3) або (12.7) є глобальним, а не локальним. Тому знаходження оптимізованих значень b_1 ... b_m бажано повторити для різних вихідних значень $b_1(0)$ ... $b_m(0)$.
3. Застосування МСП вимагає розробки програм. У разі ЛМНК, а також НМНК для низки випадків можуть бути використані звичайні розрахунки в *Excel*®.
4. Для знаходження стандартних відхилень параметрів $b_1(0)$ ... $b_m(0)$ однаково необхідно отримання інформаційно-коваріаційної матриці, яку розраховують на підставі числових значень похідний функції (12.1) за параметрами b_1 ... b_m.

12.4. ВИКОРИСТАННЯ ЧАСТКОВИХ ПОХІДНИХ РЕГРЕСІЙНОЇ ФУНКЦІЇ ЗА ПАРАМЕТРАМИ

Загальним підходом до знаходження мінімуму функціоналів (12.2-12.3) або (12.7) є використання часткових похідних за параметрами. У точці мінімуму всі частинні похідні виразів (12.3) або (12.7) за кожним з m параметрів b_1 ... b_m мають перетворитися на нуль.

У разі незваженого МНК, проводячи відповідне диференціювання виразу (12.2), отримаємо систему з m рівнянь:

$$\sum_{i,j=1}^{n,k}\left[y_i - f(x_{ij}, b_1 ... b_m)\right] \times \frac{df(x_{ij}, b_1 ... b_m)}{db_g} = 0, \quad g = 1,...m. \qquad (12.12)$$

Математичні методи розв'язання співвідношення (12.3) можна поділити на *лінійний метод найменших квадратів* (ЛМНК) і *нелінійний метод найменших квадратів* (НМНК).

ЛМНК дозволяє отримати точне розв'язання рівняння (12.3), і його застосовують у тому разі, коли функція (12.1) *лінійно залежить від параметрів* b_0, b_1 ... b_m. Водночас сама функція (12.1) може бути і нелінійною до незалежних змінних x_j (наприклад, поліном $f = \sum b_j \times x^j$). Важливо, щоб змінні були лінійно незалежними одна від одної.

НМНК застосовують у тому разі, коли функція (12.1) *нелінійно залежить від параметрів* b_1 ... b_m. Водночас сама функція (12.1) може бути як лінійною, так і нелінійною до незалежних змінних x_j.

12.5. ЛІНІЙНИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ (ЛМНК)

Нехай регресійна залежність (12.1) є лінійною відносно параметрів b_1 ... b_m:

$$y_i^{calc} = \sum_{j=1}^{j=m} x_{ij} \times b_j, \quad i = 1...n \geq m. \qquad (12.13)$$

У цьому разі число параметрів співпадає з числом незалежних змінних. Якщо є вільний член (b_1), то в співвідношенні (12.13) визначають $x_{i1} = 1$.

Завдання ЛМНК у цьому разі зводиться до розв'язання перевизначеної системи лінійних рівнянь (12.13).

Співвідношення (12.14) зручно записати в матричному вигляді:

$$y^{calc} = x \times b. \qquad (12.14)$$

Тут

y^{calc} — вектор-стовпець розрахованих значень функції (12.1) розміром $n \times 1$;

x — матриця незалежних змінних розміром $m \times n$ (див. Табл. 12.1);

b — вектор-стовпець параметрів розміром $m \times 1$.

Неважко бачити, що

$$x = \frac{\partial y^{calc}}{\partial b}, \qquad (12.15)$$

тобто в рівнянні (12.14) ЛМНК матриця незалежних змінних x являє собою матрицю часткових похідних функції y^{calc} за параметрами b. Зауважимо, що у разі ЛМНК матриця x не залежить від параметрів b.

Тоді рівняння (12.2) у матричному вигляді буде таким:

$$s_0^2 = \frac{1}{n-m} \times (y - x \times b)^T \times (y - x \times b). \qquad (12.16)$$

Тут

y — вектор-стовпець експериментальних значень функції відгуку розміром $n \times 1$;

T — символ транспонування матриці.

Проводячи диференціювання цього співвідношення за вектором параметрів b і прирівнюючи отримані похідні до нуля, отримаємо таку систему рівнянь у матричній формі:

$$(x^T \times x) \times b = x^T \times y. \qquad (12.17)$$

Звідси отримаємо оптимізований вектор параметрів b:

$$b = [ICM \times x^T] \times y = RM \times y. \qquad (12.18)$$

Тут

RM — розрахункова;

ICM — інформаційно-коваріаційна матриця вигляду:

ЛМНК: $$ICM = (x^T \times x)^{-1}. \qquad (12.19)$$

Залишкове стандартне відхилення s_0 розраховують за формулою (12.3), коефіцієнт кореляції R_c — за формулою (12.5).

Стандартні відхилення параметрів b_i розраховують із діагональних елементів інформаційно-коваріаційної матриці ICM:

$$s_{bi} = s_0 \times \sqrt{ICM_{ii}}. \qquad (12.20)$$

Лінійний метод найменших квадратів (ЛМНК) є основою розрахунків усіх лінійних залежностей під час валідації кількісних методик. Він широко застосовується для опису поліноміальних регресій, а також у багатокомпонентній спектрофотометрії. Зокрема, всі рівняння МНК, МОРК і ММНК, які описані в національній частині загальної статті *2.2.25 «Абсорбційна спектрофотометрія в ультрафіолетовій і видимій областях»*, отримані з використанням співвідношень (12.18-12.19).

Приклади застосування ЛМНК під час валідації методик кількісного визначення й у двофакторному кореляційному аналізі наведені в розділах 12.9.1 і 12.9.2.

12.6. ЛІНІЙНИЙ ЗВАЖЕНИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ (ЛЗМНК)

Рівняння ЛЗМНК відрізняються від ЛМНК додатковим введенням діагональної матриці вагових коефіцієнтів w з елементами (12.8) на діагоналі:

ЛЗМНК: $$w = diag\{w_1, w_2 \ldots w_n\}. \qquad (12.21)$$

Як вагові коефіцієнти можуть використовуватися також і квадрати зворотних значень самих ординат — у разі допущення рівності відносних дисперсій ординат — див. розділ 12.2.3.

