

Міністерство освіти і науки України
Вінницький національний технічний університет

В. І. РИНДЮК, І. В. КОЦ
МАТЕМАТИЧНІ МЕТОДИ
РОЗВ'ЯЗАННЯ ІНЖЕНЕРНИХ ЗАДАЧ
У БУДІВНИЦТВІ

Вінниця ВНТУ 2006
Міністерство освіти і науки України
Вінницький національний технічний університет

В. І. РИНДЮК, І. В. КОЦ
МАТЕМАТИЧНІ МЕТОДИ
РОЗВ'ЯЗАННЯ ІНЖЕНЕРНИХ ЗАДАЧ
У БУДІВНИЦТВІ

Вінниця ВНТУ 2006
Міністерство освіти і науки України
Вінницький національний технічний університет

В. І. РИНДЮК, І. В. КОЦ
МАТЕМАТИЧНІ МЕТОДИ
РОЗВ'ЯЗАННЯ ІНЖЕНЕРНИХ ЗАДАЧ
У БУДІВНИЦТВІ

Затверджено Вченою радою Вінницького національного технічного університету як практикум для магістрантів будівельних спеціальностей 8.092101- «Промислове та цивільне будівництво» і 8.092108 – «Теплогазопостачання і вентиляція». Протокол № 4 від 25 листопада 2004 р.

Вінниця ВНТУ 2006

Зміст

Вступ	5
I Елементи теорії похибок	5
1.1 Розв'язування задачі. Джерела і класифікація похибок	5
1.2 Похибки наближених чисел	6
1.3 Загальна формула для обчислення похибки	10
1.4 Поняття про ймовірну оцінку похибки. Обчислення без точного врахування похибок	12
1.5 Обернена задача теорії похибок	14
II Математичне моделювання, модель і експеримент	15
2.1 Інтерполяція в математичному моделюванні	16
2.2 Основні характерні риси моделювання	16
2.2.1 Основні типи математичних моделей	18
2.3 Планування експерименту	22
2.3.1 Підбір емпіричних формул та згладжування	23
2.3.2 Лінійне згладжування	25
2.3.3 Нелінійне згладжування	26
2.3.4 Лінійна кореляція і регресія	26
2.3.5 Рівняння прямих регресії	29
2.4 Інтерполяція функцій однієї і декількох змінних	32
2.4.1 Одновимірна інтерполяція	33
2.4.2 Інтерполяція сплайнами	34
2.4.3 Багатовимірна інтерполяція	36
III Елементи чисельних методів	38
3.1 Розв'язання системи лінійних алгебраїчних рівнянь методом простої ітерації	38
3.1.1 Загальні положення	38
3.1.2 Приклад	39
3.1.3 Текст програми розрахунку	40
3.2 Обчислення дійсного кореня рівняння з даною точністю із застосуванням комбінованого методу	41
3.2.1 Загальні положення	41
3.2.2 Приклад	42
3.2.3 Текст програми розрахунку	43
3.3 Наближене розв'язання систем нелінійних рівнянь методом Ньютона	43
3.3.1 Загальні положення	43
3.3.2 Приклад	45
3.3.3 Текст програми розрахунку	46
3.4 Точкове апроксимування функцій за методом найменших квадратів	47
3.4.1 Загальні положення	47
3.4.2 Приклад	48

3.4.3	Текст програми розрахунку	50
3.5	Одновимірна мінімізація функцій	51
3.5.1	Загальні положення	51
3.5.2	Приклад	52
3.5.3	Текст програми розрахунку	53
3.6	Багатовимірна безумовна мінімізація функцій. Градієнтні методи..	54
3.6.1	Загальні положення	54
3.6.1.1	Градієнтний метод з дрібненням кроку	54
3.6.1.2	Метод покоординатного спуску	55
3.6.1.3	Приклад	57
3.6.1.4	Текст програми розрахунку	61
3.7	Чисельне інтегрування та диференціювання функцій однієї змінної	62
IV	Елементи варіаційного числення та метод скінчених елементів	66
4.1	Метод скінчених елементів	68
4.1.1	Загальні положення	68
4.1.2	Побудова елемента	69
4.1.3	Побудова інтерполяційних многочленів	70
4.1.4	Одновимірний приклад варіаційного методу скінчених елементів.....	71
4.1.5	Двовимірний приклад варіаційного методу скінчених елементів	78
V	Постановка задач лінійного програмування	84
5.1	Симплекс-метод	88
5.2	Транспортна задача	98
5.2.1	Метод північно-західного кута	101
5.2.2	Метод мінімального елемента	103
5.2.3	Метод апроксимації Фогеля	104
5.2.4	Розв'язання транспортної задачі методом потенціалів.....	106
VI	Градієнтний метод розв'язування задач нелінійного програмування...	110
VII	Література.....	116

Вступ

Важливу роль у формуванні надійних і корисних для науково-технічної практики результатів відіграє етап математичного моделювання й оптимізації об'єктів управління, проектування та їх дослідження. Математичні моделі інженерних задач будівельного матеріалознавства і технології можуть бути розроблені тільки за участю інженерів-технологів з виробництва будівельних матеріалів та конструкцій.

Для цього спеціалістам необхідні знання, які дозволяють сформулювати проектні, матеріалознавчі і технологічні задачі в математичних термінах, вибрати ті чи інші шляхи розв'язання математичної задачі і отримати з результатів математичного моделювання оптимальну (корисну) інженерну інформацію. Оптимізація неможлива без формулювання конкретної математичної задачі (наведеної в математичних виразах) з обов'язковими елементами оптимізації (тобто множини варіантів розв'язків і критерію оптимальності).

Даний посібник буде корисним при вивченні дисциплін:

- математичні методи оптимізації будівельних процесів;
- вступ та основи системного аналізу;
- теорія та практика наукових досліджень.

І ЕЛЕМЕНТИ ТЕОРІЇ ПОХИБОК

1.1 Розв'язування задач. Джерела і класифікація похибок

Процес розв'язування будь-якої реальної фізичної задачі можна розбити на такі етапи:

1. Побудова математичної моделі (математичне формулювання задачі), що охоплює найважливіші для даної задачі сторони явища.

2. Вибір методу розв'язування. Для деяких найпростіших моделей вдається дістати аналітичні розв'язки задачі, а для складніших здебільшого не вдається. У цих випадках використовують наближені методи, зокрема чисельні.

3. Алгоритмізація процесу. При застосуванні чисельних методів потрібно записати алгоритм розв'язування задачі. Якщо задача розв'язуватиметься на ЕОМ, то треба також скласти програму.

4. Виконання обчислень (на ЕОМ чи вручну).

5. Аналіз результатів (осмислення математичного розв'язку і зіставлення його з експериментальними даними).

Важливо вміти оцінити точність розв'язку задачі, який здебільшого ми дістаємо з похибками. Похибки результатів зумовлюються такими причинами:

1. Математична модель лише наближено відображає реальні явища;

2. Вхідні дані, як правило, — це числа неточні (дані для обчислень часто дістають з експерименту, а кожний експеримент може дати результат лише з обмеженою точністю);

3. Метод розв'язування задачі часто є наближеним. У багатьох задачах точний результат можна дістати лише після нескінченної або досить великої кількості арифметичних операцій, які практично здійснити неможливо. Тому замість точного розв'язку здебільшого доводиться відшукувати наближений (наприклад, замість суми ряду беруть суму скінченної кількості його членів, нескінченний ітераційний процес обривають після скінченного числа ітерацій, інтеграл замінюють скінченною сумою і т. п.);

4. Округлення при обчисленнях. Усі обчислення (вручну і на ЕОМ) можна виконувати лише з обмеженою кількістю значущих цифр. Тому при виконанні арифметичних дій потрібно вдаватися до округлень, які зумовлюють похибки, що нагромаджуються в процесі обчислень.

Похибки, що породжуються вищезгаданими причинами, відповідно називаються:

- 1) похибка математичної моделі;
- 2) неусувна похибка (вона не залежить від обчислювача);
- 3) похибка методу;
- 4) похибка округлень.

Повна похибка результату дорівнює сумі всіх перерахованих похибок.

Зауважимо, що похибок, пов'язаних з особливостями будови математичної моделі, не розглядатимемо. Похибки методів досліджуватимемо в кожному окремому випадку.

При виконанні обчислень потрібно дотримуватися такого **правила**: у проміжних результатах похибка округлення має бути значно меншою від інших похибок. При обчисленнях вручну це правило забезпечується збереженням запасних цифр. На сучасних ЕОМ числа записуються також з достатньою кількістю значущих цифр, тому похибкою окремого округлення здебільшого можна нехтувати на противагу похибці методу і неусувній похибці. У процесі виконання великої кількості операцій похибка округлень збільшується. До значного нагромадження таких похибок може призвести віднімання близьких за величиною чисел, знаходження коренів многочленів високих степенів тощо.

1.2 Похибки наближених чисел

Часто на практиці з тих чи інших причин доводиться замість точного числа \tilde{a} брати його наближене значення a . При цьому виникає похибка:

$$\Delta a = \tilde{a} - a.$$

Абсолютною похибкою Δ наближеного числа a називають величину

$$\Delta = |\tilde{a} - a|.$$

Граничною абсолютною похибкою вважають будь-яке число Δ_a , що задовольняє умову:

$$\Delta \leq \Delta_a.$$

Це означає, що $-\Delta_a \leq \tilde{a} - a \leq \Delta_a$.

Надалі ці нерівності скорочено записуватимемо так:

$$\tilde{a} = a \pm \Delta_a.$$

Зауважимо, що найчастіше точне значення \tilde{a} буває невідомим, а тому невідома й похибка Δ наближеного числа a . Наприклад, неможливо точно визначити числа $\pi, \sqrt{2}, e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$ тощо, однак досить легко обчислити межі, у яких вони знаходяться. Можна знайти наближені значення їх та граничні абсолютні похибки. Зрозуміло, що чим менша гранична абсолютна похибка, тим точніше число a наближує число \tilde{a} . Та гранична абсолютна похибка не завжди достатньо характеризує точність обчислень (чи вимірювань). Так, при вимірюванні величини, значення якої дорівнює $5 \cdot 10^7$, граничну абсолютну похибку 0,01 можна вважати досить малою, однак при вимірюванні величини, значення якої дорівнює $5 \cdot 10^{-5}$, похибка 0,01 завелика. Тому часто використовують *відносну похибку наближеного числа a* , яка означається рівністю:

$$\delta = \frac{\Delta}{|a|} \quad (a \neq 0).$$

Очевидно, похибку δ , як і саме число a , не завжди можна визначити. Тому на практиці здебільшого використовують граничну відносну похибку δ_a , що визначається умовою:

$$\delta \leq \delta_a.$$

Отже,

$$\Delta = |a| \cdot \delta \leq |a| \cdot \delta_a = \Delta_a.$$

Якщо $a \neq 0$, то вважатимемо, що для Δ_a і δ_a справджуються співвідношення:

$$\delta_a = \frac{\Delta_a}{|a|}, \quad \Delta_a = |a| \delta_a. \quad (1.2.1)$$

Запишемо додатне число a у вигляді скінченного десяткового дробу:

$$a = \alpha_1 10^m + \alpha_2 10^{m-1} + \dots + \alpha_n 10^{m-n+1}.$$

Або

$$a = \alpha_1 \alpha_2 \dots \alpha_{m+1} \alpha_{m+2} \dots \alpha_n$$

(де всі коефіцієнти $\alpha_i > 0$; $a > 0$ і менші за число 10).

K-та цифра наближеного числа a називається правильною, якщо абсолютна похибка Δ цього числа не перевищує половини одиниці k -го розряду, тобто коли

$$\Delta = |\tilde{a} - a| \leq \frac{1}{2} 10^{m-k+1}.$$

В іншому випадку цифру k -го розряду називають сумнівною.

Зауважимо, що інколи цифру k -го розряду називають правильною, якщо абсолютна похибка числа не перевищує одиниці цього розряду. Цифру за останнім означенням називають *правильною у широкому розумінні*, а за першим – *у вузькому*. Надалі ми вживатимемо поняття правильних цифр в розумінні першого означення.

Значущими цифрами наближеного числа a називатимемо всі його правильні цифри, починаючи з першої зліва, що не дорівнює нулю (її десятковий розряд позначається m), до першої сумнівної цифри включно. Усі інші цифри називатимемо незначущими.

Записуючи остаточні результати наближених обчислень, незначущі цифри числа відкидатимемо. При цьому кожне число записуватимемо як добуток деякого степеня числа 10 на число, всі цифри якого значущі. Якщо з наближеними числами виконуватимуться обчислення, то в них, крім значущих, треба зберігати ще одну або дві сумнівні цифри.

На практиці при виконанні обчислень часто виникає потреба округляти числа.

Якщо округляється наближене чи точне число a до n значущих цифр, то за правилом доповнення число a замінюють числом a_1 з n значущими цифрами так, щоб похибка округлення не перевищувала половини одиниці розряду, що зберігається, тобто, щоб справджувалась умова:

$$|a_1 - a| \leq \frac{1}{2} 10^{m-n+1}.$$

Отже, якщо при округленні числа a до n значущих цифр відкидається менше ніж $-\frac{1}{2} 10^{m-n+1}$, то в числі a_1 зберігаються без зміни всі перші n

цифр числа a . А коли відкидається не менше як $-\frac{1}{2} 10^{m-n+1}$, то в числі a_1 остання цифра береться на одиницю більшою, ніж у числі a .

Зауважимо, що внаслідок округлення наближеного числа a до n значущих цифр похибка округлення Δ_0 може досягти величини $0,5 \cdot 10^{m-n+1}$. Якщо врахувати ще й похибку Δ_a числа a , то n -а цифра округленого числа може вже бути неправильною. Та коли похибка Δ_a досить мала порівняно з максимально можливим значенням похибки округлення Δ_0 (наприклад, $\Delta_a = 0,5 \cdot 10^{m-n}$), то похибкою Δ_a можна знехтувати і n -у цифру числа a після округлення вважати правильною, навіть якщо похибка округлення досягає $0,5 \cdot 10^{m-n+1}$.

Отже, щоб знайти результат з точністю до $0,5 \cdot 10^{m-n+1}$, треба виконувати обчислення так, щоб дістати результат з точністю принаймні до $0,5 \cdot 10^{m-n}$, тобто з $n+1$ правильними цифрами. Тоді після округлення можна буде вважати, що результат має n правильних цифр.

Легко встановити зв'язок відносної похибки наближеного числа a з кількістю правильних значущих цифр. Якщо додатне наближене число a має

n правильних значущих цифр, то його відносна похибка δ задовольняє співвідношення:

$$\delta \leq \frac{1}{\alpha_1} \left(\frac{1}{10} \right)^{n-1},$$

де α_1 – перша значуща цифра наближеного числа a .

Справді,

$$\Delta \leq \frac{1}{2} \cdot 10^{m-n+1},$$

тому

$$\delta = \frac{\Delta}{a} \leq \frac{\frac{1}{2} \cdot 10^{m-n+1}}{\alpha_1 10^m + \alpha_2 10^{m-1} + \dots + \alpha_n 10^{m-n+1} - \frac{1}{2} \cdot 10^{m-n+1}} \leq \frac{10^{m-n+1}}{\alpha_1 10^m} = \frac{1}{\alpha_1} \left(\frac{1}{10} \right)^{n-1}.$$

Таким чином, за граничну відносну похибку можна взяти

$$\delta_a = \frac{1}{\alpha_1} \left(\frac{1}{10} \right)^{n-1}.$$

Якщо число a має більше ніж дві правильні значущі цифри, то на практиці можна користуватись оцінкою

$$\delta_a = \frac{1}{2\alpha_1} \left(\frac{1}{10} \right)^{n-1}.$$

Навпаки, за відносною похибкою δ можна визначити кількість правильних цифр числа a . Для цього знаходимо граничну абсолютну похибку Δ_a із співвідношення (1.2.1), після чого кількість правильних цифр числа a легко встановлюється.

Приклади

1. Якою буде гранична відносна похибка, якщо замість числа $\sqrt{2}$ взяти 1,41?

Маємо: $n = 3$, $\alpha_1 = 1$. Отже $\delta_a = \frac{1}{1} \left(\frac{1}{10} \right)^{3-1} = 0,01$.

2. Знайти наближено $\sqrt{50}$ так, щоб відносна похибка не перевищувала 0,001 (або 0,1 %).

Маємо: $\alpha_1 = 7$. Отже, повинно справджуватись співвідношення $\delta_a = \frac{1}{7} \left(\frac{1}{10} \right)^{n-1} \leq 0,001$. Звідки $n-1 \geq 3$ і $n \geq 4$.

Отже, щоб відносна похибка не перевищувала 0,001, $\sqrt{50}$ треба взяти не менше як з чотирма правильними цифрами.

3. Число 7432 має граничну відносну похибку $\delta_a = 0,01$ (1%). Скільки в ньому правильних цифр?

Знайдемо граничну абсолютну похибку:

$$\Delta_a = |a|\delta_a = 7432 \cdot 0,01 < 0,75 \cdot 10^2.$$

Отже, число 7432 має лише одну правильну цифру (друга цифра сумнівна). Його треба записати так: $0,74 \cdot 10^4$, або $74 \cdot 10^2$, або $7,4 \cdot 10^3$ і т. д.

Вправи

1. Знайти граничну абсолютну та граничну відносну похибки, якщо замість точного числа \tilde{a} взято наближене число a : а) $\tilde{a} = e$, $a = 2,7$; б) $\tilde{a} = \lg 2$, $a = 0,30$; в) $\tilde{a} = \sqrt{5}$, $a = 2,236$; г) $\tilde{a} = \sqrt{99}$, $a = 9,95$.

2. Знайти граничну відносну похибку числа a :

а) $\tilde{a} = 273,4 \pm 0,1$; б) $\tilde{a} = 82\,900 \pm 500$; в) $\tilde{a} = 0,00126 \pm 0,00003$.

3. Знайти межі точного числа \tilde{a} , якщо відома гранична відносна похибка δ_a числа a : $a = 1427,394$; $\delta_a = 0,00005$; $a = 97,45$; $\delta_a = 0,015$; $a = 389 \cdot 10^5$; $\delta_a = 3\%$.

4. Записати число:

а) π з п'ятьма правильними значущими цифрами;

б) $\sqrt{2}$ з трьома правильними значущими цифрами;

в) $99,995$ з чотирма правильними значущими цифрами.

5. Скільки правильних цифр треба взяти в наближенні числа \tilde{a} , щоб відносна похибка не перевищувала δ :

$\tilde{a} = \sqrt{69}$, $\delta = 0,0001$; $\tilde{a} = \sqrt[4]{100}$, $\delta = 0,01\%$; $\tilde{a} = \lg 127$, $\delta = 1\%$

6. Число a має відносну похибку δ . Скільки в ньому правильних цифр, якщо:

а) $a = 97,432$, $\delta = 0,05\%$;

б) $a = 11,4814$, $\delta = 0,01$;

в) $a = 547,92$, $\delta = 0,001$?

1.3 Загальна формула для обчислення похибки

Нехай потрібно обчислити значення функції $\tilde{y} = f(\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n)$ при заданих значеннях незалежних змінних.

Якщо при цьому замість точних значень $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$, які нам не відомі, підставляються їх наближені значення $\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_n$ (причому $\tilde{x}_i = x_i + \Delta x_i$, $i = 1, 2, \dots, n$), то при обчисленні значення функції виникає похибка

$$\Delta y = f(x_1 + \Delta x_1, x_2 + \Delta x_2, \dots, x_n + \Delta x_n) - f(x_1, x_2, \dots, x_n).$$

Припустивши, що функція $f(x_1, x_2, \dots, x_n)$ диференційовна (достатню кількість разів), Δy можна подати у вигляді:

$$\Delta y = \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_1} \Delta x_1 + \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_2} \Delta x_2 + \dots + \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_n} \Delta x_n \quad (1.3.1)$$

плюс члени другого і вищих порядків відносно Δx_i .

Якщо похибки Δx_i за абсолютною величиною досить малі, то членами другого і вищих порядків у виразі (1.3.1) можна знехтувати. Тоді для абсолютної похибки наближеного значення матимемо:

$$|\Delta y| \leq \sum_{i=1}^n \left| \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} \right| \cdot \Delta x_i, \quad (1.3.2)$$

де через Δx_i позначені граничні абсолютні похибки аргументів. Таким чином,

$$\Delta y = \sum_{i=1}^n \left| \frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i} \right| \cdot \Delta x_i.$$

Для граничної відносної похибки δ_y маємо:

$$\delta_y = \frac{\Delta y}{|y|} = \frac{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right| \cdot \Delta x_i}{|y|} = \sum_{i=1}^n \left| \frac{\frac{\partial f}{\partial x_i}}{f} \right| \Delta x_i,$$

або

$$\delta_y = \sum_{i=1}^n \left| \frac{\partial}{\partial x_i} \ln f(x_1, x_2, \dots, x_n) \right| \Delta x_i. \quad (1.3.3)$$

З формули (1.3.2) випливає, що **гранична абсолютна похибка суми наближених чисел $a_1 + a_2 + \dots + a_n$ дорівнює сумі їх граничних абсолютних похибок**. Для різниці двох чисел $a = a_1 - a_2$ маємо: $\Delta_a = \Delta_{a_1} + \Delta_{a_2}$.

При цьому відносна похибка різниці $\delta_a = \frac{\Delta_{a_1} + \Delta_{a_2}}{|a|}$ може виявитись досить великою, якщо числа a_1 та a_2 між собою мало відрізняються і їх різниця близька до нуля. У цьому випадку для забезпечення потрібної точності треба мати в зменшуваному і від'ємнику достатню кількість цифр, щоб гранична абсолютна похибка різниці a була меншою від самої різниці. Тому, по можливості, слід уникати віднімання близьких чисел.

Для похибки добутку додатних чисел $a = a_1 a_2 \dots a_n$ з (1.3.3) маємо:

$$\delta_a = \frac{\Delta_{a_1}}{a_1} + \frac{\Delta_{a_2}}{a_2} + \dots + \frac{\Delta_{a_n}}{a_n},$$

тобто **гранична відносна похибка добутку кількох наближених додатних чисел, що відрізняються від нуля, дорівнює сумі граничних відносних похибок цих чисел**.

Легко переконатись, що **гранична відносна похибка частки дорівнює сумі граничних відносних похибок діленого і дільника**.

Справді, нехай $a = \frac{a_1}{a_2}$ ($a_1, a_2 > 0$). Тоді з (1.3.3) маємо:

$$\delta_a = \frac{\Delta_{a_1}}{a_1} + \frac{\Delta_{a_2}}{a_2} = \delta_{a_1} + \delta_{a_2}.$$

Можна переконатися також, що відносна похибка k -то степеня числа a у k раз більша, ніж відносна похибка числа a . Аналогічно, відносна похибка кореня k -го порядку з числа a у k раз менша, ніж відносна похибка числа a .

Вправи

1. Знайти суму наближених чисел так, щоб усі цифри результату були значущими (усі цифри доданків значущі):

а) $a_1 = 28,85$; $a_2 = 491,000285$; $a_3 = 1,18294$; $a_4 = 0,18498932$; $a_5 = 9544,2$;

б) $a_1 = 25,4$; $a_2 = 2900,3241$; $a_3 = 9,7 \cdot 10^2$; $a_4 = 0,00097$; $a_5 = 469,99999$.

2. Знайти абсолютну і відносну похибки компонентів дії віднімання, якщо цифри даних правильні: а) $127,434 - 48,815$; б) $281,428 - 125,341$; в) $824,481 - 73,5889$.

3. Знайти: $\sqrt{105} - \sqrt{103}$ з трьома правильними значущими цифрами;

$\sqrt{127} - \sqrt{124}$ з чотирма правильними значущими цифрами.

4. Знайти граничні абсолютну і відносну похибки добутку наближених чисел, добуток і кількість правильних цифр у ньому, якщо цифри множників правильні: $144,8 \cdot 145,12$; $19,875 \cdot 99,9996$; $21,545 \cdot 49,876$; $87,894 \cdot 39,878$.

5. Знайти граничні абсолютну і відносну похибки частки, частку і кількість правильних цифр у частці, якщо всі цифри діленого і дільника правильні: $25,343 \div 34,97$; $43,82 \div 84,42$; $39,8412 \div 481,7$; $32,872 \div 99,435$.

6. Знайти степінь числа, граничні відносну і абсолютну похибки і кількість правильних цифр, якщо всі цифри основи правильні: $(393,4)^2$; $(48.15243)^3$; $(97,4312)^3$.

1.4 Поняття про ймовірну оцінку похибки.

Обчислення без точного врахування похибок

Наближене число a може відхилитися від точного числа \tilde{a} в той чи інший бік на певну величину. Тому наближене число можна розглядати як випадкову величину з математичним сподіванням $M[a] = \tilde{a}$. Сама похибка $\Delta a = \tilde{a} - a$ також є випадковою величиною з математичним сподіванням, що дорівнює нулю, $M[\Delta a] = 0$.

Розглянемо суму наближених чисел $a = a_1 + a_2 + \dots + a_n$. Гранична абсолютна похибка суми $\Delta a = \Delta a_1 + \Delta a_2 + \dots + \Delta a_n$.

Вважатимемо, що випадкові величини $\Delta a_1, \Delta a_2, \dots, \Delta a_n$ мають один і той самий нормальний розподіл ймовірностей з математичним сподіванням, що дорівнює нулю, і середнім квадратичним відхиленням σ . Тоді, як відомо з курсу теорії ймовірностей, математичне сподівання суми випадкових величин дорівнює сумі їх математичних сподівань, тобто $M[\Delta a] = 0$. Дисперсія суми незалежних випадкових величин дорівнює сумі дисперсій цих ве-

личин, тобто $D[\Delta_a] = D[\Delta_{a_1}] + D[\Delta_{a_2}] + \dots + D[\Delta_{a_n}] = n\sigma^2$ (нагадаємо, що за означенням середнє квадратичне відхилення $\sigma[X]$ випадкової величини X дорівнює $\sqrt{D[X]}$, де $D[X]$ - дисперсія випадкової величини X , а тому $\sigma[\Delta_a] = \sigma\sqrt{n}$).

Сума незалежних випадкових величин з нормальним розподілом ймовірностей також має нормальний розподіл ймовірностей. Як відомо, випадкова величина з нормальним розподілом ймовірностей з імовірністю 0,997 відхиляється в той чи інший бік від свого математичного сподівання не більш як на три середніх квадратичних відхилення. Отже, якщо практично $|\Delta_{a_i}| < 3\sigma$ ($P(|\Delta_{a_i}| < 3\sigma) \approx 0,997 \approx 1$), то $|\Delta_{a_i}| < 3\sigma\sqrt{n}$.

Таким чином, абсолютна похибка суми наближених чисел $a = a_1 + a_2 + \dots + a_n$ пропорційна числу \sqrt{n} , а не числу n , як це впливало з формули (1.3.2). Це пояснюється тим, що практично похибки різних чисел можуть частково «гасити» одна одну, в той час як формула (1.3.2) передбачає, так би мовити, крайній випадок.

Аналогічно приходимо до висновку, що й відносна похибка добутку наближених чисел a_1, a_2, \dots, a_n зростає пропорційно числу \sqrt{n} , а не числу n .

Точний підрахунок похибок результатів обчислень наближених чисел досить громіздкий. Тому здебільшого на практиці точно не враховують похибок, а користуються **правилами підрахунку цифр за В. М. Брадїсом**:

1. При додаванні і відніманні наближених чисел молодший збережений розряд результату має бути найбільшим серед розрядів, що виражаються останніми значущими цифрами вихідних даних.

2. При множенні і діленні наближених чисел у результаті потрібно зберегти стільки значущих цифр, скільки їх має наближене дане з найменшою кількістю значущих цифр.

3. При піднесенні до квадрата або до куба в результаті треба зберегти стільки значущих цифр, скільки їх має число, яке підносять до степеня.

4. При добуванні кореня в результаті слід брати стільки значущих цифр, скільки їх у підкореневому виразі.

5. При обчисленні проміжних результатів потрібно брати на одну-дві цифри більше, ніж рекомендують попередні правила.

6. Якщо дані можна брати з довільною точністю, то, щоб знайти результат з k правильними цифрами, дані треба брати з такою кількістю цифр, яка забезпечує $k + 1$ правильну цифру в результаті відповідно до правил 1-4.

7. При обчисленнях одночленних виразів за допомогою логарифмів слід підрахувати кількість значущих цифр у тому наближеному даному, у якому їх найменше, і взяти таблицю логарифмів з кількістю знаків на одиницю більшою. У кінцевому результаті остання цифра відкидається.

Зауважимо, що в основу цих правил покладено *принцип академіка О.М. Крилова* – основний принцип звичайних обчислень, тобто обчислень без строгого врахування похибок: *наближене число слід писати так, щоб у ньому всі цифри, крім останньої, були правильними і лише остання була сумнівною, притому не більш як на одну одиницю.*

1.5 Обернена задача теорії похибок

Досить часто доводиться розв'язувати таку задачу: визначити допустимі похибки аргументів x_1, x_2, \dots, x_n так, щоб гранична абсолютна похибка функції не перевищувала заданої величини Δy .

Користуючись формулою (1.3.2), помічаємо, що ця задача не визначена, оскільки, по-різному вибравши значення Δx_i , можна дістати одне й те саме значення Δy .

Отже, задачу можна розв'язувати кількома способами:

1. Доберемо Δx_i так, щоб в (1.3.2) всі величини $\left| \frac{\partial f}{\partial x_i} \right| \Delta x_i$ були рівними між собою (принцип рівних впливів).

$$\text{Тоді, очевидно, } \left| \frac{\partial f}{\partial x_i} \right| \Delta x_i = \frac{\Delta y}{n} \text{ і } \Delta x_i = \frac{\Delta y}{n \left| \frac{\partial f}{\partial x_i} \right|}.$$

2. Поклавши, що всі Δx_i рівні між собою, дістанемо:

$$\Delta x_i = \frac{\Delta y}{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right|}.$$

3. Вважаючи всі відносні похибки аргументів рівними між собою

$$\frac{\Delta x_1}{|x_1|} = \frac{\Delta x_2}{|x_2|} = \dots = \frac{\Delta x_n}{|x_n|} = \sigma,$$

та підставляючи в (1.3.2) вирази для Δx_i , дістанемо:

$$\Delta y = \delta \sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right| \cdot |x_i|.$$

Визначивши δ , для Δx_i матимемо:

$$\Delta x_i = |x_i| \delta = \frac{|x_i| \Delta y}{\sum_{i=1}^n \left| \frac{\partial f}{\partial x_i} \right| |x_i|}.$$

При розв'язуванні задачі можуть бути використані і деякі інші варіанти.