Інформаційно-коваріаційна матриця (12.19) у цьому разі буде такою:

ЛЗМНК: $$ICM = (x^T \times w \times x)^{-1}. \qquad (12.22)$$

Оптимізований вектор параметрів (12.18) переходить у співвідношення:

ЛЗМНК:

$$b = [ICM_w \times x^T \times w] \times y = RM_w \times y. \qquad (12.23)$$

Тут RM_w — зважена розрахункова матриця.

Зважені залишкове стандартне відхилення $s_0(w)$ і загальний коефіцієнт кореляції $R_c(w)$ знаходять із співвідношень (12.7-12.9).

Для розрахунку незважених залишкового стандартного відхилення s_0 і загального коефіцієнта кореляції R_c використовують співвідношення (12.2) і (12.5).

Приклад застосування ЛЗМНК для побудови калібрувальної прямої за допущенням рівно-

сті відносних стандартних відхилень ординат (12.10-12.11) наведений у розділі 12.9.3.

12.7. НЕЛІНІЙНИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ (НМНК)

Розкладемо функцію y_i^{calc} (12.1) в ряд Тейлора з точністю до першого члена в точці y_i:

$$y_i^{calc} = \\ = y_i + \sum_{j=1}^{j=m} \frac{\partial f(x_{1i}...x_{ki};b_{10}...b_{m0})}{\partial b_j} \times (b_j - b_{j0}) = \\ = y_i + \sum_{j=1}^{j=m} E_{ij} \times \Delta b_j. \quad (12.24)$$

Тут E_{ij} — матриця часткових похідних за параметрами в точці $b_j = b_{j0}$:

$$E_{ij} = \frac{\partial f(x_{1i}...x_{ki};b_{10}...b_{m0})}{\partial b_j}. \quad (12.25)$$

Співвідношення (12.24), з урахуванням (12.25), можна записати як:

$$y_i^{calc} - y_i = \Delta y_i = \sum_{j=1}^{j=m} E_{ij} \times \Delta b_j \quad (12.26)$$

або у матричному вигляді:

$$\Delta y = E \times \Delta b. \quad (12.27)$$

Неважко бачити, що співвідношення (12.27) за формою співпадає зі співвідношенням (12.14) звичайного ЛМНК. Це дозволяє використати співвідношення (12.13-12.20) звичайного ЛМНК для отримання оптимізованого вектора b в НМНК. Однак, на відміну від матриці x ЛМНК, у разі НМНК матриця часткових E залежить від параметрів b_j, тому процес отримання вектора b в НМНК є ітераційним.

Подібно (12.19) ЛМНК, інформаційно-коваріаційну матрицю НМНК знаходять зі співвідношення:

НМНК: $$ICM = (E^T \times E)^{-1}. \quad (12.28)$$

Подібно (12.18) ЛМНК, розрахункову матрицю і зміну вектора параметрів Δb в НМНК знаходять зі співвідношень:

НМНК: $$RM = ICM \times E^T. \quad (12.29)$$

НМНК: $$\Delta b = RM \times \Delta y. \quad (12.30)$$

Залишкове стандартне відхилення s_0 розраховують за формулою:

НМНК: $$s_o = \sqrt{\frac{1}{(n-m)} \sum_{i=1}^{i=m} \Delta y_i^2}. \quad (12.31)$$

Стандартні відхилення параметрів b_i розраховують із діагональних елементів інформаційно-коваріаційної матриці ICM (12.28) і значення s_0 зі співвідношення (12.31):

$$s_{bi} = s_0 \times \sqrt{ICM_{ii}}. \quad (12.32)$$

Коефіцієнт кореляції R_c розраховують за формулою (12.5).

Алгоритм ітераційного процесу

1. Задають точність оптимізації. Оскільки оптимізація проводиться за мінімумом залишкового стандартного відхилення s_0 зі співвідношення (12.3), то задають мінімальну зміну $min\ \Delta s_0$ на останньому етапі ітерації, після якого подальша оптимізація припиняється, тобто має виконуватися нерівність:
$$\Delta s_0 < \min s_0. \quad (12.33)$$
2. Задають вихідні наближення b_{j0} і для них розраховують за (12.13-12.14) значення y_{i0}^{calc} і вихідний вектор y_0^{calc}, а також матрицю частинних похідних E_{ij} за співвідношенням (12.25).
3. За співвідношенням (12.28-12.30) розраховують оптимізовані значення b_j (перша ітерація).
4. За співвідношенням (12.14) розраховують перше наближення для вектора y^{calc}.
5. Розраховують першу зміну вектора $\Delta y = y^{calc} - y$.
6. Використовуючи Δy, за співвідношенням (12.30) розраховують зміну вектора Δb.
7. За співвідношенням (12.31) розраховують залишкове стандартне відхилення s_0 для першої ітерації.
8. За співвідношенням (12.5) розраховують коефіцієнт кореляції R_c для першої ітерації.
9. За співвідношеннями (12.28) і (12.32) розраховують стандартні відхилення параметрів s_{bi} для першої ітерації.
10. Розраховують нове наближення для вектора $b = b_0 - \Delta b$.

Ітераційну процедуру повторюють із новим наближенням вектора b. Водночас розраховують зміну Δs_0, порівнюючи з попередньою ітерацією. Процедуру ітерації продовжують, поки не буде виконуватися нерівність (12.33). Вектор b, який відповідає цьому, беруть за оптимальний. Для нього розраховують залишкове стандартне відхилення s_0 і загальний коефіцієнт кореляції R_c за співвідношеннями (12.31) і (12.5).

Приклад застосування НМНК наведений у Прикладі 12.9.4.

12.8. НЕЛІНІЙНИЙ ЗВАЖЕНИЙ МЕТОД НАЙМЕНШИХ КВАДРАТІВ (НЗМНК)

Процедура оптимізації така сама, як і для НМНК, але замість співвідношень (12.28), (12.29) і (12.31) використовують співвідношення:

НЗМНК: $ICM = (E^T \times w \times E)^{-1}$. (12.34)

НЗМНК: $RM = ICM \times E^T \times w$. (12.35)

Залишкове стандартне відхилення s_0 розраховують за співвідношенням:

НЗМНК: $$s_o = \sqrt{\frac{1}{(n-m)} \sum_{i=1}^{i=m} \Delta y_i^2 \times w_i}\ .$$ (12.36)

12.9. ПРИКЛАДИ

12.9.1. Валідація методики кількісного визначення таблеток амброксолу: розрахунок лінійної залежності методом найменших квадратів

Під час валідації за стандартизованою процедурою (див. розділ *5.3.N.2 «Валідація аналітичних методик і випробувань»*) спектрофотометричної методики кількісного визначення таблеток амброксолу були отримані такі вихідні значення нормалізованих координат X і Y (Табл. 12.П-1).

Для підтвердження лінійності необхідно отримати параметри і метрологічні характеристики лінійної залежності:

$$Y = b_1 + b_2 \times X. \qquad (12.П\text{-}1)$$

Тут X і Y — вектори-стовпці з Табл. 12.П-1. У цьому разі число дослідів $n = 9$, число параметрів прямої (12.П-1.1) $m = 2$, число ступенів свободи $n - m = 7$.

Отримання параметрів і метрологічних характеристик прямої (12.П-1) можна проводити за формулами, наведеними в розділі *7 «Розрахунок і статистична оцінка параметрів лінійної залежності»*. Однак, як приклад, продемонструємо застосування більш загального матричного підходу, описаного в розділі 12.5.

У розрахунках припускається рівноточність ординат, тобто допущення (12.6). Наведена далі схема може бути легко реалізована в *Excel*®, що забезпечує її простежуваність.

Матриці x і x^T

Ураховуючи, що в лінійній залежності (12.П-1) є вільний член b_1, ці матриці мають у цьому разі вигляд:

$$x = \begin{bmatrix} 1 & 62.51 \\ 1 & 71.77 \\ 1 & 80.29 \\ 1 & 85.92 \\ 1 & 92.41 \\ 1 & 106.62 \\ 1 & 111.68 \\ 1 & 121.29 \\ 1 & 133.66 \end{bmatrix}.$$

Матрицю x^T див. на Рис. 12П-1.

Таблиця 12.П-1

Вихідні дані і результати розрахунків за прямою $Y = b1 + b2 \times X$

N досліду	Вихідні дані		Розрахунки за прямою	
	X, %	*Y*, %	Y^{calc}	$Y - Y^{calc}$
1	62.51	63.07	62.89	0.1772
2	71.77	71.56	72.09	–0.5347
3	80.29	80.18	80.56	–0.3812
4	85.92	85.92	86.16	–0.2359
5	92.41	93.03	92.61	0.4249
6	106.62	107.6	106.73	0.8740
7	111.68	112.31	111.75	0.5558
8	121.29	121.2	121.30	–0.1039
9	133.66	132.82	133.60	–0.7762
$s_y =$		23.54	$s_0 = \sqrt{[(1/n-2) \times \sum(Y - Y^{calc})2]} =$	0.584

Матриця $x^T \times x$

$$x^T x = \begin{bmatrix} 9 & 866.15 \\ 866.15 & 87843.3 \end{bmatrix}.$$

Детермінант цієї матриці дорівнює:

$$Det = 9 \times 87843.3 - 866.15^2 = 40373.7.$$

Інформаційно-коваріаційна матриця (12.19)

$$ICM = (x^T \times x)^{-1} =$$

$$= \frac{1}{40373.7} \times \begin{bmatrix} 87843.3 & -866.15 \\ -866.15 & 9 \end{bmatrix} =$$

$$= \begin{bmatrix} 2.17576 & -0.02145 \\ -0.02145 & 0.0002229 \end{bmatrix}.$$

Розрахункова матриця (12.18)

Розрахункову матрицю (12.18) див. на Рис. 12П-2.

Вектор параметрів (12.18)

$$RM \times Y = \begin{bmatrix} 0.7752 \\ 0.9937 \end{bmatrix}.$$

Отже, ми отримали лінійну залежність:

$$Y = 0.775 + 09937 \times X.$$

Розрахунки за цією залежністю величин Y^{calc}, відхилень $Y - Y^{calc}$ і залишкового стандартного відхилення s_0 за співвідношенням (12.3) наведені в Табл. 12.П-1. Використовуючи вихідне стандартне відхилення ординат $s_Y = 23.54$ % і розраховане значення $s_0 = 0.586$ %, знайдемо за співвідношенням (12.5) загальний коефіцієнт кореляції:

$$R_c = \sqrt{1 - \frac{0.586^2}{23.54^2}} = 0.99969.$$

Використовуючи $s_0 = 0.586$ % і діагональні елементи розрахованої вище інформаційно-коваріаційної матриці, розрахуємо за формулою (12.20) стандартні відхилення для параметрів:

$$s_{b1} = 0.586 \times \sqrt{2.17576} = 0.861.$$

$$s_{b2} = 0.586 \times \sqrt{0.0002229} = 0.0087.$$

Підсумкові результати обробки експериментальних даних за допомогою МНК мають вигляд:

$b_1 = 0.775$, $s_{b1} = 0.861$; $b_2 = 0.9937$, $s_{b2} = 0.0087$, $s_0 = 0.586$; $R_c = 0.99969$.

12.9.2. Дослідження значимості впливу швидкості і сили пресування на однорідність маси і однорідність дозування одиниць ядер таблеток дезлоратадину

На стадії фармацевтичної розробки таблеток дезлоратадину виникла необхідність оцінити значимість впливу двох факторів — сили пресування (x_2) і швидкості таблетування (x_3) — на однорідність маси (y_1) і однорідність дозованих одиниць (y_2) ядер таблеток перед їх покриттям плівковою оболонкою. Для цього зручно застосувати кореляційний двофакторний аналіз, використовуючи як функцію відгуку відповідні відносні дисперсії:

$$RSD_1^2 = y_1 = b_{11} + b_{12} \times x_2 + b_{13} \times x_3. \quad (12.П\text{-}2.1)$$

$$RSD_2^2 = y_2 = b_{21} + b_{22} \times x_2 + b_{23} \times x_3. \quad (12.П\text{-}2.2)$$

Тут

x_2 — сила пресування, кН;

x_3 — швидкість пресування, табл./хв.