Аналогічно можна розв'язувати і другу обернену задачу теорії похибок, коли потрібно підібрати граничні абсолютні чи відносні похибки аргумен-

тів так, щоб відносна похибка функції не перевищувала наперед заданої величини.

Вправи

1. Визначити точність, з якою треба виконувати обчислення, щоб знайти результат з точністю до $0,5 \cdot 10^{-3}$:

а) $y = \frac{0,0327x + \sqrt{n}}{0,4231x + 0,9248}$, $x \in [0.7; 2.3]$;

б) $y = \frac{4}{3}\pi R^3$, $R \in [1; 5]$.

2. Корені рівняння $x^2 - 2x + \lg 3 = 0$ треба знайти з чотирма правильними значущими цифрами. З якою точністю треба взяти вільний член рівняння?

II МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ, МОДЕЛЬ І ЕКСПЕРИМЕНТ

Математичне моделювання – це теоретично-експериментальний метод пізнавально-складової діяльності, це метод визначення і пояснення явищ, процесів і систем (об’єктів-оригіналів) на основі створення нових об’єктів математичних моделей. Під математичною моделлю прийнято розуміти сукупність відношень (рівнянь, нерівностей, логічних умов, операторів тощо), які визначають характеристики перебування об’єкта моделювання, а через них і вихідні реакції $Y = \{y_i; i = \overline{1, m}\}$ в залежності від параметрів об’єкта-оригінала W , вхідних впливів $X = \{x_j; j = \overline{1, n}\}$, початкових і граничних умов, а також часу.

Математична модель, як правило, враховує лише ті властивості (атрибути) об’єкта-оригінала W , які відображають, визначають і є цікавими з точки зору цілей і задач конкретного дослідництва.

Відповідно, в залежності від цілей моделювання, при розгляді одного і того ж об’єкта-оригінала W з різних точок зору і в різних аспектах останній може мати різні математичні описи і, як наслідок, бути представленим різними математичними моделями.

Визначення математичної моделі найбільш загально і конструктивно сформулював П.Дж. Косн. Суть його така.

Математична модель – це формальна система, що є кінцевою сукупністю символів і досить суворих правил, визначених цими символами для певного об’єкта із деякими власними відношеннями, символами або константами.

Отже, кінцева сукупність символів (алфавіт) і досить суворі правила оперування цими символами (“граматика” і “синтаксис” математичних виразів) приводять до формування абстрактних математичних об’єктів. Тільки інтерпретація робить абстрактний об’єкт математичною моделлю.

2.1. Інтерпретації в математичному моделюванні

Інтерпретація (лат. “interpretatio” – пояснення, тлумачення) визначається як сукупність значень (змістів), що належать яким-небудь способом елементам деякої системи (теорії), наприклад, формулам і окремим символам. В математичному аспекті інтерпретація – це екстраполяція вихідних положень якої-небудь утримуваної системи, вихідне положення якої визначається незалежно від формальної системи.

Відповідно, можна запевняти, що інтерпретація – це встановлення відповідності між деякою формальною і утримуваною системою. В тих випадках, коли формальна система є застосованою (інтерпретованою) до утримуваної системи, оскільки встановлено, що між елементами формальної системи і елементами утримуваної системи існує взаємно однозначна відповідність, всі вихідні положення формальної системи отримують підтвердження в утримуваній системі. Інтерпретація вважається повною, якщо кожному елементу формальної системи відповідає деякий елемент (інтерпретант) утримуваної системи. Якщо вказана умова порушується, має місце частинна інтерпретація.

При математичному моделюванні в результаті інтерпретації задаються значення елементів математичних виразів (символів, операцій, формул) і цілих конструкцій.

Отже, інтерпретація в математичному моделюванні – інформаційний процес перетворення абстрактного математичного об’єкта (АМО) у відповідну математичну модель (ММ) конкретного об’єкта на основі відображення не пустої інформаційної кількості даних і знань, визначеної АМО і названої областю інтерпретації, в конкретну область – інформаційну кількість даних і знань, визначену предметною областю і об’єктом моделювання і названу областю значень інтерпретації.

Таким чином, інтерпретацію слід розглядати як один із основоположних механізмів (інструментів) технології математичного (наукового) моделювання.

Повний спектр елементів інтерпретації, що відображає перехід від АМО-опису до конкретної ММ, включає чотири види інтерпретації: синтаксичну (структурну), семантичну (змістовну), якісну і кількісну.

2.2 Основні характерні риси моделювання

Метою прикладної математики є математичне сприйняття дійсності [1]. З іншої сторони, інженеру-практику, наприклад, важливіше знати, чи витримає його міст заплановане навантаження. Іншими словами, отримати конкретну відповідь на запитання, а не пориватись до більш вищих цілей.

На практиці вихідним пунктом часто є деяка емпірична ситуація, яка ставить перед дослідником “задачу”, на яку необхідно знайти відповідь. Але вживання таких слів як “задача” і “відповідь” може привести до блуду. Перш за все необхідно встановити, в чому саме суть “задачі”. Це пов’язано з тим, що реальні ситуації рідко бувають чітко окресленими, а складна взаємодія з навколишнім середовищем часто робить точний опис ситуації утрудненим. Часто (але не завжди) паралельно з цією стадією постановки задачі йде процес виявлення основних або суттєвих особливостей явища (рис. 2.1), зокрема для фізичних явищ цей процес схематизації або ідеалізації грає вирішальну роль, оскільки в реальному явищі бере участь множина процесів.

Деякі риси явища є важливими, багато інших – несуттєвими. Розглянемо, наприклад, рух математичного маятника. В даній ситуації суттєвим є реальний характер коливань маятника, а не те, що нитка біла, а тягарець – чорний.

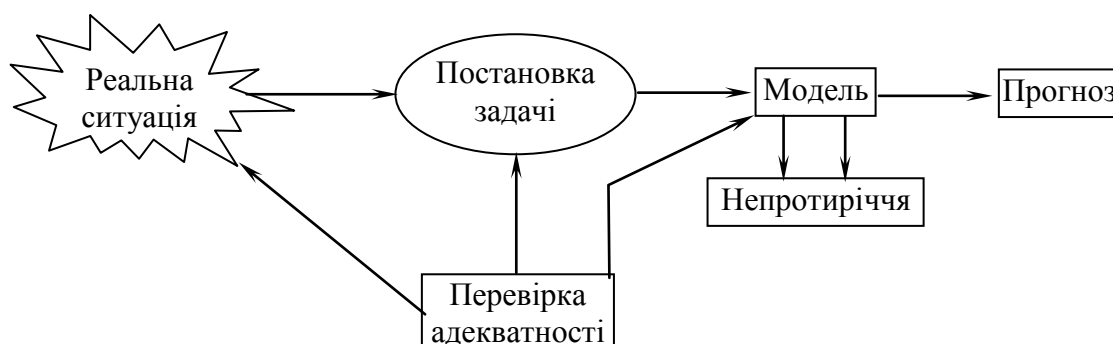


Рисунок 2.1

Після того, як суттєві факти виявлені, наступний крок полягає в переведенні цих факторів на мову математичних понять і величин та постулюванні співвідношень між цими величинами. Як правило, це найважча стадія процесу моделювання.

Після побудови моделі її необхідно перевірити на адекватність. По-перше, сама математична основа моделі повинна мати непротиворіччя і підкорятись всім звичайним законам математичної логіки. Потім справедливість моделі залежить від її здатності адекватно описати відповідну ситуацію.

Математична модель являє собою спрощення реальної ситуації. Спрощення наступають тоді, коли несуттєві особливості ситуації відкидаються і складна задача зводиться до ідеалізованої задачі, яка піддається математичному аналізу. Саме при такому підході в класичній прикладній механіці виникли блоки без тертя, ідеальні стержні, невагомні нерозтяжні нитки, нев’язкі рідини та багато інших понять. Ці поняття не існують в реальній

дійсності, вони є абстракціями. Але їх часто можна з успіхом вважати задовільним наближенням до реальної ситуації.

Це лише одна сторона спрощення. Друга сторона пов'язана в порівнянні порядку різноманітних величин, які фігурують в моделі. Наприклад, зміна деякої величини x протягом часу можна описати рівнянням виду:

$$m\ddot{x} + v\dot{x} + cx = 0.$$

Нехай в результаті спостережень було зазначено, що член $v\dot{x}$ набагато більший за величиною, ніж cx . В такому випадку можна зекономити багато часу, спростивши рівняння (відкинувши член cx), в результаті чого розв'язок знайдеться швидше, і він буде правильно відображати ситуацію. Зауважимо, що більш точне розв'язання даного рівняння в математичному відношенні в дійсності може привести до неправильних висновків.

Після отримання формул або результатів необхідно здійснити зворотний переклад з математичної мови на мову, якою спочатку формулювалась задача. Тобто, необхідно чітко зрозуміти математичний зміст отриманих розв'язків і те, що вони означають на мові реального світу.

Критерієм достовірності моделі служить інженерна або технічно-економічна корисність нової інформації, отриманої за математичною моделлю при наступній перевірці її в системі об'єкта-прототипу. Щоб модель виявилась корисною, вона повинна мати дві основні властивості.

По-перше, модель повинна бути економічнішою (дешевшою, швидкіснішою, безпечнішою тощо), ніж об'єкт-прототип.

По-друге, модель повинна забезпечувати можливість розповсюдження без суттєвих витрат.

При математичному моделюванні використовуються практично всі розділи математики. В основі вибору того чи іншого математичного методу повинні лежати: ціль моделювання, теоретичні уявлення про об'єкт, накопичення інформації про нього, обмеження на матеріальні і часові ресурси тощо.

2.2.1 Основні типи математичних моделей

З врахуванням існуючої практики розв'язання інженерних задач в області будівельного матеріалознавства і технології виділяють п'ять основних типів математичних моделей. В подальшому будемо їх позначати, відповідно, M_1, M_2, \dots, M_5 [2].

Моделі M_1 . В інженерних розрахунках використовується перехід від нескінченно малих до кінцевих змін в реальній системі. Але такий перехід, як правило, настільки аналітично складний, що інтегрування замінюється наближеними методами. В кінці XIX – на початку XX століття сформувався спеціальний науковий напрямок – теорія подібності, яка поєднала диференціальні рівняння з експериментальною інформацією про систему. В

зв'язку з розвитком ЕОМ головну роль почали грати чисельні методи розв'язання задач математичної фізики.

Зауважимо, що при складанні таких моделей – рівнянь математичної фізики – використовуються фундаментальні закони природи. Наприклад, рівняння Нав'є-Стокса для нестисливої рідини, що відображає закон збереження енергії:

$$\rho \frac{DW}{Dt} = \rho g - \text{grad} p + \mu \Delta W ,$$

де $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ – оператор Лапласа;

W – вектор швидкості руху поля потоку рідини;

ρg – масова сила, яка віднесена до одиниці об'єму (g – вектор прискорення вільного падіння в полі земного тяжіння);

D/Dt – похідна субстанції (для стаціонарного руху в напрямку осі x за-

мість $\rho DW/Dt$ можна записати $\rho \frac{\partial W}{\partial x} \cdot \frac{dx}{dt} = \rho W \frac{\partial W}{\partial x}$, де $\frac{\partial W}{\partial x}$ – є

зміна швидкості при переході від однієї точки до іншої);

P – тиск рідини.

Моделі M_2 . Мають в своїй основі перш за все деяку інженерну концептуальну модель, яка виражається на початку словами в термінах даної науки (матеріалознавство, технологія). Ця концептуальна модель описується далі абстрактно-знаково за допомогою диференціальних або алгебраїчних рівнянь, геометричних співвідношень, логічних операцій тощо. На рис. 2.2 розглянуто реологічне тіло Максвелла з періодом релаксації, який дорівнює відношенню між ньютонівською в'язкістю η_2 і модулем пружності Гука G_1 [3]: $\theta = \eta_2/G_1$.

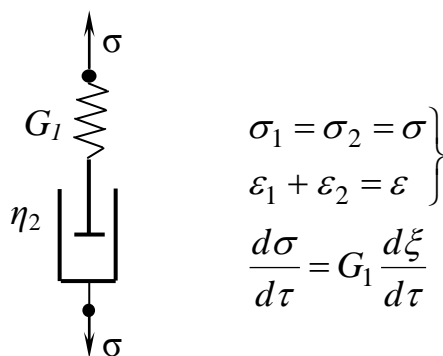


Рисунок 2.2

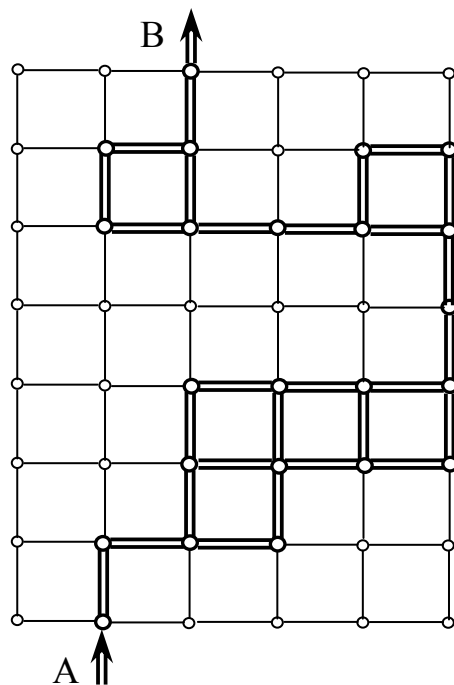
В основі математичної моделі лежить уява про те, що поведінка реального матеріалу під навантаженням аналогічна поведінці послідовно з'єднаної пружини і поршня, який рухається у в'язкій рідині. Перевага мо-

делі M_2 полягає у “фізичній” інтерпретації технологом, або ця інтерпретація вже закладена в концептуальній основі моделі. Недолік моделі M_2 полягає в тому, що описати математичними символами можна концепцію, яка далека від реальності. Для використання моделей M_2 необхідно визначати практично різні параметри об’єкта (наприклад, в’язкість η_2 і пружність G_1 конкретних рідин тощо).

Моделі M_3 . Описують з відомою точністю зв’язок між входами X і виходом Y системи на основі експериментальних даних без аналізу внутрішньої структури системи, яку можна розглядати як “чорний ящик” [4]. Модель M_3 відрізняється універсальністю методології збирання експериментальної інформації і будови моделей. До недоліків варто віднести складність різносторонньої інтерпретації параметрів моделей M_3 і обмеженість області їх дії з відповідною точністю рамок тих конкретних об’єктів, для яких вони побудовані. Моделі M_3 і надалі будуть займати основне місце в інженерній практиці в зв’язку з розширенням випуску та використанням мікроЕОМ і персональних комп’ютерів.

Моделі M_4 . Досліджують операцію як сукупність дій, направлених на досягнення деякої мети, характерні, головним чином, для техніко-економічних задач (управління запасами, розподілення ресурсів між підприємствами, задачі заміни обладнання і контролю якості тощо).

На рис. 2.3 показано використання моделі M_4 : фільтрація рідини через бетон розглядається як рух через сітку випадкової конфігурації з запірними вузлами [5].



$$\tau(A - B) = \sum_i \tau_i = \sum_i l_i V_i \rightarrow m_{1k}$$

Рисунок 2.3

Імітаційна модель M_5 включає в єдину програму для ЕОМ моделі різних типів, які об’єднані правилами взаємопереходу, а також генератор випадкових чисел, який генерує збурювальні дії середовища ξ на систему

(рис. 2.4). На сьогодні в різноманітних галузях науки створюються імітаційні моделі, до складу яких входять програмні блоки, які забезпечують активний діалог людини з ЕОМ [6].

Математичні моделі є ефективним засобом розв'язання задач у всіх основних областях діяльності інженерів-матеріалознавців; технологів – при дослідженні і проектуванні систем, при управлінні їх функціонуванням. В задачах аналізу системи, який направлений на пізнання внутрішнього механізму явищ, які в ній відбулися, особливе значення мають моделі M_1 і M_2 , природа яких забезпечує поглиблену інтерпретацію результатів моделювання. В той же час в задачах управління системою, коли головною задачею є досягнення конкретної мети функціонування системи в конкретних умовах, пріоритетними стають моделі M_3 - M_5 . При проектуванні системи на етапі пошукового синтезу перевага надається моделям M_1 і M_2 , які дають більш загальну інженерну інформацію про систему деякого класу, а на стадії підготовки синтезованої системи до реалізації підвищується значення моделей M_3 , M_4 і особливо M_5 .



Рисунок 2.4 Принципова схема управління системою і її зв'язки із зовнішнім середовищем

Робота з моделями всіх класів суттєво спрощується, якщо можливе використання їх графічного відображення.

Так однофакторна модель відображається лінією, яка може бути гладкою або кусково-гладкою. Двофакторна може бути відображена сімейством однофакторних функцій $Y = f_1(x_1)$ при деяких фіксованих значеннях іншої змінної $X_2 = C = \text{const}$ (і навпаки). Але найбільш зручне її відображення – площинна діаграма, яка складається з ліній рівного рівня відгуку системи (ізоліній). Ізолінії створюються як лінії перерізу поверхні відгуку з горизонтальними площинами, які проведені через відповідні мітки вертикальної шкали Y . Ізолінії з поверхні відгуку проектуються на площину факторного простору і створюють діаграму.

2.3 Планування експерименту

Будь-який експеримент, яким би він не був складним, закінчується поданням результатів, формулюванням висновків і, можливо, дачею рекомендацій іншим особам. Ця інформація може надаватися у вигляді графіків або кривих, математичних формул або номограм, таблиць, статичних даних або словесних описів. Із графіків отримується залежність результату R від змінної x або залежність R від x і y у випадку параметричних кривих. Для отримання залежності R від x , y і z необхідно побудувати ізометричні координати. Графічне зображення більш складних функцій неможливе, оскільки людина не в змозі наочно уявити більш складні співвідношення. Записуючи результати у вигляді формул, можна виразити залежність R від більшої кількості змінних, але лише в небагатьох експериментах одночасно досліджується більше трьох незалежних змінних.

Статичний показник може давати інформацію про всю сукупність даних і про змінність окремих елементів сукупності. Він може давати інформацію про значимість причинного співвідношення або вказувати ймовірність появи конкретної події в майбутньому на основі попереднього дослідження.

Подання результатів експериментів в словесній формі завжди було проблемою при проведенні наукових дослідів. Це найнеефективніший спосіб подання результатів, але його не можна ігнорувати повністю. Безумовно, результати деяких експериментів, які проводяться в сучасних фізичних лабораторіях, просто неможливо подати в словесній формі. В техніці такі експерименти зустрічаються порівняно рідко, і значна частина технічного звіту звичайно містить словесні описи і пояснення.

Більшість інженерних експериментів веде до конкретної дії – прийняття рішення, продовження дослідів або визнання невдач.

Найбільш складною проблемою є правильне формулювання питань, пов'язаних з побудовою плану експерименту.

План експерименту – це загальний термін. Він визначає набір інструкцій до проведення експерименту, в яких вказується послідовність робіт, характер і величина змін змінних і даються вказівки щодо проведення повторних експериментів.

Неможливо уявити собі експеримент, в якому основним продуктом не були б дані, які записані в тій або іншій формі. В простих експериментах кінцеві дані можуть бути досить елементарними і не містити взагалі будь-якої помилки. З іншого боку, у виключно складних експериментах може бути низька точність, а кінцевий результат буде містити великі помилки.

При дослідженні теоретичних питань, пов'язаних з проведенням експериментів, важливу роль грає вивчення помилок і невизначеностей. Ні один

експеримент неможливо правильно спланувати без детального вивчення цього важливого фактора.

Експериментатор, в якій би області він не працював, майже завжди при-тримується такої послідовності: на початку відбувається планування, потім додають обладнання, після цього відбуваються досліди і накінець виконується аналіз та складається звіт. При плануванні і придбанні обладнання аналіз помилок повинен бути на першому плані.

Якщо технічні обмеження приводять до досить великої невизначеності або фінансові можливості не дозволяють придбати прилади високої точності, то проведення експерименту неможливе.

План експерименту є компактним і ефективним, якщо завчасно встановлюються інтервали між значеннями змінних. Якщо нас цікавить функціональне співвідношення між незалежною змінною x і залежною змінною y , то цю функцію можна зобразити у вигляді деякої кривої (або прямої) в системі координат (x, y) . Така крива складається з нескінченної кількості окремих точок, з яких вибирається деяка кінцева необхідна кількість точок, що характеризують дану функцію. Якщо розглядається функція двох незалежних змінних, то повній сукупності даних відповідає деяка площа.

Таким чином, вибір кінцевої сукупності експериментальних точок є необхідним і важливим етапом планування.

Наступним етапом після проведення експерименту є обробка результатів.

2.3.1 Підбір емпіричних формул та згладжування

При експериментальному (емпіричному) вивченні функціональної залежності однієї величини y від другої величини x проводять ряд вимірювань величини y при різних значеннях величини x . Задача полягає в аналітичному поданні шуканої функціональної залежності, тобто в підборі формули, що описує результати експерименту. Особливість задачі полягає в тому, що наявність випадкових помилок вимірювань роблять нерозумним підбір такої формули, яка точно описувала б всі експериментальні значення. Іншими словами, графік шуканої функції не повинен проходити через всі точки, а повинен, по можливості, згладжувати “шум” (випадкові помилки вимірювань). Зрозуміло, що згладжування “шуму” буде тим більш точним і надійним, чим більша кількість проведених експериментів, тобто чим більше ми маємо надлишкової інформації.

Наприклад, для проведення прямої $y = ax + b$ досить двох точок – (x_1, y_1) і (x_2, y_2) , якщо ці точки відомі точно. Але при наявності більш або менш значного „шуму”, як на рис.2.5, для тієї ж мети може знадобитися декілька десятків точок.

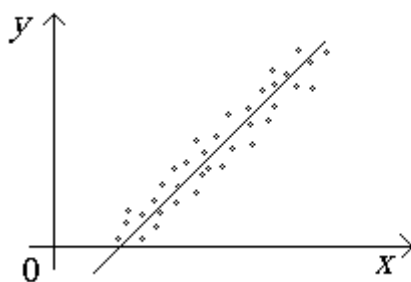


Рисунок 2.5

Емпіричну формулу звичайно вибирають з формул конкретного типу, наприклад:

$$y = ax + b, \quad y = a \cdot e^{bx} + c, \quad y = a + h \sin(\omega x + \varphi) \quad (2.3.1)$$

Іншими словами, задача зводиться до визначення параметрів a , b , c формули (2.3.1), в той же час як вигляд формули відомий наперед з яких-небудь теоретичних міркувань або з міркувань простоти аналітичного подання емпіричного матеріалу.

Позначимо вибрану функціональну залежність через

$$y = f(x; a_0, a_1, \dots, a_n) \quad (2.3.2)$$

з явним показом всіх параметрів, які необхідно визначити. Ці параметри $x; a_0, a_1, \dots, a_n$ не можна визначити точно за емпіричними значеннями функції y_1, y_2, \dots, y_n , тому що останні містять випадкові помилки. Мова йде тільки про отримання “достатньо добрих” оцінок шуканих параметрів. Метод найменших квадратів дозволяє отримати задовільні оцінки всіх параметрів a_0, a_1, \dots, a_n . Див. підрозділ 3.4.

Але не завжди, наприклад, з фізичного аналізу задачі відомий вигляд формули, щоб потім визначити невідомі параметри. Іноді доводиться мати діло з задачами, для яких попередній фізичний аналіз не дає точного вигляду формули або дозволяє вибрати формули з досить широкого кола. Крім того, для простоти подальших розрахунків часто віддають перевагу залежності у вигляді многочлена, але при цьому степінь многочлена не задається наперед.

Розглянемо деякі рекомендації щодо вибору формул серії задач, які часто зустрічаються [10].

Якщо кількість експериментальних точок велика, то підбір емпіричної формули може виявитися досить важким: формули з малим числом параметрів можуть давати більші спотворення, а велике число параметрів не зручне для аналізу. Для аналізу важливо ліквідувати “шум” експерименту, зберігаючи інформацію про істинність функції. Для цієї мети використовується згладжування емпіричних даних, тобто заміна даної таблиці експериментальних точок іншою таблицею близьких до них точок, які лежать на достатньо гладкій кривій.

Згладжування відбувається за допомогою многочленів (бажано оптимального степеня), які наближають за методом найменших квадратів вибрані групи експериментальних точок. Оскільки найкраще згладжування отримується для середніх точок (коли враховується інформація про поведінку функції по обидва боки від згладжуваної точки), то кількість точок для згладжування вибирають непарну, а групи точок – ковзними вздовж всієї таблиці: беруть, наприклад, перші п'ять точок y_1, y_2, y_3, y_4, y_5 і згладжують за їх допомогою середню точку y_3 , потім беруть наступну групу y_2, y_3, y_4, y_5, y_6 і згладжують точку y_4 , і так далі до кінця таблиці. При цьому залишаються декілька кратних точок (наприклад, дві на початку і дві в кінці таблиці), які згладжуються з меншою точністю.

Нижче наводяться найбільш вживані з простих формул згладжування для таблиць з постійним кроком. В цих формулах прийняті такі позначення. Середній точці групи приписується індекс 0, симетричні точки отримують при цьому індекси $\pm 1, \pm 2, \dots$

Згладжувані значення позначаються хвилькою зверху. Основною формулою слугує формула згладжування середньої точки, тобто формула для \tilde{y}_0 , інші формули використовуються тільки на краях таблиці.

2.3.2 Лінійне згладжування

Лінійним згладжуванням називається згладжування многочленом першого степеня.

Формули лінійного згладжування за трьома точками:

$$\tilde{y}_0 = \frac{1}{3}(y_{-1} + y_0 + y_1),$$

$$\tilde{y}_{-1} = \frac{1}{6}(5y_{-1} + 2y_0 - y_1),$$

$$\tilde{y}_1 = \frac{1}{6}(-y_{-1} + 2y_0 + 5y_1).$$

Формули лінійного згладжування за п'ятьма точками:

$$\tilde{y}_0 = \frac{1}{5}(y_{-2} + y_{-1} + y_0 + y_1 + y_2),$$

$$\tilde{y}_{-1} = \frac{1}{10}(4y_{-2} + 3y_{-1} + 2y_0 + y_1),$$

$$\tilde{y}_1 = \frac{1}{10}(y_{-1} + 2y_0 + 3y_1 + 4y_2),$$

$$\tilde{y}_{-2} = \frac{1}{5}(3y_{-2} + 2y_{-1} + y_0 - y_2),$$

$$\tilde{y}_2 = \frac{1}{5}(-y_{-2} + y_0 + 2y_1 + 3y_2).$$

2.3.3 Нелінійне згладжування

З множини формул нелінійного згладжування наведемо тільки формули згладжування многочленів третього степеня за сімома точками:

$$\tilde{y}_0 = \frac{1}{21}(-2y_{-3} + 3y_{-2} + 6y_{-1} + 7y_0 + 6y_1 + 3y_2 - 2y_3),$$

$$\tilde{y}_{-1} = \frac{1}{42}(-4y_{-3} + 16y_{-2} + 19y_{-1} + 12y_0 + 2y_1 - 4y_2 + y_3),$$

$$\tilde{y}_1 = \frac{1}{42}(y_{-3} - 4y_{-2} + 2y_{-1} + 12y_0 + 19y_1 + 16y_2 - 4y_3),$$

$$\tilde{y}_{-2} = \frac{1}{42}(8y_{-3} + 19y_{-2} + 16y_{-1} + 6y_0 - 4y_1 - 7y_2 + 4y_3),$$

$$\tilde{y}_2 = \frac{1}{42}(4y_{-3} - 7y_{-2} - 4y_{-1} + 6y_0 + 16y_1 + 19y_2 + 8y_3),$$

$$\tilde{y}_{-3} = \frac{1}{42}(39y_{-3} + 8y_{-2} - 4y_{-1} - 4y_0 + y_1 + 4y_2 - 2y_3),$$

$$\tilde{y}_3 = \frac{1}{42}(-2y_{-3} + 4y_{-2} + y_{-1} - 4y_0 - 4y_1 + 8y_2 + 39y_3).$$

Формули згладжування многочленами більш високих степенів майже не використовуються, а формули згладжування за більшим числом точок використовуються досить рідко, тому що вони залишають погано згладженими досить велику кількість точок по краях таблиці.

Зауважимо, що результат згладжування сильно залежить від використаних формул, вибір яких підказується швидше інтуїцією, ніж якими-небудь правилами.

2.3.4 Лінійна кореляція і регресія

При вивченні залежності між двома величинами, кожна з яких піддається випадковому розсіюванню (безконтрольному розподілу), використовуються методи кореляційного аналізу. Кореляційний аналіз вивчає усереднений закон поведінки кожної з величин в залежності від значень іншої величини, а також міру залежності між розглянутими величинами. Підставляючи кожному значенню однієї величини, наприклад, x , середнє з відповідних значень другої величини, наприклад, y , ми отримуємо функцію регресії або просто регресію y на x . Аналогічно визначається регресія величини x на величину y . Функція регресії зображається графічно лінією регресії. Міра залежності між величинами характеризується коефіцієнтом кореляції або кореляційними відношеннями.

Кореляція між величинами x та y називається лінійною, якщо обидві функції регресії лінійні. В цьому випадку лінії регресії перетворюються в пря-

мі регресії. Кутові коефіцієнти цих прямих виражаються через коефіцієнт кореляції, який служить також мірою лінійної залежності між величинами.

Коефіцієнтом кореляції ρ між випадковими величинами x та y називається математичне сподівання добутку їх нормованих відхилень:

$$\rho = M\left(\frac{x-a}{\sigma_x} \cdot \frac{y-b}{\sigma_y}\right),$$

де $a = Mx$ і $b = My$ – центри розподілення величин x і y , σ_x^2 і σ_y^2 – їх дисперсії.

Коефіцієнт кореляції ρ може бути також записаний в одній з таких форм:

$$\rho = \frac{1}{\sigma_x \sigma_y} M(x-a)(y-b) = \frac{M(xy) - M_x M_y}{\sigma_x \sigma_y}.$$

Величина $M(x-a)(y-b)$ називається кореляційним моментом.

Коефіцієнт кореляції – це безрозмірна величина, його абсолютна величина не перевищує одиницю: $|\rho| \leq 1$.