$y_1 = RSD_1^2$ — відносна дисперсія маси ядер таблеток;

$y_2 = RSD_2^2$ — відносна дисперсія дозованих одиниць ядер таблеток.

Для отримання прямих (12.П-2.1) і (12.П-2.2) використовували ЛМНК, описаний у розділі 12.5. Оскільки є вільний член (b_1), то у співвідношеннях (12.П-2.1) і (12.П-2.2) покладали $x_1 = 1$.

$$x^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 62.51 & 71.77 & 80.29 & 85.92 & 92.41 & 106.62 & 111.68 & 121.29 & 133.66 \end{bmatrix}$$

Рисунок 12П-1. *Матриця x^T*

$$RM = ICM \times x^T =$$

$$\begin{bmatrix} 0.83471 & 0.63605 & 0.45327 & 0.33249 & 0.19325 & -0.1116 & -0.2202 & -0.4263 & -0.6917 \\ -0.0075 & -0.0055 & -0.0036 & -0.0023 & -0.0009 & 0.00231 & 0.00344 & 0.00558 & 0.00834 \end{bmatrix}$$

Рисунок 12П-2. *Розрахункова матриця (12.18)*

Наведений нижче алгоритм легко реалізується в *Excel*®, що забезпечує простежуваність розрахунків.

Вихідні дані наведені в Табл. 12.П-2.1.

У цьому разі кількість дослідів $n = 9$, число параметрів прямих (12.П-2.1) і (12.П-2.2) $m = 3$, кількість ступенів свободи 9 – 3 = 6.

Для розрахунків за ЛМНК застосували підхід, описаний у розділі 12.5 і продемонстрований у розділі 12.7.1. Для розрахунків загальних коефіцієнтів кореляції використані значення s_y, наведені в Табл. 12.П-2.1.

Матриця $x^T \times x$ (за результатами Табл. 12.П-2.2)

9	105	315
105	1659	3675
315	3675	14775

***Det* матриці**

$x^T \times x = 9 \times (1659 \times 14775 - 3675 \times 3675) - 105 \times$
$\times (105 \times 14775 - 315 \times 3675) + 315 \times (105 \times 3675 -$
$- 315 \times 1659) = 14647500$

Інформаційно-коваріаційна матриця $ICM = (x^T \times x)^{-1}$

Це матриця алгебраїчних доповнень елементів матриці $x^T \times x$ (див. вище), поділена на *Det*:

0.751398	–0.026882	–0.009333
–0.026882	0.002304	0.000000
–0.009333	0.000000	0.000267

Розрахункова матриця $RM = (x^T \times x)^{-1} \times x^T$

0.577419	–0.019969	–0.006667
0.344086	–0.019969	0
0.110753	–0.019969	0.006667
0.335484	0.000768	–0.006667
0.102151	0.000768	0
–0.131183	0.000768	0.006667
0.12043	0.019201	–0.006667
–0.112903	0.019201	0
–0.346237	0.019201	0.006667

Результати розрахунків наведені в Табл. 12.П-2.3. У ній представлено також критичне значення загального коефіцієнта кореляції $R_c = 0.621$ для ймовірності 95 % і кількості ступенів свободи $m - n = 9 - 3 = 6$ (див. Табл. 11.6). Значення

Таблиця 12.П-2.1

Вихідні дані залежності однорідності маси й однорідності дозованих одиниць від сили і швидкості пресування

№	Сила пресування	Швидкість пресування	Однорідність маси		Однорідність дозованих одиниць	
	x_1	x_2	$yRSD_1$	$y_1 = RSD_1^2$	RSD_2	$y_2 = RSD_2^2$
1	3	10	0.60	0.3571	2.17	4.6950
2	3	35	0.40	0.1584	2.08	4.3156
3	3	60	1.10	1.2100	1.33	1.7740
4	12	10	0.89	0.7987	1.75	3.0454
5	12	35	0.70	0.4969	1.48	2.1851
6	12	60	0.60	0.3550	0.90	0.8087
7	20	10	0.80	0.6336	1.99	3.9565
8	20	35	1.30	1.6968	1.43	2.0372
9	20	60	1.19	1.4144	2.23	4.9760
			s_{y1} =	**0.533**	s_{y2} =	**1.47**

Таблиця 12.П-2.2

Розрахунок матриці $x^T \times x$

№	$x_1 \times x_1$	$x_2 \times x_2$	$x_3 \times x_3$	$x_1 \times x_2 = x_2 \times x_1$	$x_1 \times x_3 = x_3 \times x_1$	$x_2 \times x_3 = x_3 \times x_2$
1	1	9	100	3	10	30
2	1	9	1225	3	35	105
3	1	9	3600	3	60	180
4	1	144	100	12	10	120
5	1	144	1225	12	35	420
6	1	144	3600	12	60	720
7	1	400	100	20	10	200
8	1	400	1225	20	35	700
9	1	400	3600	20	60	1200
Сума	**9**	**1659**	**14775**	**105**	**315**	**3675**

$R_c < 0.621$ статистично незначуще відрізняються від нуля.

Як видно з Табл. 12.П-2.3, в обох випадках коефіцієнти кореляції є статистично незначущими. Параметри b_2 і b_3 також незначуще відрізняються від нуля.

Отже, для цього обсягу експерименту і досліджуваного діапазону, результати регресійного аналізу свідчать про те, що сила і швидкість пресування статистично значуще не впливають на фармакопейні характеристики *однорідність маси* і *однорідність дозованих одиниць*.

12.9.3. Лінійний зважений метод найменших квадратів: побудова калібрувальної прямої в припущенні рівності відносних невизначеностей ординат

Під час розробки спектрофотометричної методики кількісного визначення активного фармацевтичного інгредієнта (АФІ) в біологічному

Таблиця 12.П-2.3

Розрахунок вектора параметрів, стандартних відхилень і коефіцієнтів кореляції

Однорідність маси					
$RSD_1^2 = y_1 = b_{11} + b_{12} \times x_2 + b_{13} \times x_3$					
Матриця $RM \times y_1$				y_1^{calc}	$(y_1^{calc} - y_1)^2$
0.20621	–0.00713	–0.00238		0.25735	0.00995
0.05451	–0.00316	0		0.45568	0.08837
0.13401	–0.02416	0.00807		0.65401	0.30913
0.26795	0.00061	–0.00533		0.60579	0.03721
0.05076	0.00038	0		0.80412	0.09439
–0.04657	0.00027	0.00237		1.00245	0.41921
0.07631	0.01217	–0.00422		0.91551	0.07946
–0.19157	0.03258	0		1.11384	0.3398
–0.48973	0.02716	0.00943		1.31217	0.01046
0.06188	**0.03872**	**0.00793**	**← Сума**	$s_0(y_1) = \sqrt{(\sum / 6)} =$	**0.4810**
b_{11}	b_{12}	b_{13}			
$b_{11} =$	**0.062**		$s_{b11} =$	$0.481 \times \sqrt{0.7514} =$ **0.417**	
$b_{12} =$	**0.039**		$s_{b12} =$	$0.481 \times \sqrt{0.002304} =$ **0.023**	
$b_{13} =$	**0.0079**		$s_{b13} =$	$0.481 \times \sqrt{0.000267} =$ **0.0079**	
$s_{y1} =$	**0.533**		$s_0(y_1) =$	**0.481**	
		$R_c(y_1) =$	$\sqrt{[1 - (0.481 / 0.533)^2]} = \mathbf{0.432} < 0.621 = R_c^{crit}$		
Однорідність дозованих одиниць					
$RSD_2^2 = y_2 = b_{21} + b_{22} \times x_2 + b_{23} \times x_3$					
Матриця $RM \times y_2$				Y_1^{calc}	$(y_2^{calc} - y_2)^2$
2.7109967	–0.09376	–0.0313		3.77864	0.83976
1.4849345	–0.08618	0		3.088941	1.50467
0.1964706	–0.03542	0.011826		2.399241	0.39098
1.0216739	0.002339	–0.0203		3.777834	0.53650
0.2232066	0.001678	0		3.088135	0.81552
–0.1060928	0.000621	0.005392		2.398436	2.52713
0.476484	0.07597	–0.02638		3.777119	0.03218
–0.2300048	0.039116	0		3.08742	1.10299
–1.7228809	0.095546	0.033173		2.397721	6.64764
4.055	**–0.00009**	**–0.0276**		$s_0(y_2) = \sqrt{(\sum / 6)} =$	1.549
b_{21}	b_{22}	b_{23}			
$b_{21} =$	**4.055**		$s_{b21} =$	$1.549 \times \sqrt{0.7514} =$ **1.343**	
$b_{22} =$	**0.000**		$s_{b22} =$	$1.549 \times \sqrt{0.002304} =$ **0.074**	
$b_{23} =$	**–0.028**		$s_{b23} =$	$1.549 \times \sqrt{0.000267} =$ **0.025**	
$s_{y2} =$	**1.468**		$s_0(y_2) =$	**1.549**	
		$Rc(y_2) =$	$\sqrt{[1 - (1.549 / 1.468)^2]} = \mathbf{0} < 0.621 = R_c^{crit}$		

середовищі виникла необхідність побудови калібрувального графіка залежності концентрації АФІ (y) від вимірюваної оптичної густини (x) з однаковою відносною невизначеністю результатів у широкому інтервалі концентрацій (див. Табл. 12.П-3.1). Для цього зручно застосувати ЛЗМНК, описаний у розділі 12.6, у припущенні рівності відносних стандартних відхилень (12.10-12.11).

Ми будуємо калібрувальну залежність концентрації (x) від оптичної густини (y):

$$y = b_1 + b_2 \times x_2. \qquad (12.П\text{-}3)$$

Тут x_2 і y — вектори-стовбці з Табл. 12.П-3. У цьому разі число дослідів $n = 9$, число параметрів прямої (12.П-4) $m = 2$, число ступенів свободи $n - m = 7$. Оскільки є вільний член, то покладаємо $x_1 = 1$.

Застосовуємо той самий алгоритм, що і в розділі 12.7.1, з урахуванням появи вагових множників.

У розрахунках припускається рівність відносних стандартних відхилень ординат, тобто допущення (12.10-12.11). Наведена нижче схема може бути легко реалізована в *Excel*®, що забезпечує її простежуваність.

Матриці x і x^T

Ураховуючи, що в лінійній залежності (12.П-1) є вільний член b_1, ці матриці мають у цьому разі такий вигляд:

$$x = \begin{bmatrix} 1 & 0.135 \\ 1 & 0.269 \\ 1 & 0.407 \\ 1 & 0.531 \\ 1 & 0.635 \\ 1 & 0.783 \\ 1 & 0.893 \\ 1 & 1.040 \\ 1 & 1.180 \end{bmatrix}.$$

Матрицю x^T див. на Рис. 12П-3.

Матриця $x^T \times w \times x$ (за даними Табл. 12.П-3.2)

$$x^T \times w \times x = \begin{bmatrix} 1.3291 & 0.3479 \\ 0.3479 & 0.1566 \end{bmatrix}.$$

Детермінант цієї матриці дорівнює

$$Det = 1.3291 \times 0.1566 - 0.3479^2 = 0.087133.$$

Інформаційно-коваріаційна матриця (12.22)

$$ICM = (x^T \times w \times x)^{-1} =$$

$$= \frac{1}{0.087133} \times \begin{bmatrix} 0.1566 & -0.3479 \\ -0.3479 & 1.3291 \end{bmatrix} =$$

$$= \begin{bmatrix} 1.7976 & -3.9928 \\ -3.9928 & 15.2532 \end{bmatrix}.$$

Таблиця 12.П-3.1

Вихідні дані і результати розрахунків за прямою $Y = b_1 + b_2 \times x_2$

	Вихідні дані		Ваги	Розрахунки за ЛЗМНК*		Розрахунки за ЛМНК**	
N досліду	x_2, оптична густина	y, мкг/мл	$w = 1 / y^2$	y^{calc}	$(y - y^{calc}) / y$	y^{calc}	$(y - y^{calc}) / y$
1	0.135	1.10	0.8264	1.13	0.0228	1.24	0.1256
2	0.269	2.15	0.2163	2.10	–0.0251	2.17	0.0115
3	0.407	3.15	0.1008	3.10	–0.0172	3.14	–0.0034
4	0.531	4.20	0.0567	3.99	–0.0489	4.01	–0.0462
5	0.635	4.80	0.0434	4.75	–0.0108	4.73	–0.0140
6	0.783	5.70	0.0308	5.82	0.0211	5.77	0.0118
7	0.893	6.50	0.0237	6.62	0.0181	6.54	0.0056
8	1.04	7.80	0.0164	7.68	-0.0150	7.56	–0.0303
9	1.18	8.30	0.0145	8.70	0.0478	8.54	0.0292
$s_y =$		**2.468**		$s_0 = \sqrt{(\sum кв / 7)} =$	**0.0322**		**0.0537**