Для незалежних величин x і y коефіцієнт кореляції дорівнює нулеві. Якщо для залежних величин коефіцієнт кореляції дорівнює нулю, то в цьому випадку величини x і y називаються некорельованими.

Коефіцієнт кореляції характеризує міру лінійної залежності між величинами x і y . Пояснимо це. Нехай $Ax + B$ – лінійна функція середньоквадратичного наближення до величини y , тобто така лінійна функція, для якої математичне сподівання $M[y - (Ax + B)]^2$ досягає найменшого значення.

Для експериментального вивчення залежності між двома величинами x і y проводять деяку кількість n незалежних випробувань (дослідів, спостережень). Результат i -го дослідження дає пару значень x_i, y_i ($i = \overline{1, n}$). За цими значеннями визначаються точкові оцінки як середніх значень, так і коефіцієнта кореляції.

Незмінними і самостійними оцінками теоретичних середніх значень a і b служать емпіричні середні значення:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i. \quad (2.3.3)$$

Незмінними і самостійними оцінками дисперсії σ_x^2 і σ_y^2 служать емпіричні дисперсії:

$$S_x^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n x_i^2 - (\bar{x})^2 n \right],$$

$$S_y^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y})^2 = \frac{1}{n-1} \left[\sum_{i=1}^n y_i^2 - (\bar{y})^2 n \right]. \quad (2.3.4)$$

Для оцінки кореляційного моменту служить емпіричний кореляційний момент:

$$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1} \left[\sum_{i=1}^n x_i y_i - \bar{x} \bar{y} n \right].$$

За цими оцінками будують емпіричний коефіцієнт кореляції:

$$r = \frac{1}{S_x S_y} \cdot \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}). \quad (2.3.6)$$

Зауважимо, що емпіричний коефіцієнт кореляції, як і теоретичний, за абсолютною величиною не перевищує одиниці.

Коефіцієнт r не змінюється при зміні початку відліку і масштабу вимірювання величин x і y . Ця властивість дозволяє суттєво спростити обчислення за допомогою вибору зручного початку відліку (x_0, y_0) і одиниць масштабу.

Після заміни:

$$x = x_0 + h_1 u, \quad y = y_0 + h_2 v \quad (h_1, h_2 > 0),$$

тобто $u = \frac{x - x_0}{h_1}$, $v = \frac{y - y_0}{h_2}$, емпіричний коефіцієнт кореляції обчислюється за формулою:

$$r = \frac{\sum u_i v_i - n \bar{u} \bar{v}}{\sqrt{\sum u_i^2 - n(\bar{u})^2} \cdot \sqrt{\sum v_i^2 - n(\bar{v})^2}}, \quad (2.3.7)$$

де $\bar{u} = \frac{1}{n} \sum u_i$, $\bar{v} = \frac{1}{n} \sum v_i$, $\sum = \sum_{i=1}^n$.

Приклад розрахунку коефіцієнта кореляції

Проведено $n = 26$ вимірювань значень x і y . В перших трьох стовпцях таблиці знаходяться різні значення вимірювань (x_i, y_i) з кількістю вимірювань m_i , в яких зустрічалась кожна точка (x_i, y_i) . Для величини x вибрано початок відліку $x_0 = 26$, масштабний коефіцієнт $h_1 = 0,5$. Для величини y , відповідно, $y_0 = 0,5$ і $h_2 = 0,01$. Таким чином, величини $u_i = \frac{x_i - 26}{0,5}$, $v_i = \frac{y_i - 0,5}{0,01}$ наближають тільки цілі значення. При підрахунку сум, які входять в формулу (2.3.4), кожний доданок враховується стільки раз, скільки він зустрічається в таблиці, тому для розрахунку введені стовпці $u m, u^2 m$ і $v m, v^2 m, u v m$. В другому рядку умовно вказаний порядок дій. В останньому рядку підраховані суми, які необхідні для розрахунків:

$$\begin{aligned} \sum m_i &= n = 26, & \sum u_i v_i &= -26, & \sum u_i^2 m_i &= 248, \\ \sum v_i m_i &= 22, & \sum v_i^2 m_i &= 104, & \sum u_i v_i m_i &= 87. \end{aligned}$$

Зауважимо, що тут суми \sum результатів беруться не по всіх $n = 26$ вимірюваннях, як в формулі (2.3.7), а лише по різних точках (x_i, y_i) , що і викликає необхідність врахування множників m_i . Наприклад, повна сума $\sum_{i=1}^n u_i v_i$ тут записується у вигляді $\sum u_i v_i m_i$, а середнє значення – $\bar{u} = \frac{1}{n} \sum_{i=1}^n u_i = \frac{1}{n} \sum u_i m_i$.

Для контролю розрахунків рекомендується повторити розрахунок з іншим початком відліку.

За отриманими даними знаходимо середнє значення $\bar{u} = -1$, $\bar{v} = \frac{22}{26} = 0,846$ і емпіричний коефіцієнт кореляції:

$$r = \frac{87 - 26 \cdot (-1) \cdot 0,846}{\sqrt{248 - 26 \cdot (-1)^2} \cdot \sqrt{104 - 26 \cdot (0,846)^2}} = 0,793.$$

Таблиця 2.1

x	y	m	u	um	$U^2 m$	v	vm	$v^2 m$	Uvm
(1)	(2)	(3)	(4)	(5)=(3)(4)	(6)=(4)(5)	(7)	(8)=(3)(7)	(9)=(7)(8)	(10)=(5)(7)
23,0	0,48	2	-6	-12	72	-2	-4	8	24
24,0	0,50	4	-4	-16	64	0	0	0	0
24,5	0,49	3	-3	-9	27	-1	-3	3	9
24,5	0,50	2	-3	-6	18	0	0	0	0
25,0	0,51	1	-2	-2	4	1	1	1	-2
25,5	0,52	1	-1	-1	1	2	2	4	-2
26,0	0,49	2	0	0	0	-1	-2	2	0
26,0	0,51	1	0	0	0	1	1	1	0
26,0	0,53	2	0	0	0	3	6	18	0
26,5	0,50	1	1	1	1	0	0	0	0
26,5	0,52	1	1	1	1	2	2	4	2
27,0	0,54	2	2	4	8	4	8	32	16
27,0	0,52	1	2	2	4	2	2	4	4
28,0	0,53	3	4	12	48	3	9	27	36
Суми	26	-	-26	248	-	22		104	87

2.3.5 Рівняння прямих регресії

Пряма регресії y на x має рівняння:

$$y - b = \rho \frac{\sigma_y}{\sigma_x} (x - a), \quad (2.3.8)$$

де $a, b, \sigma_x, \sigma_y, \rho$ – параметри.

Пряма регресії x на y має рівняння:

$$x - a = \rho \frac{\sigma_x}{\sigma_y} (y - b). \quad (2.3.9)$$

Коефіцієнти

$$\beta_{y/x} = \rho \frac{\sigma_y}{\sigma_x}, \quad \beta_{x/y} = \rho \frac{\sigma_x}{\sigma_y}$$

називаються коефіцієнтами регресії y на x та x на y , відповідно. Обидва коефіцієнти регресії $\beta_{y/x}$ і $\beta_{x/y}$ мають той же знак, що і коефіцієнт кореляції ρ . Обидві прямі регресії проходять через точку (a, b) – центр розподілення. Прямі регресії y на x та x на y збігаються тільки тоді, коли $|\rho| = 1$, тобто у випадку лінійної функціональної залежності між величинами x та y .

За результатами експерименту параметри рівнянь (2.3.8) і (2.3.9) оцінюються за допомогою формул (2.3.3, 2.3.4, 2.3.6).

Ці оцінки визначають емпіричні прямі регресії. Емпірична пряма регресії y на x має рівняння:

$$y - \bar{y} = r \frac{S_y}{S_x} (x - \bar{x}), \quad (2.3.10)$$

причому

$$b_{y/x} = r \frac{S_y}{S_x} = \frac{\sum_{i=1}^n x_i y_i - n \bar{x} \bar{y}}{\sum_{i=1}^n x_i^2 - n(\bar{x})^2} = \frac{\sum_{i=1}^n u_i v_i - n \bar{u} \bar{v}}{\sum_{i=1}^n u_i^2 - n(\bar{u})^2} \cdot \frac{h_2}{h_1}$$

є емпіричним коефіцієнтом регресії y на x .

Параметри лінійної функції (2.3.10) задовольняють принцип найменших квадратів за y : сума квадратів відхилень спостережуваних значень y_i від розрахункових за рівнянням прямої регресії (2.3.10) менше, ніж сума квадратів відхилень їх від будь-якої іншої прямої. Таким чином, має місце нерівність:

$$\sum_{i=1}^n \left[y_i - \bar{y} - r \frac{S_y}{S_x} (x_i - \bar{x}) \right]^2 \leq \sum_{i=1}^n [y_i - (Ax_i + B)]^2,$$

або, з врахуванням (2.3.7), маємо:

$$\sum_{i=1}^n \left[y_i - \bar{y} - r \frac{S_y}{S_x} (x_i - \bar{x}) \right]^2 = (n-1) S_y^2 (1-r^2)$$

Емпірична пряма регресії x на y має рівняння:

$$x - \bar{x} = r \frac{S_x}{S_y} (y - \bar{y}),$$

$$\text{де } b_{x/y} = r \frac{S_x}{S_y} = \frac{\sum_{i=1}^n x_i y_i - n\bar{x}\bar{y}}{\sum_{i=1}^n y_i^2 - n(\bar{y})^2} = \frac{\sum_{i=1}^n u_i v_i - n\bar{u}\bar{v}}{\sum_{i=1}^n v_i^2 - n(\bar{v})^2} \cdot \frac{h_1}{h_2}$$

є емпіричним коефіцієнтом регресії x на y . Для цієї прямої принцип найменших квадратів виконується відносно величини x : сума квадратів відхилень спостережуваних значень x_i від розрахункових за рівнянням прямої регресії менше, ніж сума квадратів відхилень їх від будь-якої іншої прямої. Таким чином, має місце нерівність:

$$\sum_{i=1}^n \left[x_i - \bar{x} - r \frac{S_x}{S_y} (y_i - \bar{y}) \right]^2 \leq \sum_{i=1}^n [x_i - (Ay_i + B)]^2.$$

$$\text{При цьому } \sum_{i=1}^n \left[x_i - \bar{x} - r \frac{S_x}{S_y} (y_i - \bar{y}) \right]^2 = (n-1)S_x^2(1-r^2).$$

На рис. 2.6 і рис. 2.7 показано, які відхилення маютья на увазі при розгляді прямих регресії y на x і x на y .

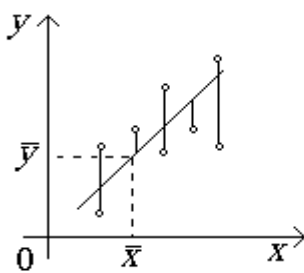


Рисунок 2.6

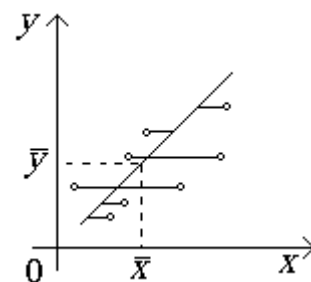


Рисунок 2.7

Обидві емпіричні прямі регресії проходять через центр емпіричного розподілення – точку (\bar{x}, \bar{y}) . Прямая регресії x на y завжди проходить крутіше (з більшим нахилом до осі x), ніж пряма регресії y на x .

Емпіричні коефіцієнти регресії $b_{y/x}$ і $b_{x/y}$ мають той же знак, що і емпіричний коефіцієнт кореляції r , і пов'язані співвідношенням $b_{y/x} \cdot b_{x/y} = r^2$.

Приклад Визначимо рівняння емпіричних прямих регресії попереднього прикладу. Маємо:

$$\bar{x} = 26 + 0,5 \cdot (-1) = 25,5,$$

$$\bar{y} = 0,5 + 0,01 \cdot 0,846 = 0,50846.$$

Емпіричний коефіцієнт регресії y на x :

$$b_{y/x} = \frac{87 - 26 \cdot (-1) \cdot 0,846}{248 - 26 \cdot (-1)^2} \cdot \frac{0,01}{0,5} = 0,00978.$$

Емпіричний коефіцієнт регресії x на y :

$$b_{y/x} = \frac{87 - 26 \cdot (-1) \cdot 0,846}{101 - 26 \cdot (0,846)^2} \cdot \frac{0,5}{0,01} = 63,8;$$

$$y - 0,508 = 0,00978 \cdot (x - 25,5),$$

$$x - 25,5 = 63,8 \cdot (y - 0,508).$$

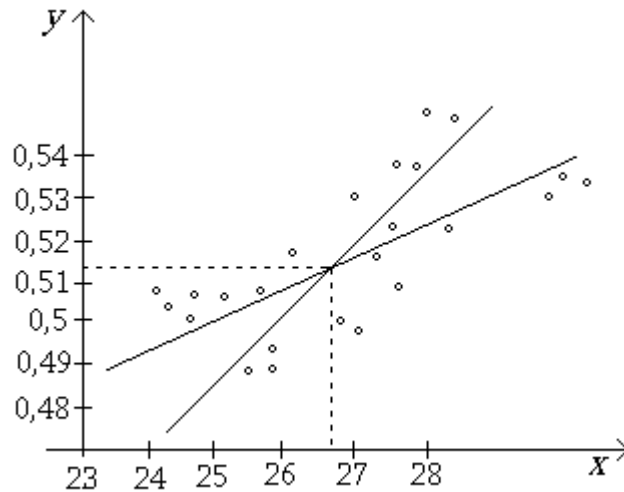


Рисунок 2.8

2.4 Інтерполяція функцій однієї і декількох змінних

Повним описом функціональної залежності відповідних результатів випробувань (вимірів) вважають задання формули або нескінченний набір чисел. Розв'язання задач на ЕОМ відбувається тільки з кінцевою сукупністю чисел. Виникає необхідність охарактеризувати функцію кінцевим набором чисел.

За методом сіток функції описують їх значеннями в кінцевому числі k -х точок. Це означає, що на деякому інтервалі $[a, b]$ у вузлах сітки маємо відповідні значення функції $f(x_1, x_2, \dots, x_k)$ ($\bar{k} = \bar{1}, \bar{n}$).

2.4.1 Одновимірна інтерполяція

На практиці найбільш часто необхідно відновити функцію за її сітковим поданням. Це роблять за допомогою інтерполяції.

Задача одновимірної інтерполяції полягає в побудові такої функції, що $f(x_k) = y_k$ для всіх k і при цьому $f(x)$ повинна приймати розумні значення для

X , які містяться між заданими точками. Критерій розумності для кожної задачі свій і, напевне, ніколи точного визначення не буде мати.

Найбільш важливим класом інтерполяційних функцій є множина алгебраїчних поліномів. Задача поліноміальної інтерполяції полягає в побудові полінома

$$P_{n-1}(x; f) = P_{n-1}(x; x_1; \dots; x_n; f_1; \dots; f_n)$$

степені $n-1$, який приймає у вузлах x_1, x_2, \dots, x_n значення f_1, f_2, \dots, f_n . З курсу математичного аналізу відомо, що будь-яка неперервна функція може бути представлена на кінцевому інтервалі $[a, b]$ поліномом P_n степені n , а через n точок можна провести єдиний поліном степені $n-1$.

Відновлення коефіцієнтів полінома $P_{n-1}(x; f)$ можна провести, розв'язуючи систему n рівнянь відносно невідомих лінійних коефіцієнтів a_k ($k=0, n-1$). Але в багатьох випадках рівняння недостатньо обумовлені. Найбільш задовільний спосіб визначення полінома, який інтерполює точки (x_k, f_k) , полягає у використанні базису лагранжевих поліномів на множині (x_k) . Це такі поліноми $l_j(x)$ ($j=0, n-1$) степені $n-1$, що

$$l_j(x) = \begin{cases} 1, & \text{якщо } j = k; \\ 0, & \text{якщо } j \neq k. \end{cases}$$

Ці умови задовольняє поліном

$$l_j(x) = \frac{(x - x_1) \cdots (x - x_{j-1})(x - x_{j+1}) \cdots (x - x_n)}{(x_j - x_1) \cdots (x_j - x_{j-1})(x_j - x_{j+1}) \cdots (x_j - x_n)}. \quad (2.4.1)$$

Кожний множник чисельника (2.4.1) обертається в нуль при $k \neq j$. Відповідні множники знаменника нормують поліномом так, що $l_j(x_j) = 1$. Поліном $l_j(x)f_j$ приймає значення f_j в точці x_j і дорівнює нулеві у всіх точках x_k ($k \neq j$). Таким чином, інтерполяційний поліном степені $n-1$, який проходить через n точок (x_k, f_k) , виражається формулою

$$P_{n-1}(x, f) = \sum_{j=1}^n l_j(x) f(x_j) \quad (2.4.2)$$

і називається інтерполяційним поліномом Лагранжа.

Похибка поліноміальної інтерполяції на інтервалі $[a, b]$ оцінюється за формулою:

$$f(x) - P_{n-1}(x, f) = \max_{[a, b]} |f^{(n)}(x)| \frac{(x - x_1) \cdots (x - x_n)}{n!}, \quad (2.4.3)$$

де $f^{(n)}$ - похідна порядку n .

Якщо похідні високих порядків великі за абсолютним значенням, то точність апроксимації, згідно з (2.4.3), буде малою, тому немає потреби вибирати велику кількість вузлів. Крім того, при підвищенні степені полінома (2.4.2) результат через погану обумовленість стає досить чутливим до помилок округлення. При недостатній функції відрізок $[a, b]$ розбивають на

часткові інтервали і на кожному з них використовують інтерполяцію невисокого порядку. Наприклад, якщо відомо, що $f(x)$ має обмежену тільки другу похідну $|f''(x)| \leq M_2$, то бажано обмежитись лінійною інтерполяцією. Тоді похибка на рівномірній сітці з кроком h буде не перевищувати $M_2 h^2/8$. Таким чином, процес кусково-лінійної інтерполяції збігається при $h \rightarrow 0$ зі швидкістю h^2 , тоді як процес інтерполяції в цілому може розбігатись.

Для наближеного відновлення функції за її значеннями на сітці використовують, крім поліномів, інші функції. Наприклад, для подання періодичних функцій використовують тригонометричну інтерполяцію.

Приклад

За дослідними даними залежність водопоглинання w від температури опалення T становить

$T, ^\circ\text{C}$	1000	1200	1300
$w, \%$	15	7	2

визначити водопоглинання для проміжного значення $T=1250^\circ\text{C}$.

Розв'язання: Інтерполяційний поліном Лагранжа для триточкової інтерполяції такий:

$$P_2 = \frac{(x-x_2)(x-x_3)}{(x_1-x_2)(x_1-x_3)} f_1 + \frac{(x-x_1)(x-x_3)}{(x_2-x_1)(x_2-x_3)} f_2 + \frac{(x-x_1)(x-x_2)}{(x_3-x_1)(x_3-x_2)} f_3. \quad (2.4.4)$$

Якщо підставити значення $x=T=1250$, $x_1=1000$, $x_2=1200$, $x_3=1300$, $f_1=w_1=15$, $f_2=w_2=7$ і $f_3=w_3=2$, отримаємо:

$$P_2(1250) = \frac{(1250-1200)(1250-1300)}{(1000-1200)(1000-1300)} \cdot 15 + \frac{(1250-1000)(1250-1300)}{(1200-1000)(1200-1300)} \cdot 7 + \frac{(1250-1000)(1250-1200)}{(1300-1000)(1300-1200)} \cdot 2 = 4,6.$$

Отже, при $T=1250^\circ\text{C}$ водопоглинання $w=4,6\%$.

2.4.2 Інтерполяція сплайнами

Розглянемо випадок кусково-поліноміальної інтерполяції, коли між сусідніми вузлами сітки функції інтерполюються кубічним поліномом (кубічним сплайном) [7]. Його коефіцієнти на кожному інтервалі визначаються з умови спряження у вузлах:

$$\begin{aligned} f_i &= y_i, \\ f'(x_i - 0) &= f'(x_i + 0), \\ f''(x_i - 0) &= f''(x_i + 0), i = \overline{1, n-1}. \end{aligned}$$

Крім цього, при $x = x_0$ і $x = x_n$ висувається гранична умова:

$$f''(x_0) = 0, f''(x_n) = 0. \quad (2.4.5)$$

Будемо шукати кубічний поліном у вигляді:

$$f(x) = a_i + \vartheta_i(x - x_{i-1}) + c_i(x - x_{i-1})^2 + d_i(x - x_{i-1})^3, \quad (2.4.6)$$

$$x_{i-1} \leq x \leq x_i.$$

З умови $f_i = y_i$ маємо:

$$\begin{aligned} f(x_{i-1}) &= a_i = y_{i-1}, \\ f(x_i) &= a_i + b_i h_i + c_i h_i^2 = y_i, \\ h_i &= x_i - x_{i-1}, \quad i = \overline{1, n-1}. \end{aligned} \quad (2.4.7)$$

Визначимо похідні:

$$\begin{aligned} f'(x) &= \vartheta_i + 2c_i(x - x_{i-1}) + 3d_i(x - x_{i-1})^2, \\ f''(x) &= 2c_i + 6d_i(x - x_{i-1}) \end{aligned}$$

і запишемо умову їх неперервності при $x = x_i$:

$$\begin{aligned} \vartheta_{i+1} &= \vartheta_i + 2c_i h_i + 3d_i h_i^2, \\ c_{i+1} &= c_i + 3d_i h_i, \quad i = \overline{1, n-1}. \end{aligned} \quad (2.4.8)$$

Загальна кількість невідомих коефіцієнтів дорівнює $4n$, але кількість рівнянь (2.4.7) і (2.4.8) дорівнює $4n-2$. Інші два рівняння отримаємо з умови (2.4.5) при $x = x_0$ і $x = x_n$: $c_1 = 0, c_n + 3d_n h_n = 0$.

Виразивши з (2.4.8) $d_i = (c_{i+1} - c_i)/3h_i$ та підставляючи його в (2.4.7) з врахуванням $a_i = y_{i-1}$, отримаємо:

$$\begin{aligned} \vartheta_i &= \frac{(y_i - y_{i-1})}{h_i} - \frac{1}{3} h_i (c_{i+1} + 2c_i), \quad i = \overline{1, n-1}, \\ \vartheta_n &= (y_n - y_{n-1})h_n - \frac{2}{3} h_n c_n. \end{aligned} \quad (2.4.9)$$

З першого рівняння (2.4.8) і (2.4.9) та виразу d_i отримуємо для визначення c_i різницеве рівняння другого порядку:

$$\begin{aligned} h_i c_i + 2(h_i + h_{i+1})c_{i+1} + h_{i+1}c_{i+2} &= 3 \left(\frac{y_{i+1} - y_i}{h_{i+1}} - \frac{y_i - y_{i-1}}{h_i} \right), \\ i &= \overline{1, n-1} \end{aligned} \quad (2.4.10)$$

з граничними умовами:

$$c_1 = 0, c_{n+1} = 0. \quad (2.4.11)$$

Різницеве рівняння (2.4.10) з умовами (2.4.11) розв'язується методом прогонки.

Використаємо кубічну сплайн-інтерполяцію для розв'язання попереднього прикладу.

Оскільки необхідно визначити значення $f(1250)$, то кубічний поліном будемо шукати у вигляді

$$f(x) = a_2 + \vartheta_2(x - x_1) + c_2(x - x_1)^2 + d_2(x - x_1)^3 \quad (2.4.12)$$

на рівномірній сітці з кроком $h = 100$ на інтервалі $1200 = x_1 \leq x \leq x_2 = 1300$.

Коефіцієнт $a_2 = y_1 = 7$.

Різницеве рівняння (2.4.10) запишеться у вигляді:

$$(c_1 + 4c_2 + c_3)h^2 = 3(y_2 - 2y_1 + y_0).$$

З врахуванням (2.4.11) $c_1 = 0, c_3 = 0$, тоді $4c_2 = \frac{3}{100^2}(2 - 14 + 15)$. Звідки

$$c_2 = \frac{9}{40000}. \text{ Тоді } d_2 = -\frac{9}{12000000}.$$

Визначимо за формулою (2.4.9) значення e_2 :

$$e_2 = \frac{y_2 - y_1}{100} - \frac{2}{3}hc_2 = \frac{2 - 7}{100} - \frac{200}{3} \cdot \frac{9}{40000} = -\frac{5}{100} - \frac{3}{200} = -\frac{13}{200}.$$

В результаті кубічний поліном набуде вигляду:

$$f(x) = 7 - \frac{13}{200}(x - 1200) + \frac{9}{40000}(x - 1200)^2 - \frac{9}{12000000}(x - 1200)^3.$$

Звідки $f(1250) = 4,22$.

Інтерполяція використовується також в задачі оберненого інтерполювання: задана таблиця $y_i = y(x_i)$; знайти x_i як функцію від y_i . Прикладом оберненого інтерполювання може служити задача про знаходження коренів рівняння.

Інтерполяційні формули використовуються також при обчисленні інтегралів, при написанні різницевої апроксимації для диференціальних рівнянь на основі інтегральних тотожностей.

2.4.3 Багатовимірна інтерполяція

В різних додатках широко використовують двовимірні, тривимірні таблиці. Наприклад, теплофізичні властивості різноманітних речовин залежать від температури і тиску, а оптичні характеристики – ще і від довжини хвилі випромінювання.

При багатовимірній інтерполяції через громіздкість таблиць необхідно брати достатньо великі кроки по аргументах, тобто сітка вузлів, на якій будують таблицю, отримується достатньо грубою. Тому необхідно вводити перетворення змінних $\eta = \eta(y)$, $\xi = \xi(x)$, $\varphi = \varphi(z)$, підбираючи відповідні формули. При вдалому виборі таких формул можна використовувати в нових змінних інтерполяційний поліном невисокого степеня.

Здійснюючи багатовимірну інтерполяцією, необхідно пам'ятати, що розміщення вузлів не може бути довільним. Наприклад, при інтерполяції полінома першого степеня $P_1(x, y)$ вузли не повинні лежати на одній прямій в площині. Дійсно, визначник системи трьох рівнянь

$$z_i = a + vx_i + cy_i, \quad i = \overline{1,3} \quad (2.4.13)$$

записується у вигляді:

$$\Delta = x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2). \quad (2.4.14)$$

Умови розміщення трьох точок на одній прямій виглядають таким чином:

$$\frac{x_2 - x_1}{y_2 - y_1} = \frac{x_3 - x_2}{y_3 - y_2}.$$

Після простих перетворень маємо:

$$x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2) = 0, \quad (2.4.15)$$

тобто, якщо вузли лежать на одній прямій, то визначник дорівнює нулю і побудувати поліном $P_1(x, y)$ вигляду (2.4.15) неможливо. Перевіряти умови потрібного типу достатньо складно, тому на практиці намагаються будувати регулярні сітки, як правило, прямокутні і рівномірні, коли вузли є точками перетину двох взаємно перпендикулярних систем паралельних прямих. На цій сітці проводять просту послідовну інтерполяцію спочатку за рядками, а потім за стовпцями.

При послідовній інтерполяції зростає степінь інтерполяційного полінома. При трикутній конфігурації розміщення вузлів степінь многочлена буде мінімальною. Многочлен n -ої степеня у формі Ньютона в даному випадку можна уявити як узагальнення одновимірного варіанта запису:

$$P_n(x, y) = \sum_{i=0}^n \sum_{j=0}^{n-1} Z(x_0, \dots, x_i, y_0, \dots, y_j) \prod_{p=0}^{i-1} (x - x_p) \prod_{q=0}^{j-1} (y - y_q). \quad (2.4.16)$$

Приклад

Записати многочлен Ньютона першого і другого степеня для двовимірної інтерполяції функції $z = z(x, y)$.

З (2.4.16) отримаємо:

$$P_1(x, y) = z(x_0, y_0) + z(x_0, y_0, y_1)(y - y_0) + z(x_0, x_1, y_0)(x - x_0),$$

$$z(x_0, y_0, y_1) = \frac{z(x_0, y_0) - z(x_0, y_1)}{y_0 - y_1}, \quad z(x_0, x_1, y_0) = \frac{z(x_0, y_0) - z(x_1, y_0)}{x_0 - x_1}.$$

$$P_2(x, y) = z(x_0, y_0) + z(x_0, y_0, y_1) \cdot (y - y_0) + z(x_0, y_0, y_1, y_2) \cdot (y - y_0) \cdot (y - y_1) + z(x_0, x_1, y_0) \cdot (x - x_0) + z(x_0, x_1, y_0, y_1) \cdot (x - x_0) \cdot (y - y_0) + z(x_0, x_1, x_2, y_0) \cdot (x - x_0) \cdot (x - x_1).$$

В деяких випадках використовують нерегулярні сітки. Тоді обмежуються інтерполяційним поліномом першого степеня $z = a + vx + cy$ і його коефіцієнти знаходять за трьома вузлами:

$$z_i = a + vx_i + cy_i, \quad i = \overline{1,3}.$$

Коефіцієнти a, v, c не обов'язково визначати, оскільки перший стовпець є лінійною комбінацією трьох інших стовпців. В результаті з цих стовпців можна скласти визначник, який буде дорівнювати нулю. Розглядаючи цей визначник за першим стовпцем, отримаємо залежність $z = z(x, y)$.

III ЕЛЕМЕНТИ ЧИСЕЛЬНИХ МЕТОДІВ

3.1 Розв'язання системи лінійних алгебраїчних рівнянь

методом простої ітерації

3.1.1 Загальні положення

Розглядається система n лінійних рівнянь з n невідомими

$$A\bar{x} = \bar{b}, \quad (3.1.1)$$

де

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \bar{b} = \begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{bmatrix}.$$

Для застосування методу простої ітерації використовується зведена система

$$\bar{x} = \alpha\bar{x} + \bar{\beta}, \quad (3.1.2)$$

де

$$\alpha = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}, \bar{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{bmatrix}.$$

Перехід від системи (3.1.1) до зведеної системи (3.1.2) здійснюється таким чином:

$$\beta_i = \frac{b_i}{a_{ii}}; \alpha_{ij} = -\frac{a_{ij}}{a_{ii}}, \quad (3.1.3)$$

$$i, j = 1, 2, \dots, n; i \neq j,$$

$$\alpha_{ij} = 0, i = j. \quad (3.1.4)$$

Довільне $(k-1)$ -е наближення обчислюється за формулою:

$$x_i^{(k+1)} = \beta_i + \sum_{j=1}^n \alpha_{ij} x_j^{(k)}, \quad (3.1.5)$$

або в розгорнутому вигляді

$$x_i^{(0)} = \beta_i, \quad (3.1.6)$$

$$\bar{x}^{(k+1)} = \bar{\beta} + \alpha\bar{x}^{(k)}$$

$$i = 1, 2, \dots, n; k = 0, 1, 2, \dots \quad (3.1.7)$$

Процес ітерації (3.1.5) стосовно зведеної системи (3.1.2) збігається незалежно від початкового наближення, якщо будь-яка канонічна норма матриці α менше 1, тобто $|\alpha| < 1$.