* $y = 0.295 + 6.989 \times x_2$ (ЛЗМНК).
** $y = 0.147 + 7.246 \times x_2$ (ЛМНК).

$$x^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0.135 & 0.269 & 0.407 & 0.531 & 0.635 & 0.783 & 0.893 & 1.040 & 1.180 \end{bmatrix}.$$

Рисунок 12П-3. *Матриця x^T*

Розрахункова матриця (12.23)

Розрахункову матрицю (12.23) див. на Рис. 12П-4.

Вектор параметрів (12.23)

$$RM \times Y = \begin{bmatrix} 0.1469 \\ 7.2460 \end{bmatrix}.$$

Лінійна залежність, отримана за допомогою ЛЗМНК:

$$Y = 0.147 + 7.246 \times X.$$

Розрахунки за цією залежністю величин Y^{calc}, відхилень $Y - Y^{calc}$ і залишкового стандартного відхилення s_0 за співвідношеннями (12.7-12.8) наведені в Табл. 12.П-3.1 (ЛЗМНК). Використовуючи вихідне стандартне відхилення ординат $s_Y = 1.468$ і розраховане значення $s_0 = 0.0322$, знайдемо за співвідношенням (12.9) загальний коефіцієнт кореляції:

$$R_c = \sqrt{1 - \frac{0.0322^2}{1.468^2}} = 0.99992.$$

Використовуючи $s_0 = 0.0322$ і діагональні елементи розрахованої вище інформаційно-коваріаційної матриці, розрахуємо за рівняннями (12.20) і (12.22) стандартні відхилення для параметрів:

$$s_{b1} = 0.322 \times \sqrt{1.7976.} = 0.254.$$

$$s_{b2} = 0.322 \times \sqrt{15.2532} = 0.740.$$

Підсумкові результати обробки експериментальних даних за допомогою ЛЗМНК мають вигляд:

$$b_1 = 0.147, \quad s_{b1} = 0.254; \quad b_2 = 7.246, \quad s_{b1} = 0.740;$$

$$s_0 = 0.0322; \quad R_c = 0.99992.$$

У Табл. 12.П-3.1 також наведені для порівняння результати розрахунків за звичайним (незваженим) ЛМНК (які виконані за схемою розділу 12.7.1). Як видно, значення параметрів b_1 і b_2, розраховані за ЛЗМНК і ЛМНК помітно різняться. Водночас відносне залишкове стандартне відхилення s_0 для ЛЗМНК (0.0322) суттєво краще, ніж для ЛМНК (0.0537). Особливо великим є відносне відхилення для першої точки, розрахованої за ЛМНК (0.1256, або 12.56 %), яке значно гірше, ніж за ЛЗМНК (0.0228, або 2.28 %).

Отже, у цьому разі калібрування з використанням ЛЗМНК дає кращі результати, ніж ЛМНК.

12.9.4. Застосування нелінійного методу найменших квадратів для опису кінетики розчинення зразків субстанції фенсукциналу з різним розміром частинок

Під час фармацевтичної розробки готових лікарських засобів (ГЛЗ) субстанції фенсукциналу (ФС) постало питання, наскільки суттєвим для фармако-технологічних показників ГЛЗ є різниця в кінетиці розчинення фракцій ФС різного ступеня подрібнення (розмір частинок). Чим менше розмір частинок, тим більше швидкість розчинення, що може бути критичним

Таблиця 12.П-3.2

Розрахунок матриці $x^T \times w \times x$

x_1	x_2	w	$x_1 \times w \times x_1$	$x_2 \times w \times x_2$	$x_1 \times w \times x_2 = x_2 \times w \times x_1$
1	0.135	0.8264	0.8264	0.01506	0.1116
1	0.269	0.2163	0.2163	0.01565	0.0582
1	0.407	0.1008	0.1008	0.01669	0.0410
1	0.531	0.0567	0.0567	0.01598	0.0301
1	0.635	0.0434	0.0434	0.01750	0.0276
1	0.783	0.0308	0.0308	0.01887	0.0241
1	0.893	0.0237	0.0237	0.01887	0.0211
1	1.040	0.0164	0.0164	0.01778	0.0171
1	1.180	0.0145	0.0145	0.02021	0.0171
		Сума	**1.3291**	**0.1566**	**0.3479**

$$RM = ICM \times x^T \times w =$$

$$-\begin{bmatrix} 1.04014 & 0.15653 & 0.17388 & -0.0183 & -0.0320 & -0.0409 & -0.04185 & -0.03871 & -0.04230 \\ -1.5980 & 0.02386 & 0.22326 & 0.23280 & 0.24709 & 0.24709 & 0.22789 & 0.19511 & 0.20331 \end{bmatrix}.$$

Рисунок 12П-4. *Розрахункова матриця (12.23)*

для технології отримання розчинів або для випробування «Розчинення» твердих лікарських засобів. З іншого боку, зменшення розміру частинок вимагає додаткових витрат. Крім того, зменшення розміру частинок субстанції часто призводить до зменшення її стабільності. Тому це питання становить практичний інтерес.