Оцінка похибки методу простої ітерації здійснюється за формулою:

$$\|\bar{x} - \bar{x}^{(k-1)}\| \leq \frac{\|\alpha\|}{1 - \|\alpha\|} \|x^{(k-1)} - \bar{x}^{(k-1)}\| = \Delta^{(k)}. \quad (3.1.8)$$

Якщо необхідно обчислити значення коренів з точністю ε , то, як тільки $\Delta \leq \varepsilon$, виконується умова $|x - x^{(k+1)}| < \varepsilon$, тобто $x^{(k+1)}$ – шукане рішення.

Алгоритм методу:

1. Вводимо $\alpha, \beta, \|\alpha\|$.
2. Визначаємо $x^{(k)}$, коли $k=0$, за (3.1.6).
3. Обчислюємо $x^{(k+1)}$ за (3.1.7).
4. Обчислюємо $\Delta^{(k)}$ за (3.1.8).
5. Перевіряємо умову $\Delta^{(k)} \leq \varepsilon$. Якщо умова не виконується, беремо $k=k+1$ і переходимо до п.3, у протилежному випадку – переходимо до п.6.
6. Виводимо $x^{(k+1)}$ на друк або на монітор.

3.1.2 Приклад

Знайти розв'язок системи алгебраїчних рівнянь за методом простої ітерації з точністю $\varepsilon=10^{-5}$, де

$$A = \begin{bmatrix} 10 & -1 & 2 & -3 \\ 1 & 10 & -1 & 2 \\ 2 & 3 & 20 & -1 \\ 3 & 2 & 1 & 20 \end{bmatrix}, \quad \bar{b} = \begin{bmatrix} 0 \\ 5 \\ -10 \\ 15 \end{bmatrix}.$$

Переходимо до зведеної системи, яка має вигляд (3.1.2):

$$\begin{cases} x_1 = 0,1x_2 - 0,2x_3 + 0,3x_4 \\ x_2 = -0,1x_1 + 0,1x_3 - 0,2x_4 + 0,5 \\ x_3 = -0,1x_1 - 0,15x_2 - 0,05x_4 - 0,5 \\ x_4 = -0,15x_1 - 0,1x_2 - 0,05x_3 + 0,75. \end{cases}$$

Звідки:

$$\alpha = \begin{bmatrix} 0 & 0,1 & -0,2 & 0,3 \\ -0,1 & 0 & 0,1 & -0,2 \\ -0,1 & -0,15 & 0 & 0,05 \\ -0,15 & -0,1 & -0,05 & 0,17 \end{bmatrix}, \quad \bar{\beta} = \begin{bmatrix} 0 \\ 0,5 \\ -0,5 \\ 0,75 \end{bmatrix}.$$

Перевіряємо умову збіжності $|\alpha| < 1$. Знайдемо, наприклад, l -норму матриці α :

$$\|\alpha\|_1 = \max_1 \sum_{j=1}^n |\alpha_{ij}|.$$

Маємо:

$$\|\alpha\|_1 = \max(0,35; 0,35; 0,35; 0,55) = 0,55 < 1.$$

Отже, процес ітерації для даної системи є збіжним.

За початкове наближення візьмемо

$$\bar{x}_0 = \bar{\beta} = \begin{bmatrix} 0 \\ 0,5 \\ -0,5 \\ 0,75 \end{bmatrix}.$$

Далі процес обчислення проводимо у відповідності з наведеним алгоритмом до досягнення заданої точності.

3.1.3 Текст програми розрахунку

```

10 REM МЕТОД ІТЕРАЦІЙ
11 DIM X(4) , A(4, 4) , XI(4) , B(4) , XP(4) , AL(4, 4) , BT(4)
12 INPUT E
20 PRINT " ВВЕДІТЬ СТОВПЕЦЬ B(I) "
25 FOR I=1 TO 4
30 INPUT B(I)
40 NEXT I
50 FOR I=1 TO 4 : FOR J=1 TO 4
60 PRINT "A("I" , "J")="
70 INPUT A(I ,J) : NEXT J : NEXT I
71 FOR I=1 TO 4 : BT(I)=B(I)/A(1, 1) : XI(I)=BT(I)
72 RI=RI+XI(I)
80 NEXT I
90 FOR I=1 TO 4 : FOR J=1 TO 4
92 AL(I, J)=-A(I, J)/A(I, I)
93 IF I=J THEN AL (I, J)=0
95 XP(I)=XP(I)+AL(I, J)*XI(J)
96NEXT J : NEXT I
97FOR I=1 TO 4
100 X(I)=XP(I)+BT(I)
101 RR=RR+X(I)
104 XI(I)=X(I) : XP(I)=0 : B(I)=0
105 NEXT I
110 IF ABS (RI-RR) < E GOTO 133
111 RI=RR
120 RR=0
130 GOTO 90
133 FOR I=1 TO 4 : PRINT "X("I")=" ; X(I) , "XI("I")=" ; XI(I) : NEXT I
136 REM ПЕРЕВІРКА РОЗРАХУНКУ

```

```

140 I=1 TO 4 : FOR J=1 TO 4 : B(I)=B(I)+A(I, J)*X(J) : NEXT J : LPRINT "B("I")=" ; B(I) :
NEXT I
150 END

```

В програмі: стовпець $B(I) \rightarrow \bar{\beta}$; елементи матриці $A(I, J) \rightarrow a_{ij}$; $BT(I) \rightarrow vi/\alpha_{ij}$; $AL(I, J) \rightarrow \alpha_{ij}$; $XI(I) \rightarrow xi^{(0)}$; $XP(I) \rightarrow xi^{(K+1)}$; $RI \rightarrow X$; $RR \rightarrow X^{(K+1)}$.
 E – точність розрахунку

3.2 Обчислення дійсного кореня рівняння з даною точністю із застосуванням комбінованого методу

3.2.1 Загальні положення

Розглядається рівняння виду:

$$f(x) = 0, \quad (3.2.1)$$

де функція $f(x)$ визначена та неперервна на якомусь кінцевому або нескінченному інтервалі.

Процес знаходження ізольованих дійсних коренів рівняння (3.2.1) складається з двох етапів: відділення коренів та обчислення кореня рівняння з даною точністю.

Комбінований метод хорд та дотичних для обчислення дійсного кореня рівняння (3.2.1) полягає в одночасному застосуванні методів хорд та дотичних, при цьому на кожному кроці здобуємо значення кореня за відповідними формами.

Нехай корінь є відділеним на відрізку $[a, b]$ та виконані умови, які гарантують єдиність кореня на $[a, b]$ та збіжність методу: $f(a) \times f(b) < 0$; $f'(x)$ і $f''(x)$ – неперервні, збігаються і мають постійні знаки на відрізку $[a, b]$.

Введемо такі позначення:

\bar{x}_n – значення наближеного кореня, здобутого за методом дотичних;

x_n – значення кореня знаходять за формулами:

$$x_{n+1} = x_n - \frac{f(x_n)}{f(\bar{x}_n) - f(x_n)} (\bar{x}_n - x_n); \quad (3.2.2)$$

$$\bar{x}_{n-1} = \bar{x}_n - \frac{f(\bar{x}_n)}{f'(\bar{x}_n)}, \quad n = 0, 1, 2, \dots \quad (3.2.3)$$

Початкове наближення $x_0 \in [a, b]$ для методу дотичних вибирають у відповідності з умовою:

$$\begin{aligned} \bar{x}_0 = a, \quad \text{якщо} \quad f(a)f''(a) > 0, \\ \text{або} \quad \bar{x}_0 = b, \quad \text{якщо} \quad f(b)f''(b) > 0. \end{aligned} \quad (3.2.4)$$

За початкове наближення $x_0 \in [a, b]$ для методу хорд тоді приймається $x_0 = b$, або $x_0 = a$, відповідно.

Процес обчислення кореня закінчується, коли виконується умова:

$$|x_{n+1} - \bar{x}_{n+1}| < \varepsilon, \quad (3.2.5)$$

де ε – дана точність.

За наближене значення кореня рівняння приймається:

$$\xi = \frac{x_{n+1} + \bar{x}_{n+1}}{2}. \quad (3.2.6)$$

Алгоритм методу містить такі етапи:

1. Вводимо початкове наближення $x_n, \bar{x}_n, n = 0$.
2. Обчислюємо наближення x_{n+1}, \bar{x}_{n+1} у відповідності з формулами (3.2.2) та (3.2.3).
3. Перевіряємо виконання умови (3.2.5). Якщо умова не виконується, припускаємо $n = n + 1$ і переходимо до п.2. В іншому випадку переходимо до п.4.
4. Обчислюємо значення кореня у відповідності з формулою (3.2.6).
5. Виводимо ξ на друк або на монітор.

3.2.2 Приклад

Обчислити корінь рівняння $f(x) = x^2 + 4\sin x = 0$ з точністю до $\varepsilon = 10^{-3}$ комбінованим методом.

Графічно відділяємо корені. Для цього задане рівняння записуємо у вигляді $x^2 = -4\sin x$. Будуємо графік функцій $y_1 = x^2$ та $y_2 = -4\sin x$ (рис.3.1).

Точний корінь рівняння $x = 0$. Другий точний корінь рівняння $\xi \in [-1; -2]$, відрізок $[-1; -2]$ – інтервал ізоляції кореня.

Перевіряємо умови, що гарантують єдиність кореня на $[-1; -2]$ і збіжність методу:

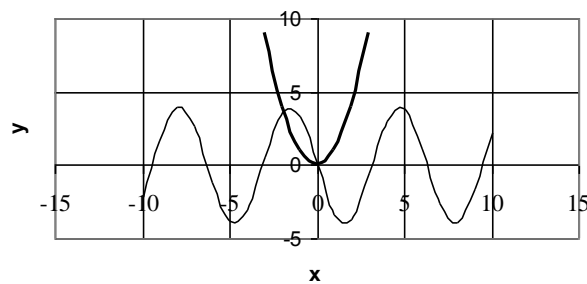


Рисунок 3.1

$$\begin{aligned} f(-1) &= -2,36; f(-2) = 0,36; \\ f'(x) &= 2x + 4\cos x; f'(-1) = -1,46; \\ f'(-2) &= -5,66; f''(x) = 2 - 4\sin x; \\ f''(-1) &= 5,36; \\ f''(-2) &= 5,64; \\ f(1)f(2) &< 0. \end{aligned}$$

Тоді маємо:

$f'(x)$ неперервна на $[-1;-2]$ і не змінює знак

($f'(x) > 0$ неперервна на $[-1;-2]$).

$f''(x)$ неперервна на $[-1;-2]$ і не змінює знак

($f''(x) > 0$ неперервна на $[-1;-2]$).

У відповідності з умовою (3.2.4) за початкове наближення для методу дотичних беремо $x_0 = -2$, для методу хорд $x_0 = -1$ ($f(-2)f''(-2)$).

Далі процес обчислення кореня здійснюється відповідно з наведеним алгоритмом до досягнення заданої точності ε .

3.2.3 Текст програми розрахунку

```
10 REM МЕТОД ДОТИЧНИХ ТА ХОРД
20 PRINT " ВВЕДІТЬ ПОЧАТКОВІ ЗНАЧЕННЯ XP , XO , E "
25 INPUT XP , XO , E
30 DEF FNA (X)=X^2+4*SIN(X)
40 DEF FNP (X)=X*2+4*COS(X)
50 D=XP- XO
60 XN=XO-(FNA (XO) / (FNA (XP)-FNA (XO)))*D
70 XR=XP-FNA (XP) / FNP (XP)
80 IF ABS (XN-XR) < E GOTO 100
90 XP=XR : XO=XN : GOTO 50
100 X=(XN+XR) / 2
110 LPRINT " НАБЛ. КОРИНЬ= " ; X , " ЗНАЧ. F("X")=" ; FNA (X)
120 END
НАБЛ. КОРИНЬ = -1.933832      ЗНАЧ. F(-1.933832)=4.155636E-04
```

Для запуску програми необхідно уточнити:

FNA(X) - вираз функції $f(x)$;

FNP(X) - вираз похідної від $f(x)$;

XP - початкове наближене значення кореня за методом дотичних;

XO - початкове наближене значення кореня за методом хорд;

E – точність розрахунку.

3.3 Наближене розв'язання систем нелінійних

рівнянь методом Ньютона

3.3.1 Загальні положення

Розглядається система n лінійних рівнянь з n невідомими:

$$\begin{cases} f_1(x_1, x_2, \dots, x_n) = 0 \\ f_2(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_n(x_1, x_2, \dots, x_n) = 0. \end{cases} \quad (3.3.1)$$

У матричному вигляді система (3.3.1) запишеться як:

$$f(\bar{x}) = 0, \quad \bar{x} = \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}, \quad \bar{f} = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_n \end{bmatrix}. \quad (3.3.2)$$

Для розв'язання системи нелінійних рівнянь (3.3.2) застосовується метод послідовних наближень, в якому значення кореня на $(k+1)$ -у кроці визначається таким чином:

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} + \bar{\Delta}^{(k)}, \quad (3.3.3)$$

де $\bar{\Delta}^{(k)}$ – вектор поправок.

Суть методу Ньютона полягає у тому, що розв'язання системи нелінійних рівнянь зводиться до знаходження послідовності розв'язків на кожному кроці системи n лінійних рівнянь відносно поправок $\bar{\Delta}^{(k)}$, в якій матриця коефіцієнтів є матрицею Якобі:

$$\bar{f}'(x) = W(\bar{x}) = \begin{bmatrix} \frac{\partial f_i}{\partial x_j} \end{bmatrix}, \quad i, j = 1, n: \quad (3.3.4)$$

$$W(\bar{x}^{(k)})\bar{\Delta}^{(k)} = -f(\bar{x}^{(k)}),$$

де

$$\bar{f}(\bar{x}^{(k)}) = \begin{bmatrix} f_1(\bar{x}^{(k)}) \\ f_1(\bar{x}^{(k)}) \\ \dots \\ f_n(\bar{x}^{(k)}) \end{bmatrix}, \quad (3.3.5)$$

$$W(\bar{x}^{(k)}) = \begin{vmatrix} \left(\frac{\partial f_1}{\partial x_1}\right)^{(k)} & \left(\frac{\partial f_1}{\partial x_2}\right)^{(k)} & \dots & \left(\frac{\partial f_1}{\partial x_n}\right)^{(k)} \\ \left(\frac{\partial f_2}{\partial x_1}\right)^{(k)} & \left(\frac{\partial f_2}{\partial x_2}\right)^{(k)} & \dots & \left(\frac{\partial f_2}{\partial x_n}\right)^{(k)} \\ \dots & \dots & \dots & \dots \\ \left(\frac{\partial f_n}{\partial x_1}\right)^{(k)} & \left(\frac{\partial f_n}{\partial x_2}\right)^{(k)} & \dots & \left(\frac{\partial f_n}{\partial x_n}\right)^{(k)} \end{vmatrix}. \quad (3.3.6)$$

За нульове наближення $\bar{x}^{(0)}$ беремо грубе значення кореня.

Процес обчислення кореня системи (3.3.2) продовжується до виконання однієї з умов:

$$\|\bar{f}^{(k+1)}\| < \varepsilon, \quad (3.3.7)$$

$$\|\bar{x}^{(k+1)} - \bar{x}^{(k)}\| < \varepsilon. \quad (3.3.8)$$

У разі виконання (3.3.7) або (3.3.8) за корінь системи (3.3.2) приймаємо $\bar{x} = \bar{x}^{(k+1)}$.

Алгоритм методу Ньютона містить такі етапи:

1. Беремо $k=0$. Вводимо початкове наближення $\bar{x}^{(k)}$ та точність ε .
2. Обчислюємо вектор правих частин системи лінійних рівнянь $\bar{f}(\bar{x}^{(k)})$ у відповідності до (3.3.5).
3. Обчислюємо матрицю Якобі $W(\bar{x}^{(k)})$ у відповідності до (3.3.6).
4. Розв'язуємо систему лінійних рівнянь (3.3.4).
5. Обчислюємо $\bar{x}^{(k+1)}$ у відповідності до (3.3.3).
6. Перевіряємо умову (3.3.7) або (3.3.8). Якщо умова виконана, беремо $k=k+1$ і переходимо до п.2, інакше – до п.7.
7. Виводимо $\bar{x}^{(k+1)}$, кількість ітерацій.

3.3.2 Приклад

Знайти розв'язок системи нелінійних рівнянь з точністю до $\varepsilon = 10^{-3}$.

$$\begin{cases} f_1(x_1, x_2, x_3) = x_1 + x_1^2 - 2x_2x_3 - 0,1 = 0 \\ f_2(x_1, x_2, x_3) = x_2 - x_2^2 + 3x_1x_3 + 0,2 = 0 \\ f_3(x_1, x_2, x_3) = x_3 + x_3^2 + 2x_1x_2 - 0,3 = 0. \end{cases}$$

Початкове наближення:

$$\bar{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Складаємо матрицю Якобі:

$$W(\bar{x}) = \begin{vmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} \end{vmatrix} = \begin{vmatrix} 1+2x_1 & -2x_3 & -2x_2 \\ 3x_3 & 1-2x_2 & 3x_1 \\ 2x_2 & 2x_1 & 1+2x_3 \end{vmatrix}.$$

Складаємо систему лінійних рівнянь вигляду (3.3.4) для початкового наближення.

Маємо:

$$\bar{f}(\bar{x}) = \begin{bmatrix} x_1 + x_1^2 - 2x_2x_3 - 0,1 \\ x_2 - x_2^2 + 3x_1x_3 + 0,2 \\ x_3 + x_3^2 + 2x_1x_2 - 0,3 \end{bmatrix},$$

$$\bar{f}(\bar{x}^{(0)}) = \begin{bmatrix} x_1^{(0)} + (x_1^{(0)})^2 - 2x_2^{(0)}x_3^{(0)} - 0,1 \\ x_2^{(0)} - (x_2^{(0)})^2 + 3x_1^{(0)}x_3^{(0)} + 0,2 \\ x_3^{(0)} + (x_3^{(0)})^2 + 2x_1^{(0)}x_2^{(0)} - 0,3 \end{bmatrix} = \begin{bmatrix} -0,1 \\ 0,2 \\ -0,3 \end{bmatrix},$$

$$W(\bar{x}^{(0)}) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Система лінійних рівнянь для $k=0$ має вигляд:

$$\begin{cases} \Delta_1^{(0)} = -0,1 \\ \Delta_2^{(0)} = 0,2 \\ \Delta_3^{(0)} = -0,3. \end{cases}$$

Розв'язуючи систему, маємо: $\Delta_1^{(0)} = -0,1; \Delta_2^{(0)} = 0,2; \Delta_3^{(0)} = -0,3.$

Звідки

$$x_1^{(1)} = x_1^{(0)} + \Delta_1^{(0)} = 0 - 0,1 = -0,1;$$

$$x_2^{(1)} = x_2^{(0)} + \Delta_2^{(0)} = 0 + 0,2 = 0,2;$$

$$x_3^{(1)} = x_3^{(0)} + \Delta_3^{(0)} = 0 - 0,3 = -0,3;$$

$$\bar{x}^{(1)} = \begin{bmatrix} -0,1 \\ 0,2 \\ -0,3 \end{bmatrix}.$$

Перевіряємо виконання умови (3.3.7):

$$\bar{f}(\bar{x}^{(1)}) = \begin{bmatrix} x_1^{(1)} + (x_1^{(1)})^2 - 2x_2^{(1)}x_3^{(1)} - 0,1 \\ x_2^{(1)} - (x_2^{(1)})^2 + 3x_1^{(1)}x_3^{(1)} + 0,2 \\ x_3^{(1)} + (x_3^{(1)})^2 + 2x_1^{(1)}x_2^{(1)} - 0,3 \end{bmatrix} = \begin{bmatrix} 0,04 \\ 0,27 \\ 0,55 \end{bmatrix}.$$

Звідки одержуємо:

$$\|\bar{f}(\bar{x}^{(1)})\| > \varepsilon.$$

Процес обчислення продовжується у відповідності з наведеним алгоритмом до досягнення точності ε .

3.3.3 Текст програми розрахунку

```
10 REM PROGRAM METON
20 PRINT " МЕТОД НЬЮТОНА "
30 PRINT " УКАЖІТЬ ПОРЯДОК НЕЛІНІЙНОЇ СИСТЕМИ – N, ТОЧНІСТЬ - E "
40 INPUT N, E
```

```

50 DIM A(N, N), B(N), X(N), Y(N)
60 PRINT " ПОЧАТКОВІ НАБЛИЖЕННЯ X(I) "
70 FOR I=1 TO N: PRINT " X("I")= " : INPUT X(I) : NEXT I
80 REM УТОЧНІТЬ ЕЛЕМЕНТИ МАТРИЦІ ЯКОБІ -A(I, J)
90 A(1, 1)=1+2*X(1)
100 A(1, 2)=-2*X(3)
110 A(1, 3)=-2*X(2)
120 A(2, 1)=3*X(3)
130 A(2, 2)=1-2*X(2)
140 A(2, 3)=3*X(1)
150 A(3, 1)=2*X(2)
160 A(3, 2)=2*X(1)
170 A(3, 3)=1+2*X(3)
180 REM УТОЧНІТЬ ВИРАЗИ ФУНКЦІЙ F(I) ДЛЯ B(I)
190 B(1)=-(X(1)+X(1)^2-2*X(2)*X(3)-.1)
200 B(2)=-(X(2)+X(2)^2-3*X(1)*X(3)+.1)
210 B(3)=-(X(3)+X(3)^2-2*X(1)*X(2)-.3)
220 REM РОЗВ'ЯЗОК СИСТЕМИ W(X(K)) Y(K) =-F(X(K))
230 FOR I=1 TO N=1 : FOR J=I+1 TO N
240 A(J, I)=-A(J, I) / A(I, I) : FOR K=I+1 TO N
250 A(J, K)=A(J, K)+A(J, I)*A(I, K) : NEXT J
260 B(J)=B(J)+A(J, I)*B(I) : NEXT J
270 Y(N)=B(N) / A(N, N)
280 NEXT I
290 FOR I=N-1 TO 1 STEP -1 : H=B(I)
300 FOR J=I+1 TO N : H=H-Y(J)*A(I, J) : NEXT J
310 Y(I)=H/A(I, I)
320 NEXT I
330 LPRINT " КОРЕНІ ДЛЯ ДЕЛЬТА - Y " : PRINT
340 FOR I=1 TO N : LPRINT " Y(";I;)"= " Y(I) : NEXT I
350 FOR I=1 TO N : X(I)=X(I)+Y(I) : NEXT I
360 KI=KI+1
370 F= ABS (B(1)+B(2)+B(3))/3
380 IF F>E GOTO 90
390 LPRINT " КОРЕНІ СИСТЕМИ РІВНЯНЬ " : PRINT
400 FOR I=1 TO N : LPRINT "X("I")=" ; X(I) , "F("I")=" ; B(I)
410 NEXT I
420 LPRINT " ІТЕРАЦІЙ - KI= " ; KI , " СЕРЕДНЄ АРИФМ. ЗНАЧЕНЬ Ф. F(I) " ; F
430 STOP

```

3.4 Точкове апроксимування функцій

за методом найменших квадратів

3.4.1 Загальні положення

Нехай функція $y=f(x)$ задається таблично на множині точок x_0, x_1, \dots, x_n , тобто, $y_0=f(x_0), y_1=f(x_1), \dots, y_n=f(x_n)$.

Потрібно апроксимувати функцію $f(x)$ поліномом

$$Q_m(x) = a_0 + a_1x + \dots + a_mx^m \quad (3.4.1)$$

таким чином, щоб відхилення функції $f(x)$ від полінома $Q(x)$ на даній множині точок було мінімальним.

Мірою відхилення будемо вважати квадратичне відхилення:

$$S = \sum_{i=0}^n [Q_m(x_i) - f(x_i)]^2. \quad (3.4.2)$$

Для побудови апроксимованого полінома потрібно вибрати його коефіцієнти a_0, a_1, \dots, a_n таким чином, щоб величина S була мінімальною.

В точці мінімуму функції її перші похідні обов'язково повинні дорівнювати нулю. Диференціюючи функцію (3.4.2) за змінними $\overline{a_i}, i = \overline{1, m}$ та прирівнюючи похідні dS/da_i до нуля, отримуємо систему лінійних рівнянь.

Коефіцієнти полінома a_0, a_1, \dots, a_m є розв'язком системи лінійних рівнянь:

$$\begin{cases} a_0 S_0 + a_1 S_1 + a_2 S_2 + \dots + a_m S_m = t_0 \\ a_0 S_1 + a_1 S_2 + a_2 S_3 + \dots + a_m S_{m+1} = t_1 \\ a_0 S_2 + a_1 S_3 + a_2 S_4 + \dots + a_m S_{m+2} = t_2 \\ \dots \\ a_0 S_m + a_1 S_{m+1} + a_2 S_{m+2} + \dots + a_m S_{2m} = t_m \end{cases}, \quad (3.4.3)$$

де

$$\begin{aligned} S_k &= x_0^k + x_1^k + \dots + x_n^k, k = 0, 1, 2, \dots; \\ t_k &= x_0^k y_0 + x_1^k y_1 + \dots + x_n^k y_n, k = 0, 1, 2, \dots \end{aligned} \quad (3.4.4)$$

Поліном (3.4.1) з коефіцієнтами a_0, a_1, \dots, a_m , що визначаються з системи (3.4.3), матиме мінімальне квадратичне відхилення від даної функції $f(x)$.

3.4.2 Приклад

Побудувати апроксимувальний поліном для функції $f(x)$, заданої таблицю:

$$x_0 = 0,11; x_1 = 0,15; x_2 = 0,2; x_3 = 0,25; x_4 = 0,35; x_5 = 0,4;$$

$$y_0 = 9,05; y_1 = 6,72; y_2 = 4,75; y_3 = 3,27; y_4 = 2,74; y_5 = 2,38.$$

Апроксимувальний поліном матиме вигляд: $Q_2(x) = a_0 + a_1 x + a_2 x^2$.

Для знаходження коефіцієнтів a_1, a_2, a_3 складаємо систему вигляду:

$$\begin{cases} S_0 a_0 + S_1 a_1 + S_2 a_2 + S_3 a_3 + S_4 a_4 + S_5 a_5 = t_0 \\ S_1 a_0 + S_2 a_1 + S_3 a_2 + S_4 a_3 = t_1 \\ S_2 a_0 + S_3 a_1 + S_4 a_2 = t_2 \end{cases}. \quad (3.4.5)$$

Знайдемо коефіцієнти системи S_0, \dots, S_4 та вільні члени t_0, t_1, t_2 .

$$S_0 = n + 1 = 4;$$

$$S_1 = \sum_{i=0}^3 x_i = 7;$$

$$S_2 = \sum_{i=0}^3 x_i^2 = 13,5;$$

$$S_3 = \sum_{i=0}^3 x_i^3 = 2,8125 + 8 + 15,625 = 27;$$

$$S_4 = \sum_{i=0}^3 x_i^4 = 5,0625 + 16 + 39,0625 = 61,125;$$

$$t_0 = \sum_{i=0}^3 y_i = 13,5;$$

$$t_1 = \sum_{i=0}^3 x_i y_i = 3 + 4,8 + 7 + 9,5 = 24,3;$$

$$t_2 = \sum_{i=0}^3 x_i^2 y_i = 3 + 2,25 \cdot 3,2 + 4 \cdot 3,5 + 6,25 \cdot 3,8 = 3 + 7,2 + 14 + 23,75 = 47,95.$$

Таким чином, система має вигляд:

$$\begin{cases} 4a_0 + 7a_1 + 13,5a_2 = 13,5 \\ 7a_0 + 13,5a_1 + 27a_2 = 24,3 \\ 12,5a_0 + 27a_1 + 61,125a_2 = 47,95. \end{cases}$$

Розв'язуючи цю систему, маємо:

$$a_0 = 2,55;$$

$$a_1 = 0,306;$$

$$a_2 = 0,088.$$

Апроксимувальний поліном має вигляд $Q(x) = 2,55 + 0,306x + 0,088x^2$.

Обчислюємо значення полінома $Q(x)$ у точках x_i :

$$Q(x_0) = 2,94;$$

$$Q(x_1) = 3,157;$$

$$Q(x_2) = 3,514;$$

$$Q(x_3) = 3,865.$$

Підрахуємо мінімальне квадратичне відхилення за формулою:

$$S_{\min} = \sum_{i=0}^3 (Q_2(x_i) - f(x_i))^2 = 0,0098.$$

Побудуємо графіки функції $f(x)$ та полінома $Q(x)$ (рис.3.2).

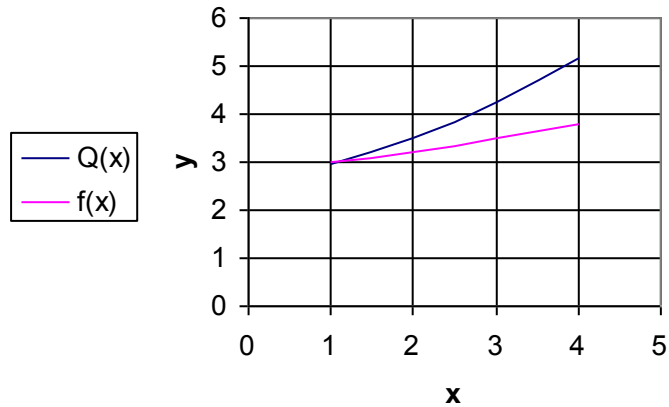


Рисунок 3.2

3.4.3 Текст програми розрахунку

```

5 REM АПРОКСИМАЦІЯ ПОЛІНОМОМ 2-ГО ПОРЯДКУ
10 DIM S(6) , X(6) , Y(6) , T(6) , Q(6)
20 FOR I=0 TO 3
30 REM ВВЕДЕННЯ ПОЧАТКОВОГО НАБЛИЖЕННЯ
40 INPUT X(I) , Y(I)
50 NEXT I
60 FOR J=0 TO 4
65 FOR I=0 TO 3
70 S(J)=S(J)+X(I)^J
80 NEXT I
90 NEXT J
100 FOR J=0 TO 2
110 FOR I=0 TO 3
120 T(J)=T(J)+X(I)^J*Y(I)
130 NEXT I
131 LPRINT " T("J")= "; T(J)
140 NEXT J
150 A(1,1)=S(0)
160 A(2,1)=S(1)
170 A(1,2)=A(2,1)
180 A(3,1)=S(2)
190 A(2,2)=S(2)
200 A(1,3)=S(2)
210 A(3,2)=S(3)
220 A(2,3)=S(3)
230 A(3,3)=S(4)
240 FOR I=1 TO 3
250 FOR J=1 TO 3
260 LPRINT " A("I","J")= "; A(I,J)
270 NEXT J : NEXT I

```

```

271 PRINT "A0,A1,A2"
275 INPUT A0,A1,A2
276 FOR I=0 TO 3
277 Q(I)=Q(I)+A0+A1*X(I)+A2*X(I)^2
278 D=(Y(I)-Q(I))^2
279 LPRINT " Q("I")= "; Q(I) , " D= "; D
280 NEXT I
290 STOP

```

3.5 Одновимірна мінімізація функцій

3.5.1 Загальні положення

Розв'язується задача

$$y(x) \rightarrow \min. \quad (3.5.1)$$

Нехай функція $y(x)$ унімодальна на відрізку $[a, b]$, точка мінімуму x^* належить $[a, b]$. Треба визначити точку мінімуму функції $y(x)$ x^* з заданою точністю ε .

Метод дихотомії. Нехай на k -му кроці точка мінімуму x^* належить відрізку $[a_k, b_k]$. Відрізок $[a_k, b_k]$ ділиться точкою $\bar{x}^{(k)}$ навпіл, потім обчислюються значення функції $y(x)$ у точках

$$x_1^{(k)} = \bar{x}^{(k)} - \frac{\varepsilon}{2}, \quad x_2^{(k)} = \bar{x}^{(k)} + \frac{\varepsilon}{2}.$$

Порівнюючи значення $y(x_1^{(k)})$ та $y(x_2^{(k)})$, робимо висновок, до якого з відрізків $[a_k, \bar{x}^{(k)}]$ чи $[\bar{x}^{(k)}, b_k]$ належить точка мінімуму, тобто звужуємо відрізок $[a_k, b_k]$. Процес пошуку точки мінімуму продовжується до досягнення заданої точності.

Алгоритм методу дихотомії містить такі етапи:

1. Ввести a, b .
2. Взяти $k=0, a_k=a, b_k=b$.
3. Визначити:

$$\bar{x}^{(k)} = \frac{a_k + b_k}{2}; \quad x_1^{(k)} = \bar{x}^{(k)} - \frac{\varepsilon}{2}, \quad x_2^{(k)} = \bar{x}^{(k)} + \frac{\varepsilon}{2}.$$

4. Обчислити $y(x_1^{(k)})$ і $y(x_2^{(k)})$.

5. Порівняти $y(x_1^{(k)})$ та $y(x_2^{(k)})$. Якщо $y(x_1^{(k)}) \geq y(x_2^{(k)})$, то $a_{k+1} = \bar{x}^{(k)}$, $b_{k+1} = b_k$, якщо $y(x_1^{(k)}) < y(x_2^{(k)})$, то $a_{k+1} = a_k$, $b_{k+1} = x_2^{(k)}$.

6. Перевірити умови завершення пошуку: $|b_{k+1} - a_{k+1}| < \varepsilon$. Якщо умова виконується, то перейти до наступного етапу, в іншому випадку – до етапу 3, припустивши $k=k+1$.

7. Взяти за точку мінімуму

$$x^* = \frac{a_{k+1} + b_{k+1}}{2}.$$

Метод золотого перерізу. На кожному кроці методу точки $x_1^{(k)}, x_2^{(k)}$

знаходяться на відрізку $[a_k, b_k]$ симетрично відносно його кінців і обчислюються за формулами, що витікають з пропорції золотого перерізу:

$$\begin{aligned}x_1^{(k)} &= a_k + \frac{1}{\tau^2}(b_k - a_k); \\x_2^{(k)} &= a_k + \frac{1}{\tau}(b_k - a_k),\end{aligned}\tag{3.5.2}$$

де

$$\frac{1}{\tau^2} \approx 3,382; \frac{1}{\tau} \approx 0,618.\tag{3.5.3}$$

Порівняння значень функції $y(x_1^{(k)})$ та $y(x_2^{(k)})$ дає змогу визначити відрізок $[a_k, b_k]$, який містить точку мінімуму x^* .

Алгоритм методу золотого перерізу складається з таких етапів:

1. Ввести a, b .
2. Взяти $k=0, a_k=a, b_k=b$.
3. Визначити $x_1^{(k)}, x_2^{(k)}$ за формулами (3.5.2), (3.5.3).
4. Обчислити $y(x_1^{(k)}), y(x_2^{(k)})$.
5. Порівняти $y(x_1^{(k)})$ та $y(x_2^{(k)})$. Якщо $y(x_1^{(k)}) \geq y(x_2^{(k)})$, то взяти $a_{k+1}=x_1^{(k)}$; $b_{k+1}=b_k$; $x_1^{(k+1)}=x_2^{(k)}$; $y(x_1^{(k+1)}) = y(x_2^{(k)})$; обчислити $x_2^{(k+1)}, y(x_2^{(k+1)})$. Якщо $y(x_1^{(k)}) < y(x_2^{(k)})$, то взяти $a_{k+1}=a_k$; $b_{k+1}=x_2^{(k)}$; $x_2^{(k+1)}=x_1^{(k)}$; $y(x_2^{(k+1)})=y(x_1^{(k)})$, обчислити $y(x_2^{(k+1)})$; $y(x_2^{(k+1)})$.
6. Перевірити виконання умови $|b_{k+1}-a_{k+1}| < \varepsilon$. Якщо умова виконується, то перейти до наступного етапу, в іншому випадку – припустити $k=k+1$ і перейти до етапу 5.
7. За точку мінімуму взяти

$$x^* = \frac{a_{k+1} + b_{k+1}}{2}.$$

8. Вивести на друк або монітор $x^*, y(x^*)$.

3.5.2 Приклад

Знайти мінімум функції

$$y = -\frac{803}{3}x^3 + 275x^2 - 74x - \frac{1}{3}$$

на $[0;0,5]$ з точністю до $\varepsilon=0,05$.

За допомогою методу дихотомії маємо:

1. $a=0, b=0,5$;
2. $k=0; a_0=a=0; b_0=b=0,5$;
3. Обчислимо:

$$\tilde{x}^{(0)} = \frac{a_0 + b_0}{2} = 0,25;$$

$$x_1^{(0)} = \tilde{x}^{(0)} - \frac{\varepsilon}{2} = 0,25 - 0,025 = 0,225;$$

$$x_2^{(0)} = \tilde{x}^{(0)} + \frac{\varepsilon}{2} = 0,25 + 0,025 = 0,275.$$

4. $y(x_1^{(0)}) = -6,11; y(x_2^{(0)}) = -5,453.$

5. $y(x_1^{(0)}) < y(x_2^{(0)})$ тому $a_1 = a_0 = 0; b_1 = \bar{x}^{(0)} = 0,25.$

6. $|b_1 - a_1| > \varepsilon \rightarrow k=1$ переходимо до пункту 3 і продовжуємо процес пошуку до досягнення потрібної точності.

За допомогою методу золотого перерізу маємо:

1. $a=0, b=0,5;$

2. $k=0; a_0=0; b_0=0,5;$

3. $x_1^{(0)} = a_0 + 0,382(b_0 - a_0) = 0,191; x_2^{(0)} = a_0 + 0,618(b_0 - a_0) = 0,309;$

4. $y(x_1^{(0)}) = -6,3; y(x_2^{(0)}) = -4,839.$

$$a_1 = a_0 = 0; b_1 = x_2^{(0)} = 0,309; x_2^{(1)} = x_1^{(0)} = 0,191;$$

$$y(x_2^{(1)}) = y(x_1^{(0)}) = -6,3; x_1^{(1)} = a_1 + 0,382(b_1 - a_1) = 0,118;$$

$$y(x_1^{(1)}) = -5,677;$$

5. $y(x_1^{(1)}) < y(x_2^{(1)}) \rightarrow$

6. $|b_1 - a_1| > \varepsilon \rightarrow$ переходимо до пункту 5, припустивши $k=1.$

Друга ітерація

5. $y(x_1^{(1)}) > y(x_2^{(1)}) \rightarrow$

$$a_2 = 0,118; b_2 = 0,309; x_1^{(2)} = 0,191; y(x_1^{(2)}) = y(x_1^{(0)}) = -6,3;$$

$$x_2^{(2)} = a_2 + 0,618(b_2 - a_2) = 0,236; y(x_2^{(2)}) = -5,999.$$

6. $|b_2 - a_2| > \varepsilon \rightarrow$ беремо $k=2$ і переходимо до пункту 5.

Процес пошуку продовжується до досягнення заданої точності.

3.5.3 Текст програми розрахунку

10 REM МЕТОД ДИХОТОМІЇ

20 INPUT A , B , E

30 K=0

40 AK=A

50 BK=B

60 XG=(AK+BK) / 2

65 DEF FNF (X1)=- (803*X1^3)/3+275* X1^2-74*X1-1/3

66 DEF FNY (X2)=- (803*X2^3)/3+275* X2^2-74*X2-1/3

70 X1=XG+E / 2

80 X2=XG-E / 2

103 F1=FNF (X1)

104 F2=FNY (X2)

110 IF F1>=F2 GOTO 180

```

120 AK=AK
130 BK=XG
140 IF ABS (BK-AK) < G GOTO 160
150 GOTO 60
160 XX=(AK+DR) / 2
165 X1=XX
170 PRINT " ТОЧКА МІНІМ.-XX= " ; XX , "F("XX")=" ; F1
175 STOP
176 GOTO 210
180 AK=XG
190 BK=BK
200 GOTO 140
210 REM МЕТОД ЗОЛОТОГО ПЕРЕРІЗУ
220 K1=.382
230 K2=.618
240 AK=A
250 BK=B
260 X1=AK+K1*(BK-AK)
270 X2=AK+K2*(BK-AK)
280 F1=FNF (X1)
290 F2=FNY (X2)
300 IF F1>=F2 GOTO 400
310 AK=AK
320 BK=X2
330 X2=X1
340 F2=F1
350 X1=AK=K1*(BK-AK)
360 F1=FNF (X1)
370 IF ABS (BK-AK) / 2>E GOTO 300
380 XX=(AK+BK) / 2
390 GOTO 165
400 AK=X1
410 BK=BK
420 X1=X2
430 F1=F2
440 X2=AK+K2*(BK-AK)
450 F2=FNY (X2)
460GOTO 370

```

3.6 Багатовимірна безумовна мінімізація функцій.

Гرادієнтні методи

3.6.1 Загальні положення

Треба розв'язати задачу

$$y(\bar{x}) \rightarrow \min_{\bar{x}},$$

тобто знайти точку мінімуму функції $y(x)$ з точністю до ε .

За допомогою градієнтних методів послідовність наближень до точки

мінімуму визначається у відповідності з виразом:

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \lambda^{(k)} \left(\frac{\partial y}{\partial \bar{x}} \right)^k, k = 0, 1, \dots,$$

де $\lambda^{(k)}$ – параметр, що регулює величину кроку, $0 < \lambda^{(k)} < 1$.

Гradientні методи відрізняються засобами вибору параметра $\lambda^{(k)}$. Процес пошуку точки мінімуму припиняється, коли необхідні умови існування точки мінімуму виконуються із заданою точністю ε .

3.6.1.1 Gradientний метод з дрібненням кроку

У цьому методі параметр $\lambda^{(k)}$ вибирається таким чином, щоб виконувалася умова $y(\bar{x}^{(k+1)}) < y(\bar{x}^{(k)})$.

Процес пошуку точки мінімуму продовжується до виконання умови:

$$\left| \left(\frac{\partial y}{\partial \bar{x}} \right)^{k+1} \right| \leq \varepsilon.$$

Алгоритм методу містить такі етапи:

1. Ввести $\bar{x}^{(k)}$, $k=0$.
2. Обчислити $y(\bar{x}^{(k)})$.
3. Взяти $\lambda^{(k)}=1$.
4. Обчислити:

$$\bar{x}^{(k+1)} = \bar{x}^{(k)} - \lambda^{(k)} \left(\frac{\partial y}{\partial \bar{x}} \right)^k.$$

5. Обчислити $y(\bar{x}^{(k+1)})$.
6. Порівняти $y(\bar{x}^{(k)})$ та $y(\bar{x}^{(k+1)})$. Якщо $y(\bar{x}^{(k)}) < y(\bar{x}^{(k+1)})$, перейти до п.7, якщо $y(\bar{x}^{(k)}) > y(\bar{x}^{(k+1)})$, то зменшити $\lambda^{(k)}$ вдвічі і перейти до п.4
7. Перевірити умову:

$$\left| \left(\frac{\partial y}{\partial \bar{x}} \right)^{k+1} \right| \leq \varepsilon.$$

Якщо умова не виконується, то, взявши $k=k+1$, перейти до п.3, у іншому випадку перейти до наступного пункту.

8. За точку мінімуму взяти $\bar{x}^* = \bar{x}^{(k+1)}$.
9. Вивести результати на друк або монітор: \bar{x}^* , $y(\bar{x}^*)$.

3.6.1.2 Метод координатного спуску

Суть цього методу полягає в тому, що на кожному r -ому кроці k -ої ітерації змінюється лише одна змінна x_r , для якої не виконується необхідна умова існування точки мінімуму:

$$\bar{x}_r^{(k+1)} = \bar{x}_r^{(k)} - \lambda_r^{(k)} \left(\frac{\partial y}{\partial x_r} \right)^k, r = 0, 1, \dots, n; k = 0, 1, \dots$$

Параметр $\lambda_r^{(k)}$ обчислюється за формулою:

$$\lambda_r^{(k)} = \frac{1}{\left(\frac{\partial^2 y}{\partial x_r^2} \right)^{(k+1)}}.$$

Алгоритм методу містить такі етапи:

1. Ввести $\bar{x}^{(k)}$, $k=0$.
2. Взяти $r=1$.
3. Порівняти r та n , якщо $r < n$, перейти до п.4, якщо $r > n$ – до п.8.
4. Обчислити

$$\left(\frac{\partial y}{\partial x_r} \right)^{(k)}.$$

5. Перевірити виконання необхідної умови існування точки мінімуму:

$$\text{якщо } \left| \left(\frac{\partial y}{\partial x_r} \right)^{(k)} \right| \leq \varepsilon, \text{ то беремо } r = r + 1 \text{ і переходимо до п.3,}$$

$$\text{якщо } \left| \left(\frac{\partial y}{\partial x_r} \right)^{(k)} \right| > \varepsilon, \text{ то переходимо до п.6.}$$

6. Обчислити:

$$x_r^{(K+1)} = x_r^{(k)} - \frac{\left(\frac{\partial y}{\partial x_r} \right)^{(k)}}{\left(\frac{\partial^2 y}{\partial x_r^2} \right)^{(k)}}.$$

7. Взяти $r=r+1$ і перейти до п.3.

8. Перевірити виконання умови:

$$\left| \left(\frac{\partial y}{\partial \bar{x}} \right)^{(k+1)} \right| \leq \varepsilon.$$

Якщо умова не виконується, то припустити $k=k+1$ та перейти до п.2. У іншому випадку перейти до п.9.

9. Взяти за точку мінімуму $\bar{x}^\bullet = \bar{x}^{(k+1)}$.

10. Вивести на друк або монітор $\bar{x}^\bullet, y(\bar{x}^\bullet)$.

3.6.1.3 Приклад

Знайти мінімум функції

$$y = x_1 x_2 - \frac{1}{x_1} - \frac{1}{x_2}$$

з точністю до $\varepsilon = 0,01$. Початкове наближення:

$$\bar{x}^{(0)} = \begin{bmatrix} -0,5 \\ -0,5 \end{bmatrix}.$$

Гرادієнтний метод з дрібненням кроку

Обчислюємо $y(\bar{x}^{(0)}) = 4,25$.

Знаходимо похідні:

$$\frac{\partial y}{\partial x_1} = x_2 + \frac{1}{x_1^2}; \frac{\partial y}{\partial x_2} = x_1 + \frac{1}{x_2^2}.$$

Обчислюємо:

$$\left(\frac{\partial y}{\partial x_1}\right)^{(0)} = x_2^{(0)} + \frac{1}{(x_1^{(0)})^2} = 3,5; \left(\frac{\partial y}{\partial x_2}\right)^{(0)} = x_1^{(0)} + \frac{1}{(x_2^{(0)})^2} = 3,5.$$

Таким чином,

$$\left(\frac{\partial y}{\partial \bar{x}}\right)^{(0)} = \begin{bmatrix} 3,5 \\ 3,5 \end{bmatrix}; \left(\frac{\partial y}{\partial \bar{x}}\right)^{(0)} > \varepsilon.$$

Нульова ітерація: $k=0$.

Беремо $\lambda^{(0)}=1$.

Знаходимо:

$$\bar{x}^{(1)} = \bar{x}^{(0)} - \lambda^{(0)} \left(\frac{\partial y}{\partial \bar{x}}\right)^{(0)},$$

$$\bar{x}^{(1)} = \begin{pmatrix} x_1^{(1)} \\ x_2^{(1)} \end{pmatrix} = \begin{pmatrix} -0,5 \\ -0,5 \end{pmatrix} - \begin{pmatrix} -3,5 \\ -3,5 \end{pmatrix} = \begin{pmatrix} -4 \\ -4 \end{pmatrix}.$$

Обчислюємо $y(\bar{x}^{(1)}) = 16,5$.

Адже $y(\bar{x}^{(1)}) > y(\bar{x}^{(0)})$, дрібномо параметр $\lambda^{(0)}$.

$$\lambda^{(0)} = \frac{1}{2}.$$

Знаходимо:

$$\bar{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} -0,5 \\ -0,5 \end{bmatrix} - 0,5 \begin{bmatrix} 3,5 \\ 3,5 \end{bmatrix} = \begin{bmatrix} -2,25 \\ -2,25 \end{bmatrix},$$

$$y(\bar{x}^{(1)}) = 5,95.$$

Умова $y(\bar{x}^{(1)}) < y(\bar{x}^{(0)})$ не виконується, знову дрібномо $\lambda^{(0)}$:

$$\lambda^{(0)} = \frac{1}{4}.$$

$$\bar{x}^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} -0,5 \\ -0,5 \end{bmatrix} - 0,25 \begin{bmatrix} 3,5 \\ 3,5 \end{bmatrix} = \begin{bmatrix} -1,375 \\ -1,375 \end{bmatrix};$$

$$y(\bar{x}^{(1)}) = 3,345.$$

Тепер умова $y(\bar{x}^{(1)}) < y(\bar{x}^{(1)})$ виконана, тобто

$$\bar{x}^{(1)} = \begin{bmatrix} -1,375 \\ -1,375 \end{bmatrix}.$$

Обчислюємо:

$$\left(\frac{\partial y}{\partial \bar{x}} \right)^{(1)} = \begin{pmatrix} -0,846 \\ -0,846 \end{pmatrix},$$

$$\left(\frac{\partial y}{\partial \bar{x}} \right)^{(1)} > \varepsilon,$$

тобто, задану точність не досягнуто, переходимо до наступної ітерації.

Перша ітерація: $k=1$.

Беремо $\lambda^{(0)}=1$.

Знаходимо:

$$\bar{x}^{(2)} = \begin{bmatrix} -1,375 \\ -1,375 \end{bmatrix} - \begin{bmatrix} -0,846 \\ -0,846 \end{bmatrix} = \begin{bmatrix} -0,529 \\ -0,529 \end{bmatrix};$$

$$y(\bar{x}^{(2)}) = 4,06.$$

Оскільки $y(\bar{x}^{(1)}) < y(\bar{x}^{(2)})$, беремо $\lambda^{(0)}=0,5$.

Обчислюємо знову:

$$\bar{x}^{(2)} = \begin{bmatrix} -1,375 \\ -1,375 \end{bmatrix} - 0,5 \begin{bmatrix} -0,846 \\ -0,846 \end{bmatrix} = \begin{bmatrix} -0,952 \\ -0,952 \end{bmatrix};$$

$$y(\bar{x}^{(2)}) = 3,007.$$

Умова $y(\bar{x}^{(1)}) < y(\bar{x}^{(2)})$ виконана, тобто

$$\bar{x}^{(2)} = \begin{bmatrix} -0,952 \\ -0,952 \end{bmatrix}.$$

Знаходимо:

$$\left(\frac{\partial y}{\partial \bar{x}} \right)^{(2)} = \begin{pmatrix} 0,152 \\ 0,152 \end{pmatrix},$$

$$\left(\frac{\partial y}{\partial \bar{x}} \right)^{(2)} > \varepsilon,$$

тобто, задану точність розв'язку не досягнуто, треба взяти $k=2$ і перейти до наступної ітерації.

Метод координатного спуску

Обчислюємо $y(\bar{x}^{(0)})=4,25$.

Знаходимо вирази для перших та других похідних:

$$\frac{\partial y}{\partial x_1} = x_2 + \frac{1}{x_1^2}; \frac{\partial y}{\partial x_2} = x_1 + \frac{1}{x_2^2};$$

$$\frac{\partial^2 y}{\partial x_1^2} = -\frac{2}{x_1^3}; \frac{\partial^2 y}{\partial x_1 \partial x_2} = -\frac{1}{x_2^2}.$$

Нульова ітерація: $k=0$.

I-й крок: $x=1$.

Обчислюємо:

$$\left(\frac{\partial y}{\partial x_1}\right)^{(0)} = x_2^{(0)} + \frac{1}{x_1^{(0)2}} = 3,5.$$

Задану точність за змінною x_1 не досягнуто, адже

$$\left(\frac{\partial y}{\partial x_1}\right)^{(0)} > \varepsilon.$$

Знаходимо значення $x_1^{(1)}$:

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(0)} = -\frac{2}{x_1^{(0)3}} = 16;$$

$$x_1^{(1)} = x_1^{(0)} - \frac{\left(\frac{\partial y}{\partial x_1}\right)^{(0)}}{\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(0)}} = 1 - \frac{3,5}{16} = 0,78125.$$

Перевіряємо виконання умови:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(0)}\right| > \varepsilon.$$

Враховуючи, що змінну x_1 вже обчислено (дорівнює $0,78125$),

$$\left(\frac{\partial y}{\partial x_1}\right)^{(0)} = x_1 + \frac{1}{x_1^{(0)2}} = 3,28$$

задану точність за змінною x_2 не досягнуто.

II-й крок. Знаходимо $x_2^{(1)}$:

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(0)} = -\frac{2}{x_1^{(0)3}} = 16;$$

$$x_2^{(1)} = x_2^{(0)} - \frac{\left(\frac{\partial y}{\partial x_2}\right)^{(0)}}{\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(0)}} = -0,5 - \frac{3,28}{16} = -0,705.$$

Перевіримо, що після ітерації ми наблизимося до точки мінімуму:

$$y(\bar{x}^{(1)}) = 3,316 < y(\bar{x}^{(0)}) = 4,25.$$

Перевіримо умову:

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(1)}\right| > \varepsilon.$$

Маємо:

$$\left(\frac{\partial y}{\partial x_1}\right)^{(1)} = 1,229; \left(\frac{\partial y}{\partial x_2}\right)^{(1)} = 1,293.$$

Необхідна умова існування точки мінімуму не виконується, тому переходимо до першої ітерації.

Перша ітерація: $k=k+1=1$.

I-й крок

$$\left(\frac{\partial y}{\partial x}\right)^{(1)} = 1,229; \left|\left(\frac{\partial y}{\partial x_1}\right)^{(1)}\right| > \varepsilon.$$

Знаходимо $x_1^{(2)}$:

$$\left(\frac{\partial^2 y}{\partial x_1^2}\right)^{(1)} = 5,376; x_1^{(2)} = -0,719 - \frac{1,229}{5,376} = 0,948.$$

II-й крок

$$\left(\frac{\partial y}{\partial x_2}\right)^{(1)} = 1,064; \left|\left(\frac{\partial y}{\partial x_2}\right)^{(1)}\right| > \varepsilon.$$

Знаходимо $x_2^{(2)}$:

$$\left(\frac{\partial^2 y}{\partial x_2^2}\right)^{(1)} = 5,714; x_2^{(2)} = -0,705 - \frac{1,064}{5,414} = -0,891.$$

Оскільки $y(\bar{x}^{(2)}) = 3,022 < y(\bar{x}^{(1)}) = 3,316$, то в результаті ітерації ми наближаємося до точки мінімуму (значення функції змінюється).

Перевіряємо виконання необхідної умови існування точки мінімуму:

$$\left(\frac{\partial y}{\partial x_1}\right)^{(2)} = 0,221; \left(\frac{\partial y}{\partial x_2}\right)^{(2)} = 0,311.$$

Умова

$$\left|\left(\frac{\partial y}{\partial x_1}\right)^{(2)}\right| > \varepsilon,$$

не виконується, переходимо до наступної ітерації.

Процес пошуку мінімуму функції $y(\bar{x})$ продовжується до досягнення заданої точності ε .

3.6.1.4 Текст програми розрахунку

```

10 REM ГРАДІЄНТНИЙ МЕТОД З ДРІБНЕННЯМ КРОКУ
20 DEF FNXX (X1, X2)=X1*X2-1/X1-1/X2
30 DEF FNXA (X1, X2)=X2-1/X1^2
40 DEF FNXB (X1, X2)=X1-1/X2^2
45 E=.001 : L=1
50 X1=-.8
60 X2=-.8
65 FS=FNXX (X1, X2)
70 F1=FNXA (X1, X2)
75 IF R=1 GOTO 230
80 F2=FNXB (X1, X2)
90 IF N=2 GOTO 310
95 GR= SQR(F1^2+F2^2)
100 IF GR<E GOTO 180
101 K=K+1
110 X1=X1-L*F2
120 X2=X2-L*F2
130 FH=FNXX (X1, X2)
140 IF FH>FS GOTO 160
150 GOTO 65
160 L=L/2
170 GOTO 65
180 PRINT " X1=" ; X1, " X2=" ; X2, " FH="; FH , " ЧИСЛО ІТЕР.= " ; K
185 STOP
190 REM МЕТОД КООРДИНАТНОГО СПУСКУ
200 DIM F(2)
210 DEF FNDD (X1)=-2/X1^3
220 DEF FNDE (X2)=-2/X2^3
221 N=2 : R=1 : K=0 : GOTO 50
230 IF ABS (F1)<E GOTO 290
240 F(1)=FNDD(X1)
250 L=1/F(1)
260 X1=X1-F1*L
270 FF=FNXX (X1, X2)
271 K=K+1

```

```

280 GOTO 65
290 R=N
300 GOTO 80
310 IF ABS (F2)<E GOTO 350
315 F(2)=FNDE (X2)
320 L=1/F(2)
330 X2=X2-F2*L
331 K=K+1
340 GOTO 80
350GOTO 180

```

3.7 Чисельне інтегрування та диференціювання функцій однієї змінної

Задача чисельного інтегрування полягає в знаходженні наближеного інтеграла

$$I(f) = \int_a^b f(x) dx,$$

де $f(x)$ – задана функція.

На відрізку $[a, b]$ вводиться сітка: $x_0 = a < x_1 < \dots < x_i < x_{i+1} < \dots < x_N = b$, і за наближене значення інтеграла розглядається число

$$I_N(f) = \sum_{i=0}^N c_i f(x_i), \quad (3.7.1)$$

де $f(x_i)$ б – значення функції $f(x)$ у вузлах $x = x_i$, c_i б – деякі числа (вагові множники), які залежать тільки від вузлів, але не залежать від вибору $f(x)$. Формула (3.7.1) називається *квадратурною формулою*.

Найчастіше використовуються в чисельному інтегруванні квадратурні формули прямокутників, трапецій і Сімпсона.

Розглянемо найпростіші квадратурні формули.

Формула прямокутників:

$$I(f) \approx \sum_{i=0}^{N-1} x_i f(x_{i+1}).$$

Формула трапецій:

$$I(f) \approx \frac{1}{2} \sum_{i=0}^{N-1} (f_{i+1} + f_i)(x_{i+1} - x_i).$$

Формула Сімпсона є наслідком інтерполяції $f(x)$ на інтервалі (x_k, x_{k+2}) . Нехай числова область складається з двох інтервалів довжиною h : $x_{k+1} = x_k + h$, $x_{k+2} = x_k + 2h$. Тоді, замінюючи криву АВС інтерполяційною параболою, яка проведена через точки f_k , f_{k+1} , f_{k+2} , отримують формулу Сімпсона (рис.3.3):

$$I(f) \approx \frac{h}{3} [f(x_k) + 4f(x_{k+1}) + f(x_{k+2})].$$

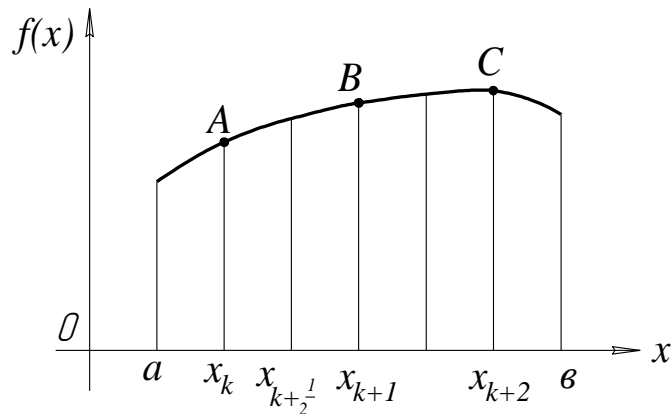


Рисунок 3.3

На відрізку $[a, b]$ при $h = \frac{b-a}{2n}$; $x_k = a + kh$; $y_k = f(x_k)$ ($k = \overline{0, 2n}$) формулу Сімпсона можна подати в такому вигляді:

$$I(f) \approx \frac{h}{3} \left[y_0 + y_{2n} + \sum_{k=1}^{2n-1} (3 + c_k) y_k \right],$$

де $c_k = \begin{cases} 1 & \text{при непарному } k; \\ -1 & \text{при парному } k. \end{cases}$

При грубому оцінюванні інтеграла від остаточного члену квадратичної інтерполяції похибка формули складає $O(h^4)$.