У процесі досліджень було показано, що за фіксованого середнього розміру частинок L залежність концентрації ФС ($C(L, t)$ мг/л), який перейшов у розчин, від часу розчинення (t, час) описується одночастевою експоненціальною моделлю:

$$y = C(L,t) = C(L,t=\infty) \times \times [1 - \exp(-k \times t)], \qquad (12.\text{П-}4.1)$$

яка є функцією двох параметрів: $b_1 = C(L, t=\infty)$ — екстраполяційна максимальна концентрація ФС в розчині для цього середнього розміру частинок за безкінечного часу розчинення; $b_2 = k$ — параметр, що характеризує швидкість розчинення.

Експоненціальна модель (12.П-4.1) не може бути приведена до лінійного вигляду відносно параметрів $C(L, t = \infty)$ і k, тому для опису експериментальних даних за цим рівнянням необхідно використовувати нелінійний метод найменших квадратів (НМНК), який описаний у розділах 12.7-12.8. НМНК передбачає числову оптимізацію параметрів.

Експериментальні дані залежності концентрації фенсукциналу, який перейшов до розчину (0.1 М кислота хлористоводнева), від часу розчинення для різних фракцій фенсукциналу наведені в Табл. 12.П-4.1.

Дослідження кінетики розчинення проводили для часу: t = 5, 10, 20, 30, 40, 50, 60, 90 і 120 хв. Для розрахунків за допомогою НМНК доцільно використовувати час у годинах, що також зазначено в Табл. 12.П-4.1. Це суттєво зменшує часткові похідні за параметрами, що покращує збіжність процедури оптимізації.

Як видно з Табл. 12.П-4.1, стандартні відхилення не відрізняються значуще за критерієм Кокрена (див. Табл. 10.4), тобто виконуються вимоги (12.6) рівноточності ординат. Тому для розрахунків ми використовували звичайний (незважений) НМНК, описаний у розділі 12.7.

Нижче наведений приклад розрахунку за ітераційної процедури, описаної в розділі 12.7, для фракції 0.05-0.10 мм фенсукциналу. Ця процедура для двопараметричної одночастевої експоненціальної моделі (12.П-4.1) може бути легко реалізована в *Excel*®, що забезпечує її простежуваність.

Похідні за параметрами (12.26)

Проводячи відповідне диференціювання функції (12.П-4.1) за параметрами $b_1 = C(L, t=\infty)$ і $b_2 = k$, отримаємо:

Таблиця 12.П-4.1

Вихідні результати дослідження кінетики розчинення в 0.1 М кислоті хлористоводневій різних фракцій фенсукциналу залежно від часу

Час	**Концентрація фенсукциналу в розчині, $y = C$ мг/л для фракцій (мм):**				**Стандартне відхилення $s(y) = SD$ мг/л (n = 3)* для фракцій (мм):**			
Години (хвилини)	**0.05-0.10**	**0.10-0.16**	**0.16-0.315**	**0.315-0.50**	**0.05-0.10**	**0.10-0.16**	**0.16-0.315**	**0.315-0.50**
0.0833 (5 хв)	4.07	5.8	4.80	4.2	1.16	1.50	0.87	0.71
0.167 (10 хв)	13.2	15.7	11.5	7.2	3.06	1.86	1.75	0.36
0.333 (20 хв)	20.9	21.5	16.0	10.2	2.65	1.54	2.31	0.58
0.50 (30 хв)	24.1	24.2	18.4	12.8	1.73	1.64	1.78	1.66
0.667 (40 хв)	26.2	26.5	20.2	14.3	1.21	2.05	1.70	2.11
0.833 (50 хв)	27.3	27.4	22.1	15.9	1.08	2.19	1.60	2.63
1.0 (60 хв)	27.8	28.6	22.6	17.0	0.85	1.73	1.12	2.75
2.0 (120 хв)	29.7	30.3	25.4	19.1	0.85	2.26	1.77	0.14
3.0 (180 хв)	30.4	30.8	25.0	21.0	0.78	1.48	0.00	0.07
$s_y = s_C =$	**8.73**	**7.98**	**6.75**	**5.50**				
Сума квадратів					25.41	30.06	22.16	22.68
Критерій Кокрена $G(0.95, v = 2) = SD^2_{max}$ / ∑квадратів < 0.5157					0.369	0.160	0.241	0.333

* Підкреслені максимальні значення, необхідні для перевірки гіпотези рівноточності ординат за Кокреном (Табл. 10.4).

$$E_{i1} = \frac{\partial y^{calc}}{\partial b_1} = \frac{\partial C(L,t_i)}{\partial C(L,t=\infty)} = \quad (12.П\text{-}4.2)$$
$$= 1 - \exp(-k \times t_i),$$

$$E_{i2} = \frac{\partial y^{calc}}{\partial b_2} = \frac{\partial C(L,t_i)}{\partial k} = \quad (12.П\text{-}4.3)$$
$$= C(L,t=\infty) \times \exp(-k \times t_i)] \times t_i.$$

Кількість дослідів $\boldsymbol{n = 9}$, число параметрів $\boldsymbol{m = 2}$, число ступенів свободи $\boldsymbol{n - m = 7}$.

Точність оптимізації $\boldsymbol{\Delta s_0 = 0.001}$.

Вихідні наближення для параметрів: $\boldsymbol{b_{10} = 100}$, $\boldsymbol{b_{20} = 10}$.

Матриця $E^T \times E$ (за даними Табл. 12.П-4.2)

7.8916	6.192
6.192	24.539

***Det* матриці** $E^T \times E = 7.8916 \times 24.539 + 6.192 \times 6.192 = 155.315$.

Інформаційно-коваріаційна матриця $ICM = (E^T \times E)^{-1}$

Це матриця алгебраїчних доповнень, поділена на *Det*:

$$\begin{bmatrix} 24.539 / 155.315 & -6.192 / 155.315 \\ -6.192 / 155.315 & 7.8916 / 155.315 \end{bmatrix} = \begin{bmatrix} 0.1579975 & -0.0398671 \\ -0.0398671 & 0.05081042 \end{bmatrix}$$