Приклад

За дослідними даними про залежність ентропії S від температури процесу T

$T(S), \text{ К}$	278	303	323	340	350	354	347	330	325
$S, \text{ кДж/(кг}\cdot\text{К)}$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9

визначити середньоінтегральну температуру газу $T_{c.i.}$, використовуючи для цього термодинамічне співвідношення:

$$T_{c.i.} = \left(\int_{t_0}^{S_{2n}} T(S) dS \right) / (S_{2n} - S_0).$$

Інтеграл в даній формулі обчислюється за методом Сімпсона:

$$I \approx \frac{h}{3} [f(S_0) + f(S_8) + 2f(S_2) + 4f(S_3) + 2f(S_4) + f(S_0) + 4f(S_5) + \\ + 2f(S_6) + 4f(S_7)] = \frac{0.1}{3} [287 + 325 + 4 \cdot 303 + 2 \cdot 323 + 4 \cdot 340 + 2 \cdot 350 + \\ + 4 \cdot 354 + 2 \cdot 347 + 4 \cdot 330] = 265.$$

Тоді:

$$T_{c.i.} = \frac{I}{(S_{2n} - S_0)} = \frac{265}{0.8} = 331K.$$

Чисельне диференціювання або визначення похідних є теж не менш важливою задачею в прикладному аналізі.

Нехай відомі значення функції $f(x)$ в точках x_1, x_2, \dots, x_n , близьких до точки x (точка x може збігатися з однією з точок x_1, x_2, \dots, x_n). Необхідно знати наближене значення похідної $f'(x)$ та $f''(x)$.

Один із способів розв'язання цієї задачі полягає в заміні функції $f(x)$ її інтерполяційним поліномом:

$$\frac{df}{dx} \approx \frac{d}{dx} P_{n-1}(x; f).$$

В іншому способі – невизначених коефіцієнтів – використовується формула розкладання в ряд Тейлора функції $f(x)$, неперервної в околі точки x_k :

$$f(x_k) = f(x) + (x_k - x)f'(x) + \frac{1}{2!}(x_k - x)^2 f''(x) + \dots$$

Звідки

$$f'(x) \approx \frac{f(x_k) - f(x)}{x_k - x}.$$

Оскільки $x_k = x \pm h$, то похідна

$$f'(x) \approx \frac{f(x \pm h) - f(x)}{\pm h} \quad (3.7.2)$$

називається, відповідно, правою і лівою похідною. Похибка обчислення похідної буде порядку $O(h)$.

Формула (3.7.2) може бути покращена за рахунок триточкової апроксимації, яка має другий порядок точності. На підставі формули Тейлора:

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + O(h^3);$$

$$f(x+2h) = f(x) + 2hf'(x) + 2h^2 f''(x) + O(h^3).$$

Помножимо перший рядок на 4 і віднімемо його від другого, отримаємо:

$$f'(x) = (-f(x+2h) + 4f(x+h) - 3f(x))/2h + O(h^2),$$

де $O(h^2)$ - похибка обчислення похідної.

Отримана формула може бути використана для визначення похідної в крайній лівій точці, а для останньої (кінцевої) точки використовується формула:

$$f'(x_n) = (3f(x_n) - 4f(x_n - h) + f(x_n - 2h))/2h + O(h^2).$$

Найбільш практичною формулою для внутрішніх точок є центральна різницева похідна:

$$f'(x_k) = (f(x_k + h) - f(x_k - h))/2h + O(h^2).$$

Визначимо симетричну апроксимацію другої похідної функції $f(x)$.

Нехай $x_1 = x - h$, $x_2 = x$, $x_3 = x + h$. На підставі формули Тейлора маємо:

$$f(x + h) = f(x) + hf'(x) + \frac{h^2}{2} f''(x) + \frac{h^3}{6} f'''(x) + O(h^4);$$

$$f(x - h) = f(x) - hf'(x) + \frac{h^2}{2} f''(x) - \frac{h^3}{6} f'''(x) + O(h^4).$$

Складемо ці рівняння і визначимо $f''(x)$:

$$f''(x) = \frac{f(x + h) - 2f(x) + f(x - h)}{h^2} + O(h^2).$$

Отримані формули мають здатність примножувати ефекти завад як при надмірно малому кроці сітки h , так і при великих його значеннях. Отже, в кожному випадку необхідний підбір оптимального кроку для максимально можливого уникнення завад.

Так, при розрахунку похідних на ЕОМ величину h (практичні результати) вибирають в межах 0,001 або 0,005. Покажемо це на наступному прикладі.

Визначимо першу та другу похідні функції $f(x) = e^x + 2x^3$ в точці $x = 0$ при $h = 0,001$; 0,005 та порівняємо з аналітичними розрахунками:

$$f'(x) = e^x + 6x^2; f''(x) = e^x + 12x; f'(0) = 1; f''(0) = 1.$$

При $h = 0,001$:

$$f'(0) = \frac{-(e^{0,002} + 2 \cdot 0,002^3) + 4(e^{0,001} + 2 \cdot 0,001^3) - (e^0 + 2 \cdot 0,002^3)}{0,002} = 1,000047$$

$$f''(0) = \frac{(e^{0,001} + 2 \cdot 1 + e^{-0,001})}{0,001^2} = 1,013279$$

При $h = 0,005$:

$$f'(0) = \frac{-(e^{0,01} + 2 \cdot 0,01^3) + 4(e^{0,005} + 2 \cdot 0,005^3) - 3}{0,01} = 0,999898$$

$$f''(0) = \frac{(e^{0,005} + 2 \cdot 1 + e^{-0,005})}{0,005^2} = 1,001358$$

Отже, правомірність наших передбачень очевидна.

Приклад

Використовуючи дані про залежність границі міцності R від водоцементного відношення B/C

B/C	0,3	0,35	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75
R , МПа	32	36	41	41	38	36	34	32	30	28

оцінити за значенням $f'(B/C)$ чутливість R до зміни B/C при значеннях останнього: 0,3; 0,45 і 0,6.

Підстановкою відповідних значень в триточкову формулу для $f'(x)$ отримаємо:

$$f'(0,3) \approx [-f(0,3 + 2 \cdot 0,05) + 4f(0,3 + 0,05) - 3f(0,3)] / (2 \cdot 0,05) =$$
$$= (-41 + 4 \cdot 36 - 3 \cdot 32) / 0,1 = 70;$$

$$f'(0,45) \approx \frac{1}{0,05} \left(\frac{38 - 41}{2} \right) = -30;$$

$$f'(0,6) \approx \frac{1}{0,05} \left(\frac{32 - 36}{2} \right) = -40.$$

Отже, в великій частині розглянутого діапазону B/C з його збільшенням міцність бетону зменшується, причому при великих B/C такого спаду менше (рис.3.4).

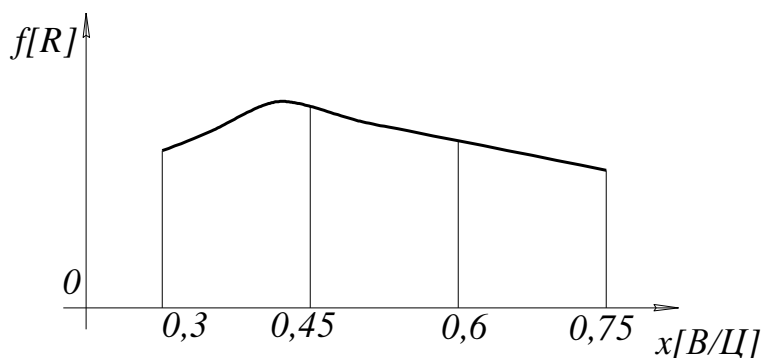


Рисунок 3.4

IV ЕЛЕМЕНТИ ВАРІАЦІЙНОГО ЧИСЛЕННЯ ТА МЕТОД КІНЦЕВИХ ЕЛЕМЕНТІВ

Розглянемо деякі поняття варіаційного числення [13].

Основна теорема математичного аналізу встановлює, що будь-яка функція $y=f(x)$, неперервна в області $a \leq x \leq b$, досягає своїх мінімальних і максимальних значень в цій області. Тобто виконується рівність:

$$df(x)/dx = 0, \quad (4.1)$$

з якої можна визначити значення x_0 , де $f(x)$ досягає свого оптимального значення.

Твердження (4.1) може бути сформульоване в такій формі. Якщо функція $y=f(x)$, неперервна в області $a \leq x \leq b$, досягає оптимального значення в точці $x=x_0$, то похідна df/dx в цій точці повинна дорівнювати нулеві. Отже, перша варіація функції - δf , яка іноді означається як диференціал df , дорівнює нулеві для будь-якої змінної δx в x , тобто

$$\delta f = (df/dx)\delta x = 0. \quad (4.2)$$

Подальше диференціювання функції $y=f(x)$ визначає відносний мінімум, максимум або мінімаксну умову функції $f(x)$ в точці $x=x_0$:

$$\begin{aligned} &\text{мінімум, якщо } d^2f/dx^2 > 0 \quad \text{при } x=x_0, \\ &\text{максимум, якщо } d^2f/dx^2 < 0 \quad \text{при } x=x_0, \\ &\text{мінімакс, якщо } d^2f/dx^2 = 0 \quad \text{при } x=x_0. \end{aligned} \quad (4.3)$$

Рівняння (4.1), (4.2) дають необхідні умови існування екстремуму функції $f(x)$ в точці x_0 , але вирази (4.3) вказують на недостатність умови. Рівняння (4.2) є необхідною і достатньою умовою стаціонарності функції $f(x)$ в точці $x=x_0$, тобто для $f(x)$ виконуються умови (4.3).

Розглянемо умову стаціонарності функції двох незалежних змінних $z=f(x,y)$. Якщо функція z стаціонарна в точці (x_0, y_0) , то x_0 і y_0 будуть розв'язками системи рівнянь:

$$\partial f / \partial x = 0, \quad (4.4)$$

$$\partial f / \partial y = 0. \quad (4.5)$$

З іншого боку, якщо $z=f(x,y)$ стаціонарна в точці (x_0, y_0) , то варіація цієї функції повинна бути нульовою для будь-яких варіацій δx і δy , тобто

$$\delta f = \left(\frac{\partial f}{\partial x} \right) \delta x + \left(\frac{\partial f}{\partial y} \right) \delta y = 0 \quad (4.6)$$

в точці $(x,y)=(x_0,y_0)$. Рівняння (4.1) і (4.2) справедливі через довільність δx і δy .

Формулу (4.6) можна узагальнити і для функції n незалежних змінних: $f(x_1, x_2, \dots, x_n)$. Якщо існує точка $(x_1, x_2, \dots, x_n) = (x_1^0, x_2^0, \dots, x_n^0) = X^0$, в якій функція $f(x_i)$ стаціонарна, то її необхідною і достатньою умовою стаціонарності є рівність нулеві в цій точці першої варіації для довільних варіацій δx_i :

$$i=\overline{1,n}, \text{ тобто } \delta f = \left\{ \sum_{j=1}^n \frac{\partial f(x_i)}{\partial x_j} \Big|_{X^0} \cdot \delta x_j \right\} = \sum_{j=1}^n \frac{\partial f(X^0)}{\partial x_j} \delta x_j = 0, \quad (4.7)$$

де $\frac{\partial f(X^0)}{\partial x_j} = \frac{\partial f(x_i)}{\partial x_j} \Big|_{X^0}$ – частинна похідна від $f(x_i)$ на x_i , яка вирахована в точці X^0 .

Рівняння (4.7) справедливе для будь-яких варіацій $\delta x_j, j = \overline{1, n}$. Нехай для незалежної змінної x_k вибрана нульова варіація $\delta x_k (\delta x_j = 0, j = \overline{1, n}, j \neq k)$. Тоді (4.7) запишеться у вигляді:

$$\delta f = \frac{\partial f(X^0)}{\partial x_k} \delta x_k = 0. \quad (4.8)$$

Звідси маємо, що:

$$\frac{\partial f(X^0)}{\partial x_k} = 0. \quad (4.9)$$

Оскільки незалежна змінна x_k була вибрана довільно, то вираз (4.7) справедливий для всіх k . Тому необхідна і достатня умова стаціонарності функції $f(x_i)$ в точці X^0 зводиться до системи сумісних рівнянь:

$$\begin{aligned} \frac{\partial f(X^0)}{\partial x_i} &= 0, \\ i &= \overline{1, n}. \end{aligned} \quad (4.10)$$

4.1 Метод скінченних елементів

4.1.1 Загальні положення

МСЕ містить основні концепції методу сіток, пов'язані з дискредитацією областей неперервної зміни методу Гальоркіна, в яких використовуються варіаційні принципи для знаходження функцій у вузлах. В результаті розв'язання крайової задачі зводиться до розв'язання системи лінійних алгебраїчних рівнянь (порядок яких визначається числом вузлів).

Основна ідея МСЕ полягає в тому, що неперервну величину (температуру, тиск і переміщення) можна апроксимувати дискретною моделлю, яка будується на множині кусково-неперервних підобластей. Кусково-неперервні функції визначаються за допомогою значень неперервної величини в кінцевому числі точок розглядуваної області.

Для побудови дискретної моделі неперервної величини чинять так.

- 1) В розглядуваній області фіксують кінцеве число точок (вузлів).
- 2) Значення неперервної величини в кожній вузловій точці вважають змінною, яка повинна бути визначена.
- 3) Область визначення неперервної величини розбивають на кінцеве число підобластей – елементів. Елементи мають загальні вузлові точки і в сукупності апроксимують форму області.
- 4) Неперервна величина апроксимується на кожному елементі поліномом, який визначають за допомогою вузлових значень цієї величини. Для кожного елемента вибирають свій поліном таким чином, щоб збереглась неперервність величини вздовж меж елемента.

4.1.2 Побудова елемента

Розбити область на елементи можна різними способами. Якщо обмежити кожний елемент двома сусідніми вузловими точками, то будуть створені чотири елементи (рис.4.1,а), або розбити область на два елементи, кожний з яких містить три вузли (рис.4.1,б). Відповідний поліном для кожного елемента визначають за значеннями $T(x)$ у вузлових точках елемента. Так, для елементів (рис.4.1,а) функції будуть лінійні за x , а для елементів з трьома вузловими точками функції будуть у вигляді полінома другого степеня (рис.4.1,б).

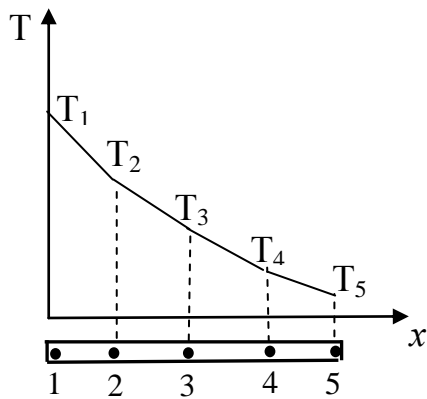


Рисунок 4.1,а

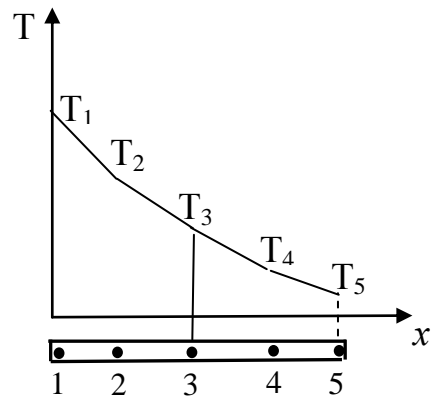


Рисунок 4.1,б

В цьому випадку апроксимацією $T(x)$ є сукупність двох кусково-неперервних квадратичних функцій.

Якщо значення функції $T(x)$ у вузлових точках невідомі, то визначають множину вузлів і значення температури у вузлах T_1, T_2, T_3, \dots , які тепер є змінними, оскільки явно невідомі. Область розбивають на елементи, на кожен з яких визначають відповідну функцію елемента. Вузлові значення $T(x)$ повинні забезпечувати найкраще наближення до справжнього розподілення функцій $T(x)$. Для цього потрібно мінімізувати деяку величину, яка пов'язана з фізичною суттю задачі. Якщо розглядається задача розповсюдження, наприклад, теплоти, то мінімізується функціонал, пов'язаний з відповідними диференціальними функціями. Процес мінімізації зводиться до розв'язування системи лінійних алгебраїчних рівнянь відносно вузлових значень $T(x)$.

При побудові дискретної моделі неперервної величини, яка визначена в дво- або тривимірній області, основну ідею методу скінченних елементів використовують аналогічно.

В двовимірному випадку елементи описуються функціями x і y (трикутники або чотирикутники, тобто функція елемента подається площиною).

Якщо число вузлів, що використовується, більше мінімального, то функції елемента відповідає криволінійна поверхня. Крім того, надлишкове число вузлів дозволяє розглянути елементи з криволінійними границями. Кінцевою апроксимацією двовимірної неперервної величини $\varphi(x,y)$ служить сукупність кусково-неперервних поверхностей, кожна з яких визначається на окремому елементі за допомогою значень $\varphi(x,y)$ у відповідних вузлових точках.

4.1.3 Побудова інтерполяційних многочленів

а) Одновимірний лінійний елемент.

Елемент містить два вузли, невідому функцію записують у вигляді:

$$\varphi = \alpha_1 + \alpha_2 x. \quad (4.1.1)$$

Коефіцієнти α виражаються через значення функцій у вузлах:

$$\varphi = N_i \Phi_i + N_j \Phi_j,$$

де N_i, N_j – функції форми;

Φ_i, Φ_j значення φ у вузлах i, j .

$$N_i = (x_j - x) / l, \quad N_j = (x - x_i) / l,$$

де x_i, x_j – координати вузлів елементів;

l – довжина елемента.

б) Двовимірний лінійний елемент, містить три вузли, шукана функція має вигляд:

$$\varphi = \alpha_1 + \alpha_2 x + \alpha_3 y. \quad (4.1.2)$$

Коефіцієнти α виражаються через значення шуканої функції у вузлах (рис.4.2):

$$\varphi = N_i \Phi_i + N_j \Phi_j + N_k \Phi_k, \quad (4.1.3)$$

$$N_i = \frac{1}{2\Delta} (a_i + b_i x + c_i y), \quad N_j = \frac{1}{2\Delta} (a_j + b_j x + c_j y), \quad N_k = \frac{1}{2\Delta} (a_k + b_k x + c_k y),$$

$$2\Delta = \det \begin{vmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_k & y_k \end{vmatrix} \neq 0 \quad (\text{оскільки площа трикутника ніколи не дорівнює ну-$$

леві).

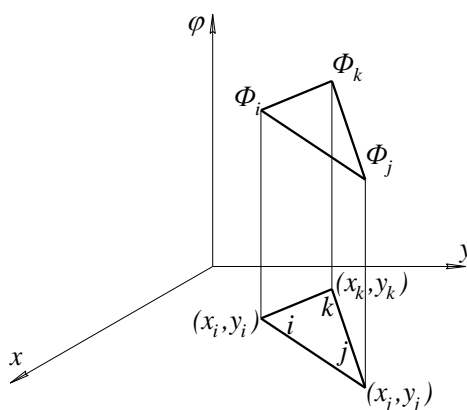


Рисунок 4.2

$$\begin{aligned}\Phi_i &= \alpha_1 + \alpha_2 x_i + \alpha_3 y_i, \quad \Phi_j = \alpha_1 + \alpha_2 x_j + \alpha_3 y_j, \\ \Phi_k &= \alpha_1 + \alpha_2 x_k + \alpha_3 y_k.\end{aligned}\tag{4.1.4}$$

Розв'язуючи систему (4.1.4) відносно $\alpha_1, \alpha_2, \alpha_3$, отримуємо такі вирази:

$$\begin{aligned}\alpha_1 &= \left(\frac{1}{2\Delta}\right)(a_i \Phi_i + a_j \Phi_j + a_k \Phi_k), \\ \alpha_2 &= \left(\frac{1}{2\Delta}\right)(b_i \Phi_i + b_j \Phi_j + b_k \Phi_k), \\ \alpha_3 &= \left(\frac{1}{2\Delta}\right)(c_i \Phi_i + c_j \Phi_j + c_k \Phi_k),\end{aligned}\tag{4.1.5}$$

де

$$\begin{aligned}a_i &= x_j y_k - x_k y_j, \quad b_i = y_j - y_k, \quad c_i = x_k - x_j, \\ a_j &= x_k y_i - x_i y_k, \quad b_j = y_k - y_i, \quad c_j = x_i - x_k, \\ a_k &= x_i y_j - x_j y_i, \quad b_k = y_i - y_j, \quad c_k = x_j - x_i.\end{aligned}\tag{4.1.6}$$

4.1.4 Одновимірний приклад варіаційного методу скінченних елементів

Розглянемо одиничну масу, яка рухається під дією сили, що пропорційна пройденому шляху y (коефіцієнт пропорційності $k = 1$). Тоді диференціальне рівняння руху запишеться у вигляді:

$$\frac{d^2 y}{dt^2} - y = 0.\tag{4.1.7}$$

Необхідно знайти пройдений шлях за проміжок часу від $t = 0$ до $t = 2$, якщо дано:

$$\begin{aligned}y(0) &= 1, \\ y(2) &= e^2 = 7,389.\end{aligned}\tag{4.1.8}$$

Відомо, що розв'язок рівняння (4.1.7) можна отримати подвійним інтегруванням (4.1.7) з врахуванням початкових і кінцевих умов:

$$\begin{aligned}\frac{\dot{y} dy}{dy} &= y, \\ \dot{y} dy &= y dy, \\ \dot{y} &= y + c_1, \\ \int \frac{dy}{y + c_1} &= dt, \\ \ln|y + c_1| &= t + c_2, \\ y + c_1 &= e^{t+c_2}, \quad y = e^{t+c_2} - c_1.\end{aligned}$$

При $t = 0$, $y(0) = 1 = e^{c_2} - c_1$,

$t = 2$, $y(2) = 7,389 = 7,389e^{c_2} - c_1$.

Звідки $6,389 = 6,389e^{c_2}$,

$1 = e^{c_2} \Rightarrow c_2 = 0$,

$1 = e^0 - c_1 \Rightarrow c_1 = 0$.

Отже, $y = e^t$.

Розв'язок рівняння (4.1.7) збігається з функцією, що мінімізує функціонал

$$X = \frac{1}{2} \int_0^2 \left[\hat{y}^2 + \left(\frac{d\hat{y}}{dt} \right)^2 \right] dt \quad (4.1.9)$$

за умови, що пробні функції \hat{y} неперервні, мають кусково-неперервні перші похідні і задовольняють умови (4.1.8).

Область $[0 \leq t \leq 2]$ розіб'ємо на n кінцевих елементів e_1, e_2, \dots, e_n і у будемо апроксимувати пробною функцією \hat{y}^{e_i} в середині кожного елемента (рис.4.3).

Функціонал X запишемо у вигляді суми елементів інтегралів:

$$X = \sum_{i=1}^n \frac{1}{2} \int_{t_i}^{t_{i+1}} \left[(\hat{y})^2 + \left(\frac{d\hat{y}}{dt} \right)^2 \right] dt. \quad (4.1.10)$$

Визначаючи елементний внесок в формулу (4.1.10) через X^{e_i} , можна записати:

$$X = \sum_{i=1}^n X^{e_i}. \quad (4.1.11)$$

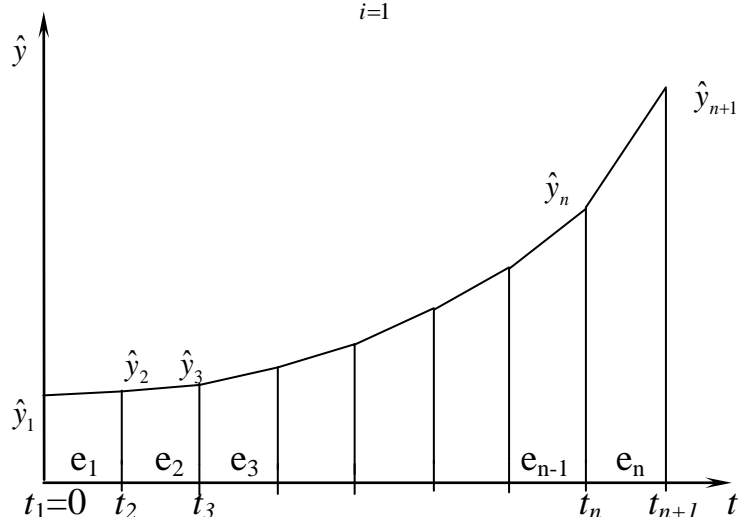


Рисунок 4.3

Виберемо пробну функцію \hat{y}^{e_i} у вигляді:

$$\begin{aligned}\hat{y}^{e_i} &= \alpha_1^{e_i} + \alpha_2^{e_i} t, \\ t_i &\leq t \leq t_{i+1},\end{aligned}\tag{4.1.12}$$

де $\alpha_1^{e_i}$, $\alpha_2^{e_i}$ - постійні, які можуть бути визначені із умов у вузлах:

$$\begin{aligned}\hat{y}^{e_i}(t_i) &= \bar{y}_i, \\ \hat{y}^{e_i}(t_{i+1}) &= \bar{y}_{i+1}.\end{aligned}\tag{4.1.13}$$

Отримаємо:

$$\begin{aligned}\alpha_1^{e_i} &= \frac{t_{i+1}\bar{y}_i - t_i\bar{y}_{i+1}}{t_{i+1} - t_i}, \\ \alpha_2^{e_i} &= \frac{\bar{y}_{i+1} - \bar{y}_i}{t_{i+1} - t_i}.\end{aligned}\tag{4.1.14}$$

З врахуванням (4.1.14) і (4.1.12) отримуємо елементну апроксимацію у вигляді:

$$\hat{y}^{e_i} = \left(\frac{t_{i+1} - t}{t_{i+1} - t_i} \right) \bar{y}_i + \left(\frac{t - t_i}{t_{i+1} - t_i} \right) \bar{y}_{i+1}.\tag{4.1.15}$$

Визначаючи коефіцієнти при \bar{y}_i і \bar{y}_{i+1} як базисні функції $N_i^{e_i}$ і $N_{i+1}^{e_i}$, рівність (4.1.15) запишемо у вигляді апроксимації базисними функціями:

$$\hat{y}^{e_i} = N_i^{e_i} \bar{y}_i + N_{i+1}^{e_i} \bar{y}_{i+1}.\tag{4.1.16}$$

Зауважимо, що $N_i^{e_i}$ і $N_{i+1}^{e_i}$ залежать тільки від t і зовні елемента e_i пробна функція \bar{y}^{e_i} дорівнює нулеві.

Рівність (4.1.16) запишемо у матричній формі:

$$\bar{y}^{e_i} = N^{e_i} y^{e_i},\tag{4.1.17}$$

$$\text{де } N^{e_i} = \begin{bmatrix} N_i^{e_i} & N_{i+1}^{e_i} \end{bmatrix}, \quad y^{e_i} = \begin{bmatrix} \bar{y}_i \\ \bar{y}_{i+1} \end{bmatrix}^{e_i}.\tag{4.1.18}$$

Величину N^{e_i} називають матрицею базисних функцій, а y^{e_i} - вузловим вектором.

Якщо елементи e_i вибрані однакової довжини, тоді

$$t_{i+1} - t_i = (t_{n+1} - t_1)/n, \quad i = \overline{1, n}.\tag{4.1.19}$$

З врахуванням значень t з умови задачі рівність (4.1.19) запишеться у вигляді:

$$t_{i+1} - t_i = 2/n.\tag{4.1.20}$$

Підставимо (4.1.20) в (4.1.15), отримаємо елементарну пробну функцію:

$$\hat{y}^{e_i} = \frac{n}{2} [(t_{i+1} - t)\bar{y}_i + (t - t_i)\bar{y}_{i+1}]\tag{4.1.21}$$

Підстановка (4.1.21) в X^{e_i} дає такий вираз елементного вкладу:

$$X^{e_i} = \frac{n^2}{8} \int_{t_i}^{t_{i+1}} \left([(t_{i+1} - t)\bar{y}_i + (t - t_i)\bar{y}_{i+1}]^2 + (\bar{y}_{i-1} - \bar{y}_i)^2 \right) dt. \quad (4.1.22)$$

Проінтегруємо (4.1.22) з врахуванням (4.1.20), маємо:

$$X^{e_i} = \frac{n}{4} \left[\left(\frac{4}{3n^2} + 1 \right) \bar{y}_i^2 + \left(\frac{4}{3n^2} - 2 \right) \bar{y}_i \bar{y}_{i+1} + \left(\frac{4}{3n^2} + 1 \right) \bar{y}_{i+1}^2 \right]. \quad (4.1.23)$$

Підставляючи (4.1.23) в (4.1.11), приходимо до висновку, що X є функцією вузлових значень $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n, \bar{y}_{n+1}$, тобто

$$X = X(\bar{y}_1, \bar{y}_2, \dots, \bar{y}_n, \bar{y}_{n+1}). \quad (4.1.24)$$

З варіаційного числення відомо, що мінімумом для функціонала X є:

$$\frac{\partial X}{\partial \bar{y}_p} = \sum_{i=1}^n \frac{\partial X^{e_i}}{\partial \bar{y}_p} = 0, \quad (4.1.25)$$

$$p = \overline{1, n+1}.$$

Вузлові значення \bar{y}_1 і \bar{y}_{n+1} замінимо постійними з врахуванням (4.1.8): $\bar{y}_1 = 1, \bar{y}_{n+1} = 7,389$. Отже, в (4.1.25) p буде приймати значення 2, 3, ..., n . З рівності (4.1.23) визначимо $\frac{\partial X^{e_i}}{\partial \bar{y}_i}$:

$$\frac{\partial X^{e_i}}{\partial \bar{y}_i} = \frac{n}{4} \left[2 \left(\frac{4}{3n^2} + 1 \right) \bar{y}_i + \left(\frac{4}{3n^2} - 2 \right) \bar{y}_{i+1} \right]. \quad (4.1.26)$$

Аналогічно для елемента e_{i-1} маємо:

$$\frac{\partial X^{e_{i-1}}}{\partial \bar{y}_i} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) \bar{y}_{i-1} + 2 \left(\frac{4}{3n^2} + 1 \right) \bar{y}_i \right]. \quad (4.1.27)$$

Підставляючи (4.1.26) і (4.1.27) в (4.1.25), приходимо до висновку, що в рівняння (4.1.25) дають внесок тільки елементи e_i, e_{i-1} , $i = \overline{2, n}$:

$$\frac{\partial X}{\partial \bar{y}_i} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) \bar{y}_{i-1} + 4 \left(\frac{4}{3n^2} + 1 \right) \bar{y}_i + \left(\frac{4}{3n^2} - 2 \right) \bar{y}_{i+1} \right] = 0. \quad (4.1.28)$$

Для $i = 1$ та $i = n+1$ маємо:

$$\frac{\partial X}{\partial \bar{y}_1} = \frac{n}{4} \left[2 \left(\frac{4}{3n^2} + 1 \right) \bar{y}_1 + \left(\frac{4}{3n^2} - 2 \right) \bar{y}_2 \right] = 0, \quad (4.1.29)$$

$$\frac{\partial X}{\partial \bar{y}_{i+1}} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) \bar{y}_n + 2 \left(\frac{4}{3n^2} + 1 \right) \bar{y}_{i+1} \right] = 0. \quad (4.1.30)$$

Внесок в формули (4.1.29), (4.1.30) дають, відповідно, тільки елементи e_1 і e_n .

Підставляючи рівняння (4.1.28) – (4.1.30) в (4.1.25), отримуємо систему $n+1$ лінійних алгебраїчних рівнянь, яку запишемо в матричній формі:

$$\begin{cases} \frac{\partial X^{e_2}}{\partial y_2} = \frac{n}{4} \left[2 \left(\frac{4}{3n^2} + 1 \right) y_2 + \left(\frac{4}{3n^2} - 2 \right) y_3 \right] = \frac{104}{48} y_2 - \frac{92}{48} y_3 \\ \frac{\partial X^{e_2}}{\partial y_3} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) y_2 + 2 \left(\frac{4}{3n^2} + 1 \right) y_3 \right] = -\frac{92}{48} y_2 + \frac{104}{48} y_3 \end{cases}, \quad (4.1.34)$$

$$\begin{cases} \frac{\partial X^{e_3}}{\partial y_3} = \frac{n}{4} \left[2 \left(\frac{4}{3n^2} + 1 \right) y_3 + \left(\frac{4}{3n^2} - 2 \right) y_4 \right] = \frac{104}{48} y_3 - \frac{92}{48} y_4 \\ \frac{\partial X^{e_3}}{\partial y_4} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) y_3 + 2 \left(\frac{4}{3n^2} + 1 \right) y_4 \right] = -\frac{92}{48} y_3 + \frac{104}{48} y_4 \end{cases}, \quad (4.1.35)$$

$$\begin{cases} \frac{\partial X^{e_4}}{\partial y_4} = \frac{n}{4} \left[2 \left(\frac{4}{3n^2} + 1 \right) y_4 + \left(\frac{4}{3n^2} - 2 \right) y_5 \right] = \frac{104}{48} y_4 - \frac{92}{48} y_5 \\ \frac{\partial X^{e_4}}{\partial y_5} = \frac{n}{4} \left[\left(\frac{4}{3n^2} - 2 \right) y_4 + 2 \left(\frac{4}{3n^2} + 1 \right) y_5 \right] = -\frac{92}{48} y_4 + \frac{104}{48} y_5 \end{cases}. \quad (4.1.36)$$

$$\frac{\partial X^{e_1}}{\partial y^{e_1}} = \begin{bmatrix} \frac{104}{48} & -\frac{92}{48} \\ -\frac{92}{48} & \frac{104}{48} \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}; \quad \frac{\partial X^{e_2}}{\partial y^{e_2}} = \begin{bmatrix} \frac{104}{48} & -\frac{92}{48} \\ -\frac{92}{48} & \frac{104}{48} \end{bmatrix} \begin{pmatrix} y_2 \\ y_3 \end{pmatrix}; \quad (4.1.37)$$

$$\frac{\partial X^{e_3}}{\partial y^{e_3}} = \begin{bmatrix} \frac{104}{48} & -\frac{92}{48} \\ -\frac{92}{48} & \frac{104}{48} \end{bmatrix} \begin{pmatrix} y_3 \\ y_4 \end{pmatrix}; \quad \frac{\partial X^{e_4}}{\partial y^{e_4}} = \begin{bmatrix} \frac{104}{48} & -\frac{92}{48} \\ -\frac{92}{48} & \frac{104}{48} \end{bmatrix} \begin{pmatrix} y_4 \\ y_5 \end{pmatrix}.$$

Матричну систему рівнянь (4.1.31) п'ятого порядку отримуємо за рахунок асамблювання матриць (4.1.37) та з врахуванням (4.1.25).

Асамблювання матриць для першого та другого елементів таке:

$$e_1 \Rightarrow \begin{pmatrix} \frac{104}{48} & -\frac{92}{48} & 0 & 0 & 0 \\ \frac{92}{48} & \frac{104}{48} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \quad e_2 \Rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{104}{48} & -\frac{92}{48} & 0 & 0 \\ 0 & \frac{92}{48} & \frac{104}{48} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix};$$

$$e_3 \Rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{104}{48} & -\frac{92}{48} & 0 \\ 0 & 0 & \frac{92}{48} & \frac{104}{48} & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}; \quad e_4 \Rightarrow \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{104}{48} & -\frac{92}{48} \\ 0 & 0 & 0 & \frac{92}{48} & \frac{104}{48} \end{pmatrix}.$$

Складемо отримані матриці. В результаті маємо систему рівнянь:

$$\begin{pmatrix} 104 & -92 & 0 & 0 & 0 \\ 48 & 48 & 0 & 0 & 0 \\ 92 & 208 & -92 & 0 & 0 \\ 48 & 48 & 48 & 0 & 0 \\ 0 & -92 & 208 & -92 & 0 \\ 0 & 48 & 48 & 48 & 0 \\ 0 & 0 & -92 & 208 & -92 \\ 0 & 0 & 48 & 48 & 48 \\ 0 & 0 & 0 & -92 & 104 \\ 0 & 0 & 0 & 48 & 48 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Після елементарних перетворень системи отримаємо:

$$\begin{pmatrix} 52 & -46 & 0 & 0 & 0 \\ -46 & 104 & -46 & 0 & 0 \\ 0 & -46 & 104 & -46 & 0 \\ 0 & 0 & -46 & 104 & -46 \\ 0 & 0 & 0 & -46 & 52 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Використовуючи перше правило для вузлових точок з умовами Діріхле (якщо p - вузол, в якому вузлове значення задається точно, тобто $\bar{y}_p = y_p$, то процедура введення умов Діріхле в матричне рівняння системи полягає в наступному: в p -ий рядок матриці системи вносяться нулі, крім діагональної позиції, на яку поміщається 1, в p -ий рядок стовпця вноситься y_p), маємо:

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -46 & 104 & -46 & 0 & 0 \\ 0 & -46 & 104 & -46 & 0 \\ 0 & 0 & -46 & 104 & -46 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \bar{y}_3 \\ \bar{y}_4 \\ \bar{y}_5 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ e^2 \end{pmatrix}.$$

Використавши метод Гаусса в програмному забезпеченні, маємо:

$$\bar{y}_1 = 1, \quad \bar{y}_2 = 1,643897, \quad \bar{y}_3 = 2,716637,$$

$$\bar{y}_4 = 4,498064, \quad \bar{y}_5 = 7,4529,$$

де $\bar{y}_1 \Rightarrow t_1 = 0$, $\bar{y}_2 \Rightarrow t_2 = 0,5$, $\bar{y}_3 \Rightarrow t_3 = 1$, $\bar{y}_4 \Rightarrow t_4 = 1,5$, $\bar{y}_5 \Rightarrow t_5 = 2$.

Аналітичний розв'язок:

$$y_1 = 1, \quad y_2 = e^{0,5}, \quad y_3 = e, \quad y_4 = e^{1,5}, \quad y_5 = e^2.$$

4.1.5 Двовимірний приклад варіаційного методу скінченних елементів

Як приклад розглянемо рівняння Лапласа:

$$\frac{\partial^2 T}{\partial x^2} + \frac{\partial^2 T}{\partial y^2} = 0 . \quad (4.1.38)$$

З граничними умовами Діріхле на одній ділянці границі:

$$T=50 , \quad y=0, \quad (4.1.39)$$

$$T=100 , \quad y=L. \quad (4.1.40)$$

З умовами Неймана на іншій ділянці границі:

$$\frac{\partial T}{\partial x} = 0, \quad x = 0, \quad (4.1.41)$$

$$\frac{\partial T}{\partial x} = 0, \quad x = L. \quad (4.1.42)$$

Задача визначена на множині R , яка складається з області D і межі S , тобто $R=D+S$.

За допомогою варіаційного числення розв'язок $T(x,y)$, що задовольняє рівняння (4.1.38)-(4.1.42), збігається з функцією, яка мінімізує функціонал

$$X = \frac{1}{2} \iint_D \left[\left(\frac{\partial \hat{T}}{\partial x} \right)^2 + \left(\frac{\partial \hat{T}}{\partial y} \right)^2 \right] dx dy, \quad (4.1.43)$$

де $\hat{T}(x,y)$ – функція з допустимої множини пробних функцій, заданих в D .

Граничні умови Неймана (4.1.41-4.1.42) виконуються автоматично для функції, яка мінімізує функціонал, і, як наслідок, варіаційні формулювання будуть називатися натуральними граничними умовами.

Розіб'ємо область на i скінченних елементів (трикутників). Загальне число вузлів позначимо через n .

В двовимірному і тривимірному випадках немає чіткого зв'язку між загальним числом елементів і загальним числом вузлів. Розбиття області і умови неперервності, що накладаються на пробні функції, дозволяють записати функціонал у вигляді:

$$X = \sum_{i=1}^n X^{e_i} , \quad (4.1.44)$$

де X^{e_i} – елементний вклад, що визначається рівністю:

$$X^{e_i} = \frac{1}{2} \iint_{e_i} \left[\left(\frac{\partial \hat{T}^{e_i}}{\partial x} \right)^2 + \left(\frac{\partial \hat{T}^{e_i}}{\partial y} \right)^2 \right] dx dy . \quad (4.1.45)$$

Для довільного елемента e_i пробну функцію виберемо у вигляді (4.1.2).

Враховуючи, що в даному випадку $\Phi_i = \hat{T}_i$, $\Phi_j = \hat{T}_j$, $\Phi_k = \hat{T}_k$, згідно з формулами (4.1.5-4.1.6) для кожного елемента отримуємо:

$$\hat{T}^{e_i}(x, y) = \frac{1}{2\Delta} \left[(a_i + \varepsilon_i x + c_i y) \hat{T}_i + (a_j + \varepsilon_j x + c_j y) \hat{T}_j + (a_k + \varepsilon_k x + c_k y) \hat{T}_k \right]. \quad (4.1.46)$$

Тоді маємо:

$$\frac{\partial \hat{T}^{e_i}}{\partial x} = \frac{1}{2\Delta} \left[\varepsilon_i \hat{T}_i + \varepsilon_j \hat{T}_j + \varepsilon_k \hat{T}_k \right], \quad (4.1.47)$$

$$\frac{\partial \hat{T}^{e_i}}{\partial y} = \frac{1}{2\Delta} \left[c_i \hat{T}_i + c_j \hat{T}_j + c_k \hat{T}_k \right]. \quad (4.1.48)$$

Підставимо (4.1.47)-(4.1.48) в (4.1.45):

$$X^{e_i} = \frac{1}{8\Delta} \iint_{e_i} \left[(\varepsilon_i \hat{T}_i + \varepsilon_j \hat{T}_j + \varepsilon_k \hat{T}_k)^2 + (c_i \hat{T}_i + c_j \hat{T}_j + c_k \hat{T}_k)^2 \right] dx dy. \quad (4.1.49)$$

Оскільки підінтегральний вираз в (4.1.49) не залежить від x і y , і, крім того, $\iint_{e_i} dx dy = \Delta$, то рівність (4.1.49) перепишемо в такому вигляді:

$$X^{e_i} = \frac{1}{8\Delta} \left[(\varepsilon_i \hat{T}_i + \varepsilon_j \hat{T}_j + \varepsilon_k \hat{T}_k)^2 + (c_i \hat{T}_i + c_j \hat{T}_j + c_k \hat{T}_k)^2 \right]. \quad (4.1.50)$$

Вираз вигляду (4.1.50) може бути отриманий для кожного елемента. Підставляючи (4.1.50) в (4.1.44), перетворимо функціонал (4.1.42) в функцію всіх вузлових значень $T_1 - T_n$, тобто

$$X = X(\hat{T}_1, \hat{T}_2, \dots, \hat{T}_n). \quad (4.1.51)$$

Умови мінімуму X можна записати у вигляді:

$$\frac{\partial X}{\partial \hat{T}_p} = 0, \quad p = \overline{1, n}. \quad (4.1.52)$$

Підставляючи (4.1.44) в (4.1.52), маємо:

$$\frac{\partial X}{\partial \hat{T}_p} = \sum_{i=1}^n \frac{\partial X^{e_i}}{\partial \hat{T}_p} = 0, \quad p = \overline{1, n}. \quad (4.1.53)$$

Диференціювання (4.1.50) за \hat{T}_p приводить до виразу:

$$\frac{\partial X^{e_i}}{\partial \hat{T}_p} = \frac{1}{4\Delta} \left[\varepsilon_p (\varepsilon_p \hat{T}_p + \varepsilon_q \hat{T}_q + \varepsilon_r \hat{T}_r) + c_p (c_p \hat{T}_p + c_q \hat{T}_q + c_r \hat{T}_r) \right], \quad (4.1.54)$$

де значення p, q, i, r – відповідно, номери вузлів i, j, k елемента e_i в множині номерів вузлів системи.

Об'єднання компонентів елементарних рівнянь, що задаються рівністю (4.1.54), називається об'єднанням за вузлами, тому що процес об'єднання повинен бути виконаний окремо для кожного вузла системи. Формування вкладів $\partial X^{e_i} / \partial T_p$ та їх об'єднання проілюструємо нижче.

Спочатку область D розіб'ємо на 16 елементів з загальним числом вузлів 15 (рис. 4.4) та запишемо номери вузлів i, j, k для кожного елемента системи в таблицю 4.1. Обхід кожного елемента будемо робити проти го-

динникової стрілки. Розміри всіх елементів однакові. Координати вузлів запишемо в таблицю 4.2.

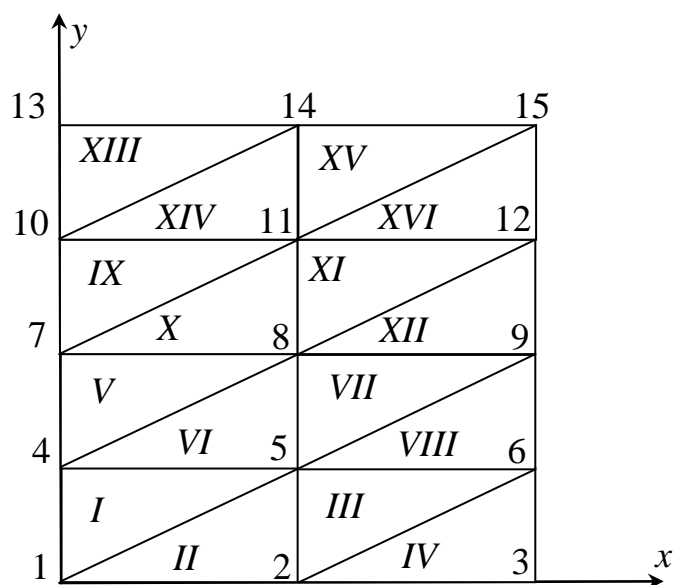


Рисунок 4.4

Таблиця 4.1 – Співвідношення між глобальними і локальними номерами вузлів

Елемент	Глобальний номер вузла			Елемент	Глобальний номер вузла		
	<i>I</i>	<i>J</i>	<i>k</i>		<i>I</i>	<i>J</i>	<i>K</i>
1	4	1	5	9	10	7	11
2	2	5	1	10	8	11	7
3	5	2	6	11	11	8	12
4	3	6	2	12	9	12	8
5	7	4	8	13	13	10	14
6	5	8	4	14	11	14	10
7	8	5	9	15	14	11	15
8	6	9	5	16	12	15	11

В таблиці 4.3 згідно з формулою (4.1.6) наведені v_i , v_j , v_k , c_i , c_j і c_k . Для ілюстрації розглянемо елемент 9. З таблиці 4.1 видно, що його вузли i , j , k мають номери 10, 7 і 11, відповідно.

Таблиця 4.2 – Координати вузлів

Вузол	Координати		Вузол	Координати		Вузол	Координати	
	x	Y		x	Y		X	Y
1	0	0	6	2	0,5	11	1	1,5
2	1	0	7	0	1	12	2	1,5
3	2	0	8	1	1	13	0	2
4	0	0,5	9	2	1	14	1	2
5	1	0,5	10	0	1,5	15	2	2

Таблиця 4.3 – Характерні параметри елементів

Елементи	Параметри					
	ν_i	ν_j	ν_k	c_i	c_j	c_k
1	-0,5	0	0,5	1	-1	0
2	0,5	0	-0,5	-1	1	0
3	-0,5	0	0,5	1	-1	0
4	0,5	0	-0,5	-1	1	0
5	-0,5	0	0,5	1	-1	0
6	0,5	0	-0,5	-1	1	0
7	-0,5	0	0,5	1	-1	0
8	0,5	0	-0,5	-1	1	0
9	-0,5	0	0,5	1	-1	0
10	0,5	0	-0,5	-1	1	0
11	-0,5	0	0,5	1	-1	0
12	0,5	0	-0,5	-1	1	0
13	-0,5	0	0,5	1	-1	0
14	0,5	0	-0,5	-1	1	0
15	-0,5	0	0,5	1	-1	0
16	0,5	0	-0,5	-1	1	0

Підставляючи ці дані і відповідні значення параметрів для елемента 9 з таблиці 4.3 в рівність (4.1.50), отримуємо вираз:

$$X^9 = \frac{1}{8\Delta^9} \left[(-0,5 \cdot T_{10} + 0 \cdot T_7 + 0,5 \cdot T_{11})^2 + (1 \cdot T_{10} - 1 \cdot T_7 + 0 \cdot T_{11})^2 \right]. \quad (4.1.55)$$

Верхній індекс 9 у площі Δ вказує на номер елемента, для якого вона визначена. Оскільки площі всіх елементів однакові $\Delta^{e_i} = \frac{1}{2} \cdot 1 \cdot \frac{1}{2} = 0,25$, то в подальшому верхній не будемо зображувати.

Згідно з (4.1.55) визначимо похідні $\partial X^{e_i} / \partial T_p$:

$$\frac{\partial X^9}{\partial T_{10}} = \frac{1}{4\Delta} \left[-0,5(-0,5 \cdot T_{10} + 0 \cdot T_7 + 0,5 \cdot T_{11}) + 1(1 \cdot T_{10} - 1 \cdot T_7 + 0 \cdot T_{11}) \right], \quad (4.1.56)$$

$$\frac{\partial X^9}{\partial T_7} = \frac{1}{4\Delta} [0(-0,5 \cdot T_{10} + 0 \cdot T_7 + 0,5 \cdot T_{11}) - 1(1 \cdot T_{10} - 1 \cdot T_7 + 0 \cdot T_{11})], \quad (4.1.57)$$

$$\frac{\partial X^9}{\partial T_{11}} = \frac{1}{4\Delta} [0,5(-0,5 \cdot T_{10} + 0 \cdot T_7 + 0,5 \cdot T_{11}) + 0(1 \cdot T_{10} - 1 \cdot T_7 + 0 \cdot T_{11})]. \quad (4.1.58)$$

Рівняння (4.1.56-4.1.58) можна записати у вигляді елементарного матричного рівняння:

$$\begin{bmatrix} \frac{\partial X^9}{\partial T_{10}} \\ \frac{\partial X^9}{\partial T_7} \\ \frac{\partial X^9}{\partial T_{11}} \end{bmatrix} = \frac{1}{16\Delta} \begin{bmatrix} 5 & -4 & -1 \\ -4 & 4 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_{10} \\ T_7 \\ T_{11} \end{bmatrix}, \quad (4.1.59)$$

або
$$\frac{\partial X^e}{\partial T^e} = k^e T^e, \quad (4.1.60)$$

де
$$k^e = \begin{bmatrix} k_{ii}^e & k_{ij}^e & k_{ik}^e \\ k_{ji}^e & k_{jj}^e & k_{jk}^e \\ k_{ki}^e & k_{kj}^e & k_{kk}^e \end{bmatrix} \quad (4.1.61)$$

елементарна матриця жорсткості елемента e і T^e – елементарний вузловий вектор. Для спрощення записів в подальшому верхній індекс e_i в рівняннях (4.1.60) і (4.1.61) будемо опускати.

Оскільки елементи 1, 3, 5, 7, 9, 11, 13, 15 мають однакові розміри і орієнтацію відносно системи Oxy , то елементарна матриця жорсткості k однакова з врахуванням (4.1.60) у відповідних номерах вузлів. Так, для елемента 5 матричне рівняння має вигляд:

$$\begin{bmatrix} \frac{\partial X^5}{\partial T_7} \\ \frac{\partial X^5}{\partial T_4} \\ \frac{\partial X^5}{\partial T_8} \end{bmatrix} = \frac{1}{16\Delta} \begin{bmatrix} 5 & -4 & -1 \\ -4 & 4 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_7 \\ T_4 \\ T_8 \end{bmatrix}. \quad (4.1.62)$$

Інші елементи 2, 4, 6, 8, 10, 12, 14, 16 також мають загальну елементну матрицю жорсткості k , яка в даному випадку тотожна матриці жорсткості з непарними номерами. Так, для елемента 10, матричне рівняння має вигляд:

$$\begin{bmatrix} \frac{\partial X^{10}}{\partial T_8} \\ \frac{\partial X^{10}}{\partial T_{11}} \\ \frac{\partial X^{10}}{\partial T_7} \end{bmatrix} = \frac{1}{16\Delta} \begin{bmatrix} 5 & -4 & -1 \\ -4 & 4 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} T_8 \\ T_{11} \\ T_7 \end{bmatrix}. \quad (4.1.63)$$

Після визначення похідних $\partial X^e / \partial T_p$ для всіх елементів покажемо їх об'єднання по вузлах згідно з формулою (4.1.53) для отримання матричного рівняння системи. Наприклад, для вузла 7 дають внесок лише елементи 5, 9 і 10, отже, маємо:

$$\frac{\partial X}{\partial T_7} = \frac{\partial X^5}{\partial T_7} + \frac{\partial X^9}{\partial T_7} + \frac{\partial X^{10}}{\partial T_7}. \quad (4.1.64)$$

Підставляючи в (4.1.64) внески елементів 5, 9 і 10 з врахуванням (4.1.59), (4.1.62) і (4.1.63), отримаємо:

$$\begin{aligned} \frac{\partial X}{\partial T_7} &= \frac{1}{16\Delta} \left\{ [5 \quad -4 \quad -1] \cdot \begin{bmatrix} T_7 \\ T_4 \\ T_8 \end{bmatrix} + [-4 \quad 4 \quad 0] \cdot \begin{bmatrix} T_{10} \\ T_7 \\ T_{11} \end{bmatrix} + [-1 \quad 0 \quad 1] \cdot \begin{bmatrix} T_8 \\ T_{11} \\ T_7 \end{bmatrix} \right\} = \\ &= \frac{1}{16\Delta} [-4T_4 + 10T_7 - 2T_8 - 4T_{10}] = 0. \end{aligned} \quad (4.1.65)$$

Рівняння (4.1.65) запишемо в розширеній формі:

$$\frac{\partial X}{\partial T_7} = \frac{1}{16\Delta} [0 \quad 0 \quad 0 \quad -4 \quad 0 \quad 0 \quad 10 \quad -2 \quad 0 \quad -4 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \cdot T = 0, \quad (4.1.66)$$

де вузловий вектор T системи має вигляд:

$$T = \begin{bmatrix} T_1 \\ T_2 \\ \vdots \\ T_{15} \end{bmatrix}. \quad (4.1.67)$$

Для кожного вузла системи рівняння аналогічні (4.1.65).

В результаті матричне рівняння системи має вигляд:

$$\frac{\partial X}{\partial T} = \begin{bmatrix} \partial X / \partial T_1 \\ \partial X / \partial T_2 \\ \vdots \\ \partial X / \partial T_{15} \end{bmatrix} = kT = 0. \quad (4.1.68)$$

Для корегування отриманої системи необхідно враховувати граничні умови Діріхле. Оскільки у вузлах 1, 2, 3 і 13, 14, 15 значення задані явно

((4.1.39),(4.1.40)), то, згідно з методом Пейна-Айронса, в рядках матриці k 1, 2, 3, 13, 14 і 15 внесемо нулі, крім діагональної позиції, на яку запишемо одиницю, а в відповідні рядки стовпця в правій частині вводимо значення вузлів. В результаті отримаємо таку матричну систему рівнянь:

$$\begin{bmatrix}
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 -4 & 0 & 0 & 10 & -2 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & -8 & 0 & -2 & 20 & -2 & 0 & -8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & -4 & 0 & -2 & 10 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & -4 & 0 & 0 & 10 & -2 & 0 & -4 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & -8 & 0 & -2 & 20 & -2 & 0 & -8 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & -4 & 0 & -2 & 10 & 0 & 0 & -4 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 10 & -2 & 0 & -4 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & -8 & 0 & -2 & 20 & -2 & 0 & -8 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & -2 & 10 & 0 & 0 & -4 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix} \cdot \begin{bmatrix} T_1 \\ T_2 \\ T_3 \\ T_4 \\ T_5 \\ T_6 \\ T_7 \\ T_8 \\ T_9 \\ T_{10} \\ T_{11} \\ T_{12} \\ T_{13} \\ T_{14} \\ T_{15} \end{bmatrix} = \begin{bmatrix} 50 \\ 50 \\ 50 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 100 \\ 100 \\ 100 \end{bmatrix} \quad (4.1.69)$$

Використовуючи чисельні методи (програми) розрахунку системи алгебраїчних рівнянь, визначаємо розв'язок (4.1.69):

$$\begin{aligned}
 T_1 &= 50, & T_2 &= 50, & T_3 &= 50 \\
 T_4 &= 62,5, & T_5 &= 62,5, & T_6 &= 62,5 \\
 T_7 &= 75, & T_8 &= 75, & T_9 &= 75 \\
 T_{10} &= 87,5, & T_{11} &= 87,5, & T_{12} &= 87,5 \\
 T_{13} &= 100, & T_{14} &= 100, & T_{15} &= 100.
 \end{aligned}$$

Отже, наприклад, для досліджуваної області в точці A з координатами $(l/2, l/4)$ температура $T/A \approx T_{11} = 87,5^\circ\text{C}$.

У ПОСТАНОВКА ЗАДАЧ ЛІНІЙНОГО ПРОГРАМУВАННЯ

Одним з методів розв'язання задач оптимізації є метод програмування, який полягає в побудові, аналізі і чисельному розв'язанні лінійних оптимізаційних моделей. Задача лінійного програмування має місце лише тоді, коли обмеження і цільова функція об'єкта управління лінійні відносно факторів X . Метод використовується для розв'язання виробничих і управлінських задач в різноманітних галузях народного господарства, якщо параметри об'єкта моделювання відповідають вимогам аксіом пропорційності (ділення) і адитивності.

Принцип ділення (пропорційності), наприклад, у виробничому процесі означає, що загальний (сумарний) об'єм застосування кожного виду ресурсів і, відповідно, отриманий при цьому прибуток підприємства прямо пропорційні об'єму виробленої продукції. Так, якщо збільшити вдвічі об'єм продукції, то вживання усіх ресурсів при цьому також виростуть вдвічі.

Адитивність в цих випадках полягає в тому, що прибуток підприємства від реалізації продукції є сума вкладів від реалізації різноманітних її видів. Одночасно повна кількість кожного з використовуваних ресурсів дорівнює сумі однойменних ресурсів, які затрачуються на кожному виробництві всіх видів продукції (наприклад, загальне використання на заводі цементу за день складається з його витрат на виробництво всієї денної номенклатури виробів).

Приклад

Завод ЗБВ випускає декоративні плити двох типів – Π_1 і Π_2 , витрачаючи кольорові мінеральні наповнювачі (КМН) і кольорові мінеральні заповнювачі (КМЗ), одночасні ресурси яких обмежені; прибуток від реалізації продукції Π_1 і Π_2 різний (табл.5.1).

Таблиця 5.1

Показник	Тип плит		Ресурси матеріалу, кг
	Π_1	Π_2	
Витрата кольорових матеріалів, кг			
наповнювач КМН	70	40	168000
заповнювач КМЗ	75	100	240000
Прибуток від випуску однієї плити, грн.	3	2	—

Змінні X , тобто величини, які належить визначити в результаті розв'язання задачі ЛП, в даному випадку – об'єми виробництва плит X_1 , шт., типу Π_1 і X_2 , шт., типу Π_2 .

Прибуток від реалізації всіх плит Π_1 буде $3X_1$ грн. (виконується аксіома пропорційності), а від всіх плит Π_2 складе $2X_2$ грн., тобто загальний прибуток у відповідності з принципом адитивності:

$$C = 3X_1 + 2X_2. \quad (5.1)$$

Ця функція називається цільовою. Тобто, необхідно знайти такі значення X_1 і X_2 , щоб прибуток C був максимальним. При цьому не можна перевищувати ресурси ні по одному з матеріалів:

$$70X_1 + 40X_2 \leq 168000, \quad (5.2)$$

$$75X_1 + 100X_2 \leq 240000. \quad (5.3)$$

До отриманих рівнянь необхідно додати умову невід'ємності змінних як натуральних результатів виробництва: $X_1 \geq 0, X_2 \geq 0$.

Розв'яжемо задачу геометрично в координатах X_1 і X_2 . Область допустимих значень $\Omega\{X_1, X_2\}$ визначається на першому етапі розв'язання задачі. Внаслідок невід'ємності змінних вона знаходиться в першому квадранті і обмежена координатними осями ($X_1=X_2=0$) і двома прямими:

$75X_1 + 100X_2=240000$ (проходить через точки $M\{3200;0\}$ і $M'\{0;240000\}$) і

$70X_1 + 40X_2=168000$ (проходить через точки $K\{2400;0\}$ і $K'\{0;1680\}$), які створюють опуклий багатокутник. В кожній точці всередині нього (незаштрихована область Ω) виконуються всі обмеження задачі.