Таблиця 12.П-4.2

Розрахунок матриці $E^T \times E$

№	t (час)	E_{i1}	E_{i2}	$E_{i1} \times E_{i1}$	$E_{i2} \times E_{i2}$	$E_{i1} \times E_{i2} = E_{i2} \times E_{i1}$
1	0.0833	0.5652569	3.6214	0.319515	13.11461	2.047027
2	0.167	0.8117529	3.1437	0.658943	9.8830131	2.5519288
3	0.333	0.9642069	1.1919	0.929695	1.4206504	1.1492482
4	0.5	0.9932621	0.3369	0.98657	0.1134998	0.3346274
5	0.667	0.9987316	0.0846	0.997465	0.0071575	0.0844949
6	0.833	0.9997588	0.0201	0.999518	0.0004036	0.0200848
7	1	0.9999546	0.0045	0.999909	$2.061E^{-05}$	0.0045398
8	2	1	$4E^{-07}$	1	$1.699E^{-13}$	$4.122E^{-07}$
9	3	1	$3E^{-11}$	1	$7.881E^{-22}$	$2.807E^{-11}$
Сума				**7.891614**	**24.53936**	**6.191951**

Таблиця 12.П-4.3

Розрахунок параметрів, стандартних відхилень, коефіцієнта кореляції і їх змін

№	y_i	y_i^{calc}	$\Delta y_i = y_i^{calc} - y_i$	$RM_{1i} \times \Delta y_i$	$RM_{2i} \times \Delta y_i$
1	4.07	56.52569	52.45569	–2.8885201	8.4700324
2	13.2	81.175293	67.975293	0.1987401	8.6581372
3	20.9	96.420689	75.520689	7.9163996	1.6706199
4	24.1	99.326205	75.226205	10.795094	–1.6911277
5	26.2	99.87316	73.67316	11.37692	–2.6167119
6	27.3	99.975883	72.675883	11.421629	–2.8224921
7	27.8	99.99546	72.19546	11.393115	–2.8614379
8	29.7	100	70.3	11.10722	–2.8026546
9	30.4	100	69.6	10.996623	–2.7747492
$s_y =$	**8.813**	$s_0 =$	79.680*		
$R_c = \sqrt{[1 - (79.680 / 8.813)^2]}$ — не має сенсу, тобто дорівнює 0					
$\Delta b_1 = \Delta C(L, t = \infty) =$				72.317**	
$\Delta b_2 = \Delta k =$					3.230**
$b_1 = b_{10} - \Delta b_1 =$				100 – 72.317 = **27.683**	
$b_2 = b_{20} - \Delta b_2 =$					10 – 3.230 = **6.770**
$s_{b1} =$				$79.680 \times \sqrt{0.1580}$ = **31.67**	
$s_{b2} =$					$79.680 \times \sqrt{0.0508}$ = **17.96**

* $\sqrt{(\Sigma \Delta y^2 / 7)}$.

** Сума наведених вище елементів.

Розрахункова матриця $RM = (E^T \times E)^{-1} \times E^T$

RM_{i1}	RM_{i2}
–0.05506591	0.1614702
0.00292371	0.1273718
0.10482425	0.0221214
0.14350177	–0.0224806
0.15442422	–0.0355178
0.15715844	–0.0388367
0.15780929	–0.0396346
0.15799745	–0.0398671
0.15799746	–0.0398671

Процедуру ітерацій повторюють, використовуючи отримані значення $b_1 = 27.683$ і $b_2 = 6.770$ замість вихідних значень $b_{10} = 100$, $b_{20} = 10$. Результати цих ітерацій наведені в Табл. 12.П-4.4.

Як видно, для отримання заданої точності залишкового стандартного відхилення ($\Delta s_0 < 0.001$) достатньо 6 ітерацій.

Підсумкові результати розрахунків за допомогою НМНК за рівнянням (12.П-4.1) для різних фракцій ФС, зазначених у Табл. 12.П-4.1, наведені в Табл. 12.П-4.5.

Як видно, у досліджуваному проміжку часу (5-180 хв) в 0.1 М кислоті хлористоводневій спостерігається дуже значна залежність максимального ступеня розчинення фенсукциналу $C(L, t = \infty)$ від розміру частинок. Це може суттєво впливати на технологічний процес отримання рідких лікарських засобів і на виконання фармакопейного тесту «Розчинення» для твердих лікарських засобів.

Таблиця 12.П-4.4

Результати знаходження параметрів рівнянь 12.П-4.1 за запропонованим алгоритмом

	Ітерації						
	0	**1**	**2**	**3**	**4**	**5**	**6**
$b_1 = C(L, t = \infty)$	100	27.683	28.64	28.168	29.737	29.788	29.788
$\Delta b_1 = \Delta C(L, t = \infty)$	72.4	–0.954	0.472	–0.422	–0.051	-3.19×10^{-5}	-3.19×10^{-5}
$b_2 = k$	10	6.770	1.549	2.792	3.214	3.221	3.2209
$\Delta b_2 = \Delta k$	3.23	5.221	–1.243	–0.424	–0.007	$1.43E^{-05}$	1.43×10^{-5}
s_0	79.7	4.269	7.046	2.741	1.3267	1.3254	1.32537
Δs_0		–75.411	2.777	–4.305	–1.4143	–0.0013	-3.0×10^{-5}
R_c	0	0.875	0.601	0.950	0.989	0.989	0.989
ΔR_c		0.875	–0.274	0.350	0.038	0.000	0.000
$s_{b1} = s_{C(L, t = \infty)}$	31.7	1.846	5.936	1.718	0.777	0.776	0.776
$s_{b2} = s_k$	18.0	2.204	0.784	0.544	0.290	0.290	0.290

Таблиця 12.П-4.5

Метрологічні характеристики кінетики розчинення фракцій фенсукциналу з різним розміром частинок в 0.1 М кислоті хлористоводневій (t — години)

$$C(L,t) = C(L,t = \infty) \times [1 - \exp(-k_t \times t)]$$

Розмір частинок, мм	$C(L, t = \infty)$ *мг/л*	s_C	k_t	s_k	s_0	R_c	**Рівень значимості, R_c, %**
0.05–0.10	29.8	0.77	3.22	0.29	1.32	0.989	100.00
0.10–0.16	27.9	0.72	2.97	0.25	1.19	0.989	100.00
0.16–0.315	24.5	0.61	3.00	0.25	1.00	0.989	100.00
0.315–0.50	19.8	0.62	2.11	0.18	0.87	0.987	100.00