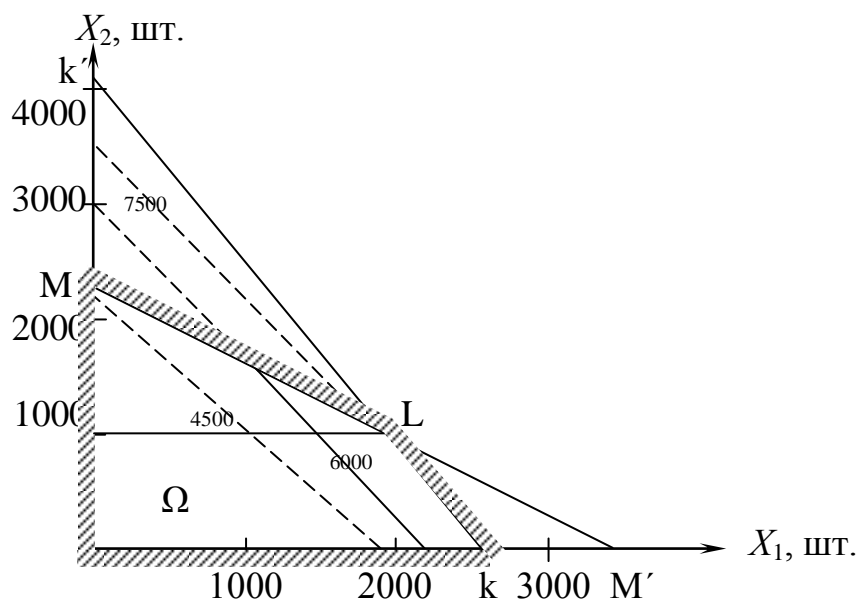


Рисунок 5.1

Наступний етап в розв'язанні задачі оснований на аналізі цільової функції C – на діаграму наносяться її ізолінії при послідовно зростаючих на $\Delta C = \text{const}$ значеннях загального прибутку. З рис.5.1 видно, що максимально можливе значення C_{\max} в допустимій області $\Omega\{X_1, X_2\}$ досягається у вершині L багатокутника – на перетині прямих. Координати її визначаються з рівнянь (5.2) і (5.3), прирівнюючи їх відносно однієї із змінних (наприклад, X_1):

$$X_1 = (168000 - 40X_2)/70;$$

$$X_1 = (24000 - 100X_2)/75;$$

$$70 \cdot (24000 - 100X_2) = 75 \cdot (168000 - 40X_2).$$

Звідки $X_2\{L\}=1050$ і $X_1\{L\}=1800$. Тоді значення цільової функції в точці максимуму $C\{L\}=3 \cdot 1800 + 2 \cdot 1050 = 7500$ грн.

Отримані результати можна перевірити за рахунок аналітичного розв'язання задачі.

Запишемо задачу в канонічній формі. Для цього введемо в ліву частину (5.2), (5.3) додаткові невід'ємні змінні \bar{X}_3 і \bar{X}_4 :

$$70X_1 + 40X_2 + \bar{X}_3 = 168000; \quad (5.4)$$

$$75X_1 + 100X_2 + \bar{X}_4 = 240000; \quad (5.5)$$

$$X_1 \geq 0, X_2 \geq 0, \bar{X}_3 \geq 0, \bar{X}_4 \geq 0.$$

Цільова функція для системи (5.4)-(5.5) аналогічна (5.1):

$$C = 3 \cdot X_1 + 2 \cdot X_2 + 0 \cdot \bar{X}_3 + 0 \cdot \bar{X}_4. \quad (5.6)$$

Отже, для визначення невідомих $X_1, X_2, \bar{X}_3, \bar{X}_4$ маємо меншу кількість рівнянь, тобто система є невизначеною.

Одне з можливих рішень невизначеної системи можна знайти, прирівнюючи які-небудь дві змінних до нуля. Наприклад, $X_1=X_2=0$, тоді $X_3=168000$ і $X_4=240000$, причому цільова функція (5.6) в цій точці $C=0$.

Розглянемо можливі варіанти розв'язання системи (5.4), (5.5):

$$X_1 = X_3 = 0 \Rightarrow X_1 = 0, X_2 = 2400, X_3 = 0, X_4 = -180000.$$

Отже, в даному випадку, отримуємо недопустимий розв'язок ($X_4 < 0$).

$$X_1 = X_4 = 0 \Rightarrow X_1 = 0, X_2 = 2400, X_3 = 72000, X_4 = 0.$$

Значення цільової функції $C=4800$.

$$X_2 = X_3 = 0 \Rightarrow X_1 = 2400, X_2 = 0, X_3 = 0, X_4 = 60000,$$

$$C = 7200,$$

$$X_2 = X_4 = 0 \Rightarrow X_1 = 3200, X_2 = 0, X_3 = -56000, X_4 = 0,$$

Знову отримали недопустимий розв'язок ($X_3 < 0$).

$$X_3 = X_4 = 0 \Rightarrow X_1 = 1800, X_2 = 1050, X_3 = 0, X_4 = 0, C = 7500.$$

Отже, отримали шість розв'язків задачі, з яких один оптимальний. Кількість розв'язків в загальному випадку можна отримати за формулою:

$$k \leq n! / [m!(n-m)!], \quad (5.7)$$

де n – кількість рівнянь;

m – кількість невідомих.

Ми впевнились, що навіть при $n=2$ число варіантів розв'язання дорівнює 6. Тобто, з точки зору геометричного розв'язання задачі прямий перебір всіх вершин багатокутника є занадто громіздким.

Систематичний обчислювальний метод, який дозволяє знаходити чисельні рішення для лінійних оптимізаційних моделей за менше число кроків, розроблений Дж. Данцігом і називається симплекс-методом [11]. Він дозволяє аналізувати і правильно інтерпретувати отримані розв'язки на кожному кроці виконання алгоритму.

Процес розв'язання починається з визначення початкової вершини багатогранника розв'язків, координати якої створюють початковий допустимий базисний розв'язок, або початковий опорний план. Оптимальний розв'язок знаходиться як результат послідовного покращення проміжного

де

$$P_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}; P_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}; \dots; P_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix};$$

$$F_{m+1} = \begin{pmatrix} a_{1m+1} \\ a_{2m+1} \\ \vdots \\ a_{mm+1} \end{pmatrix}; P_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}; P_0 = \begin{pmatrix} \vartheta_1 \\ \vartheta_2 \\ \vdots \\ \vartheta_m \end{pmatrix}.$$

Оскільки

$$\vartheta_1 P_1 + \vartheta_2 P_2 + \dots + \vartheta_m P_m = P_0, \quad (5.1.4)$$

то за означенням опорного плану $x=(\vartheta_1; \vartheta_2; \dots; 0; \dots; 0)$ є опорним планом даної задачі. План $x=(\vartheta_1; \vartheta_2; \dots; \vartheta_n)$ називається опорним планом основної задачі лінійного програмування, якщо система векторів P_j в (5.1.4) з додатними коефіцієнтами ϑ_j , лінійно незалежна. Цей план визначається системою одиничних векторів P_1, P_2, \dots, P_m , які створюють базис m -вимірного простору. Тому кожний з векторів P_1, P_2, \dots, P_n а також вектор P_0 можуть бути подані у вигляді лінійної комбінації векторів даного базису. Нехай

$$P_j = \sum_{i=1}^m x_{ij} P_i, \quad (j = \overline{0, n}).$$

Покладемо: $z_j = \sum_{i=1}^m c_i x_{ij}, \quad (j = \overline{1, n}); \Delta_j = z_j - c_j, \quad (j = \overline{1, n}).$

Оскільки вектори P_1, P_2, \dots, P_m – одиничні, то:

$$x_{ij} = a_{ij} \text{ і } z_j = \sum_{i=1}^m c_i a_{ij}, \text{ а } \Delta_j = \sum_{i=1}^m c_i a_{ij} - c_j.$$

Ознаку оптимальності опорного плану визначають за теоремами:

Теорема 1.

Опорний план $X^* = (x_1^*, x_2^*, \dots, x_m^*; 0; \dots, 0)$ задачі (5.1.1) – (5.1.3) є оптимальним, якщо $\Delta_j \geq 0$ для будь-якого $j(j = \overline{1, k})$.

Теорема 2.

Якщо $\Delta_k \leq 0$ для деякого $j = k$ і серед чисел $a_{ik}(j = \overline{1, m})$ немає додатних ($a_{ik} \leq 0$), то цільова функція (5.1.1) задачі (5.1.1) – (5.1.3) не обмежена на множині її планів.

Теорема 3.

Якщо опорний план X задачі (5.1.1) – (5.1.3) невироджений і $\Delta_k \leq 0$, але серед чисел a_{ik} є додатні (не всі $a_{ik} \leq 0$), то існує опорний план X' такий, що $F(X') \geq F(X)$.

Дослідження опорного плану на оптимальність, а також подальший обчислювальний процес зручно вести, якщо умову задачі і початкові дані отримані після визначення відповідного опорного плану в таблиці 5.2.

Таблиця 5.2

i	Базис	$C_{\bar{0}}$	P_0	C_1	C_2	...	C_r	...	C_m	C_{m+1}	...	C_k	...	C_n
				P_1	P_2	...	P_r	...	P_m	P_{m+1}	...	P_k	...	P_n
1	P_1	c_1	θ_1	1	0	...	0	...	0	a_{1m+1}	...	a_{1k}	...	a_{1n}
2	P_2	c_2	θ_2	0	1	...	0	...	0	a_{2m+1}	...	a_{2k}	...	a_{2n}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots		\vdots		\vdots	\vdots		\vdots		\vdots
r	P_r	c_r	θ_r	0	0	...	1	a_{rm+1}	...	a_{rk}	...	a_{rn}
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots		\vdots		\vdots	\vdots		\vdots		\vdots
m	P_m	c_m	θ_m	0	0	...	0	...	1	a_{mm+1}	...	a_{mk}	...	a_{mn}
$m+1$			F_0	0	0		0			Δ_{m+1}		Δ_k		Δ_n

В стовпці $C_{\bar{0}}$ цієї таблиці записують коефіцієнти при невідомих цільової функції, які мають такі ж індекси, що і вектори даного базису. В стовпці P_0 записують додатні компоненти опорного плану, в ньому ж в результаті розрахунків отримують додатні компоненти опорного плану. Стовпці векторів P_j являють собою коефіцієнти розкладання цих векторів за векторами даного базису.

В таблиці перші m рядків визначаються даними задачі, а показники $(m+1)$ -ого рядка визначають. В цьому рядку в стовпці вектора P_0 записують значення цільової функції, які вона приймає при даному опорному плані, а в стовпці P_j – значення $\Delta_j = z_j - c_j$.

Значення z_j знаходяться як скалярний добуток вектора P_j ($j = \overline{1, m}$) на вектор $C_{\bar{0}} = (c_1; c_2; \dots; c_m)$:

$$z_j = \sum_{i=1}^m c_i a_{ij}, \quad (j = \overline{1, n}).$$

Значення F_0 дорівнює скалярному добутку вектора P_0 на вектор $C_{\bar{0}}$:

$$F_0 = \sum_{i=1}^m c_i \theta_i.$$

Після заповнення таблиці опорний план перевіряють на оптимальність. Для цього переглядають елементи $(m+1)$ -ого рядка таблиці. В результаті можуть мати місце такі випадки:

1) $\Delta_j \geq 0$ для $j = m+1, m+2, \dots, n$ (при $j = \overline{1, m}$ $z_j = c_j$). Тому в даному випадку числа $\Delta_j \geq 0$ для всіх j від 1 до n ;

2) $\Delta_j \leq 0$ для деякого j , і всі відповідні цьому індексу величини $a_{ij} \leq 0$ ($i = \overline{1, m}$);

3) $\Delta_j \leq 0$ для деяких індексів j , і для кожного такого j у крайньому випадку одне з чисел a_{ij} додатне.

Перший випадок є оптимальним. В другому випадку цільова функція не обмежена зверху на множині планів, а в третьому випадку можна перейти до нового опорного плану, при якому значення цільової функції збільшиться. Цей перехід від одного опорного плану до другого відбувається вилученням з базису якого-небудь з векторів, введенням в нього нового вектора. Як новий можна взяти будь-який з векторів P_j , для якого $\Delta_j \leq 0$. Нехай, наприклад, $\Delta_k < 0$, і введемо в базис вектор P_k .

Для визначення вектора, який належить вилученню з базису, знаходять $\min(e_i/a_{ik})$ для всіх $a_{ik} > 0$. Нехай цей мінімум досягається при $i = r$. Тоді з базису вилучають вектор P_r , а число a_{rk} називають елементом розв'язку.

Стовпці, на перетині яких знаходиться елемент розв'язку, називають напрямними.

Після виділення напрямного рядка і напрямного стовпця знаходять новий опорний план, коефіцієнти розкладу векторів P_j через вектори нового базису, відповідного новому опорному плану. Це реалізується методом Жордана-Гаусса. При цьому додатні компоненти нового опорного плану обчислюються за формулами:

$$e'_i = \begin{cases} e_i - (e_r/a_{rk}) \cdot a_{ik}, & \text{при } i \neq r, \\ e_r / a_{rk}, & \text{при } i = r, \end{cases} \quad (5.1.5)$$

а коефіцієнти розкладу векторів P_j – через вектори нового базису, відповідного новому опорному плану, за формулами:

$$a'_{ij} = \begin{cases} a_{ij} - (a_{rj}/a_{rk}) \cdot a_{ik}, & \text{при } i \neq r, \\ a_{rj}/a_{rk}, & \text{при } i = r. \end{cases} \quad (5.1.6)$$

Метод Жордана-Гаусса полягає в тому, що після приведення відповідної матриці до трикутного вигляду продовжують елементарні перетворення рядків останньої до діагональної форми. В результаті за методом Жордана-Гаусса розрахунок відбувається аналогічно способу Гаусса, але розв'язок системи отримується одразу без зворотного ходу (в лівій частині буде одинична матриця).

Після знаходження e'_i і a'_{ij} згідно з формулами (5.1.5)-(5.1.6) їх значення заносять в таблицю 5.3. Елементи $(m+1)$ -го рядка цієї таблиці можуть бути визначені за формулами:

$$F'_0 = F_0 - (e_r/a_{rk}) \cdot \Delta_k, \quad (5.1.7)$$

$$\Delta'_j = \Delta_j - (a_{rj}/a_{rk}) \cdot \Delta_k. \quad (5.1.8)$$

Наявність двох стовпців знаходження елементів $(m+1)$ -ого рядка дозво-

ляє провести контроль правильності проведення розрахунків.

Таблиця 5.3

i	Базис	$C_{\bar{b}}$	P_0	C_1	C_2	...	C_r	...	C_m	C_{m+1}	...	C_k	...	C_n
				P_1	P_2	...	P_r	...	P_m	P_{m+1}	...	P_k	...	P_n
1	P_1	c_1	v_1	1	0		a'_{1r}		0	a'_{1m+1}		0		a'_{1n}
2	P_2	c_2	v_2	0	1		a'_{2r}		0	a'_{2m+1}		0		a'_{2n}
⋮	⋮	⋮	⋮	⋮	⋮		⋮		⋮	⋮		⋮		⋮
r	P_r	c_r	v_r	0	0		a'_{rr}		0	a'_{rm+1}		1		a'_{rn}
⋮	⋮	⋮	⋮	⋮	⋮		⋮		⋮	⋮		⋮		⋮
m	P_m	c_m	v_m	0	0		a'_{mr}		1	a'_{mm+1}		0		a'_{mn}
$m+1$			F_0	0	0					$z'_{m+1} - c_{m+1}$		0		$z'_n - c_n$

З формули (5.1.7) маємо, що при переході від одного опорного плану до іншого найбільш раціонально ввести в базис вектор P_j , при якому максимальним за абсолютною величиною є число $(v_r/a_{rj}) \cdot \Delta_j$ ($\Delta_j < 0, a_{rj} > 0$). В подальшому вектор, який вводиться в базис, визначати, виходячи з максимальної абсолютної величини від'ємних чисел Δ_j . Якщо таких чисел декілька, то в базис будемо вводити вектор, який має такий же індекс, як і максимальне з чисел C_j , що визначаються даними числами Δ_j ($\Delta_j < 0$).

Перехід від одного опорного плану до другого зводиться до переходу від однієї симплекс-таблиці до іншої. Елементи її вираховуються за допомогою рекурентних формул (5.1.5)-(5.1.8) за такими правилами:

Правило 1.

В стовпцях векторів, які входять в базис, на перетині рядків і стовпців однойменних векторів проставляються одиниці, а всі інші елементи даних стовпців вважають рівними нулеві.

Правило 2.

Елементи векторів P_0 і P_j в рядку нової таблиці, в якому записаний вектор, введений в базис, отримують з елементів цього ж рядка попередньої таблиці діленням їх на величину вирішуваного елемента. В стовпці $C_{\bar{b}}$ в рядку вектора, що вводиться, проставляють величину C_k , де k – індекс вирішуваного вектора.

Правило 3.

Інші елементи стовпців вектора P_0 і P_j нової таблиці обчислюють за правилом трикутника. Для знаходження будь-якого з цих елементів знаходять три числа:

1) число, яке стоїть в даній таблиці на місці шуканого елемента нової таблиці;

2) число, яке стоїть в даній симплекс-таблиці на перетині рядка, в якому знаходиться шуканий елемент нової таблиці, і стовпця відповідного вектора, що вводиться в базис;

3) число, яке стоїть в новій таблиці на перетині стовпця, в якому знаходиться шуканий елемент, і рядка знову введеного в базис вектора (цей рядок отримується із рядка даної таблиці діленням її елементів на вирішуваний елемент).

Ці три числа створюють трикутник, дві вершини якого відповідають числам, що знаходяться в даній таблиці, а третя – числу, що знаходиться в новій симплекс-таблиці. Для визначення шуканого елемента нової таблиці від першого числа віднімають добуток другого і третього.

Після заповнення нової симплекс-таблиці переглядають елементи $(m+1)$ -го рядка. Якщо всі $z_j - c_j \geq 0$, то новий опорний план є оптимальним. Якщо серед вказаних чисел є від'ємні, то, використовуючи описану вище послідовність дій, знаходять новий опорний план. Цей процес продовжують до тих пір, поки не отримають оптимальний план задачі або встановлять неможливість її розв'язання.

Як приклад розглянемо розв'язання ще однієї задачі.

Для виготовлення різноманітних виробів А, В і С підприємство використовує три види сировини. Норма витрат сировини на виготовлення одного виробу кожного виду, ціна одного виробу А, В і С, а також загальна кількість сировини кожного виду, що може бути використана підприємством, наведені в таблиці 5.4.

Таблиця 5.4

Вид сировини	Норма витрат сировини (кг) на один виріб			Загальна кількість сировини (кг)
	А	В	С	
І	18	15	12	360
ІІ	6	4	8	192
ІІІ	5	3	3	180
Ціна одного виробу (грн.)	9	10	16	

Скласти план виготовлення виробів, при якому загальна вартість всієї виробленої підприємством продукції є максимальною.

Розв'язання.

Складемо математичну модель задачі. Позначимо кількість виробів А через x_1 , виробів В – через x_2 , виробів С – через x_3 . Згідно з умовою задачі маємо таку систему нерівностей:

$$\begin{aligned} 18x_1 + 15x_2 + 12x_3 &\leq 360, \\ 6x_1 + 4x_2 + 8x_3 &\leq 192, \\ 5x_1 + 3x_2 + 3x_3 &\leq 180. \end{aligned} \quad (5.1.9)$$

Загальна вартість виготовленої підприємством продукції запишеться у вигляді:

$$F = 9x_1 + 10x_2 + 16x_3. \quad (5.1.10)$$

За своїм економічним змістом змінні x_1 , x_2 і x_3 можуть приймати тільки невід'ємні значення:

$$x_1, x_2, x_3 \geq 0. \quad (5.1.11)$$

Отже, необхідно визначити серед розв'язків (5.1.9) такі, при яких функція (5.1.10) приймає максимальне значення.

Введемо три додаткові змінні: x_4 , x_5 , x_6 – невикористана кількість сировини, відповідно. Тоді систему нерівностей (5.1.9) запишемо у вигляді системи рівнянь:

$$\begin{aligned} 18x_1 + 15x_2 + 12x_3 + x_4 &= 360, \\ 6x_1 + 4x_2 + 8x_3 + x_5 &= 192, \\ 5x_1 + 3x_2 + 3x_3 + x_6 &= 180. \end{aligned} \quad (5.1.12)$$

Систему (5.1.12) запишемо у векторній формі:

$$x_1 P_1 + x_2 P_2 + x_3 P_3 + x_4 P_4 + x_5 P_5 + x_6 P_6 = P_0,$$

де

$$P_1 = \begin{pmatrix} 18 \\ 6 \\ 5 \end{pmatrix}; \quad P_2 = \begin{pmatrix} 15 \\ 4 \\ 3 \end{pmatrix}; \quad P_3 = \begin{pmatrix} 12 \\ 8 \\ 3 \end{pmatrix}; \quad P_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; \quad P_5 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}; \quad P_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}; \quad P_0 = \begin{pmatrix} 360 \\ 192 \\ 180 \end{pmatrix}.$$

Оскільки серед векторів P_1 – P_6 є три одиничних вектори, то для даної задачі можна записати опорний план, який вважається системою тривимірних одиничних векторів P_4 , P_5 , P_6 , що створюють базис тривимірного векторного простору, тобто

$$x = (0; 0; 0; 360; 192; 180).$$

Складемо симплексну таблицю для першої ітерації, підрахуємо значення F_0 , $z_j - c_j$ і перевіримо початковий опорний план на оптимальність:

$$F_0 = (c, P_0) = 0 \cdot 360 + 0 \cdot 192 + 0 \cdot 180 = 0;$$

$$z_1 = (c, P_1) = 0 \cdot 18 + 0 \cdot 6 + 0 \cdot 5 = 0;$$

$$z_2 = (c, P_2) = 0 \cdot 15 + 0 \cdot 4 + 0 \cdot 3 = 0;$$

$$z_3 = (c, P_3) = 0 \cdot 12 + 0 \cdot 8 + 0 \cdot 3 = 0;$$

$$z_1 - c_1 = 0 - 9 = -9; \quad z_2 - c_2 = 0 - 10 = -10; \quad z_3 - c_3 = 0 - 16 = -16.$$

Для векторів базису $z_j - c_j = 0$, $j = \overline{4-6}$.

Таблиця 5.5

i	Базис	C_6	P_0	$C_1=9$	$C_2=10$	$C_3=16$	$C_4=0$	$C_5=0$	$C_6=0$
				P_1	P_2	P_3	P_4	P_5	P_6
1	P_4	0	360	18	15	12	1	0	0
2	P_5	0	192	6	4	8	0	1	0
3	P_6	0	180	5	3	3	0	0	1
4			$F_0=0$	-9	-10	-16	0	0	0

З таблиці видно, що значення цільової функції дорівнює нулеві, тобто ніщо не виробляється, сировина не використовується. Отже, план не є оптимальним.

В 4-ому рядку таблиці 5.5 від'ємні числа свідчать про можливість збільшення загальної вартості виробленої продукції і показують, на скільки збільшиться ця сума при введенні в план одиниці того чи іншого виду продукції. Так, число 9 означає, що при включенні в план виробництва одного виробу А забезпечується збільшення вартості випуску продукції на 9 грн. Якщо включити в план виробництва по одному виробу В і С, то загальна вартість виготовленої продукції виросте, відповідно, на 10 і 16 грн. Тому з економічної точки зору найбільш раціональним є включення в план виробництва вибір С. Оскільки число 16 стоїть в 4-ому рядку стовпця P_3 , то в базис вводимо вектор P_3 . Визначимо вектор, який належить виключенню з базису. Для цього знаходимо:

$$\theta_0 = \min(360/12; 192/8; 180/3) = 192/8 = 24.$$

Отже, вектор P_5 належить вилученню з базису. Стовпець вектора P_3 і другий рядок є напрямними. Складемо таблицю для 2-ї ітерації (табл.5.6).

Таблиця 5.6

i	Базис	C_6	P_0	9	10	16	0	0	0
				P_1	P_2	P_3	P_4	P_5	P_6
1	P_4	0	72	9	9	0	1	$-3/2$	0
2	P_3	16	24	$3/4$	$1/2$	1	0	$1/8$	0
3	P_6	0	108	$11/4$	$3/2$	0	0	$-3/8$	1
4			384	3	-2	0	0	2	0

Спочатку заповнюємо рядок вектора P_3 . Елементи цього рядка отримуються з відповідних елементів рядка 2 таблиці 5.3 діленням на вирішуваний елемент 8. При цьому в стовпці C_6 записуємо коефіцієнт $C_3=16$. Потім заповнюємо елементи стовпців для векторів, які входять в повний базис. В цих стовпцях на перетині рядків і стовпців однойменних векторів проставляємо одиниці, а всі інші елементи вважають рівними нулеві.

Для визначення інших елементів таблиці 5.6 використовуємо правило трикутника.

Визначаємо елементи таблиці 5.6, які стоять в стовпці вектора P_0 .

Для першого елемента маємо:

$$360 - 12 \cdot 24 = 72.$$

Для третього елемента - $180 - 24 \cdot 3 = 180$.

Значення F_0 в 4-ому рядку стовпця цього ж вектора визначимо за формулою:

$$F_0 = (C, P_0) = 0.72 + 16 \cdot 24 + 0 \cdot 108 = 384.$$

За правилом трикутника визначимо елементи стовпця вектора P_1 . Перші два числа з таблиці 5.5, а третє число – з таблиці 5.6. Це число стоїть на перетині 2-ого рядка і стовпця вектора P_1 останньої таблиці. В результаті маємо значення невідомих елементів: $18 - 12 \cdot (3/4) = 9; 5 - 3 \cdot (3/4) = 11/4$.

Число $z_1 - c_1$ в 4-ому рядку стовпця вектора P_1 таблиці 5.6 визначимо за формулою:

$$z_1 - c_1 = (C_1 P_1) - C_1 = 0 \cdot 9 + 16 \cdot 3/4 + 0 \cdot 11/4 - 9 = 3.$$

Аналогічно знаходимо елементи стовпця вектора P_2 . Елементи стовпця вектора P_3 визначаємо за правилом трикутника.

Для першого елемента маємо:

$$0 - 12 \cdot (1/8) = -3/2.$$

Для третього елемента маємо:

$$0 - 3 \cdot (1/8) = -3/8.$$

В результаті з таблиці 5.6 отримуємо новий опорний план задачі $X = (0; 0; 24; 72; 0; 108)$. При даному плані виробництва виготовляються 24 виробу C і залишаються невикористаними 72 кг сировини 1-го виду і 108 кг сировини 3-го виду. Вартість всієї виробленої продукції при цьому плані дорівнює 384 грн. Але план не є оптимальним. Це видно з 4-ого рядка таблиці 5.6, оскільки у стовпці вектора P_2 цього рядка стоїть від'ємне число 2. Отже, в базис необхідно ввести вектор P_2 , тобто в новому плані необхідно передбачити випуск виробів B . Можливий випуск виробів B визначається $\min(v'_i / a'_{i2})$ для $a'_{i2} > 0$, тобто знаходимо:

$$\theta_0 = \min\left(\frac{72}{9}; \frac{24 \cdot 2}{1}; \frac{108 \cdot 2}{3}\right) = \frac{72}{9} = 8.$$

Отже, виключенню з базису підпадає вектор P_4 . Число 9 є вирішуваним елементом, а стовпець вектора P_2 і перший рядок таблиці 5.6 є напрямними. Складаємо таблицю для третьої ітерації (таблиця 5.7).

В таблиці 5.7 спочатку заповнюємо елементи першого рядка, який являє собою рядок з вектором P_2 в базисі. Елементи цього рядка отримуємо з елементів першого рядка таблиці 5.6 діленням останніх на вирішуваний елемент (тобто на 9). При цьому в стовпці C_6 даного рядка записуємо $C_2 = 10$.

Потім заповнюємо елементи стовпців векторів базису і за правилом трикутника вираховуємо елементи останніх стовпців.

Так, для стовпця вектора P_0 маємо:

другий елемент: $24 - 8 \cdot 0.5 = 20$;

третьій елемент: $108 - 3/2 \cdot 8 = 96$.

Значення F_0 в 4-ому рядку стовпця цього ж вектора визначимо за формулою:

$$F_0 = (C, P_0) = 8 \cdot 10 + 16 \cdot 20 - 0 \cdot 96 = 400.$$

Таблиця 5.7

i	Базис	C_b	P_0	9	10	16	0	0	0
				P_1	P_2	P_3	P_4	P_5	P_6
1	P_2	10	8	1	1	0	1/9	-1/6	0
2	P_3	16	20	1/4	0	1	-1/18	5/24	0
3	P_6	0	96	5/4	0	0	-1/6	-1/8	1
4			400	5	0	0	2/9	5/3	0

Для стовпця вектора P_1 визначимо елементи 2-4 рядків.

$$\text{Другий елемент: } \frac{3}{4} - \frac{1}{2} \cdot 1 = \frac{1}{4}.$$

$$\text{Третій елемент: } \frac{11}{4} - \frac{3}{2} \cdot 1 = \frac{5}{4}.$$

Аналогічно вираховуємо елементи для векторів P_4 і P_5 (елементи другого і третього рядків).

Наприклад, для вектора P_5 :
елемент другого рядка:

$$\frac{1}{8} - \frac{1}{2} \cdot \left(-\frac{1}{6}\right) = \frac{1}{8} + \frac{1}{12} = \frac{5}{24};$$

елемент третього рядка:

$$-\frac{3}{8} - \frac{3}{2} \cdot \left(-\frac{1}{6}\right) = -\frac{1}{8}.$$

Розглянемо четвертий рядок таблиці 5.7. В цьому рядку серед чисел Δ_j немає від'ємних. Це означає, що знайдений опорний план є оптимальним і $F_{\max} = 400$.

Отже, план випуску продукції, включаючи виготовлення 8 виробів В і 20 виробів С, є оптимальним. Сировина першого і другого видів використовується повністю. При цьому залишається 96 кг сировини третього виду, а вартість виробленої продукції рівна 400грн.

5.2 Транспортна задача

Математична постановка задачі.

Загальна постановка транспортної задачі полягає у визначенні оптимального плану перевезення деякого однорідного вантажу з m пунктів відправлення A_1, A_2, \dots, A_m у n пунктів призначення B_1, B_2, \dots, B_n . При цьому як критерій оптимальності звичайно береться або мінімальна вартість перевезень усього вантажу, або мінімальний час його доставки. Розглянемо транспортну задачу, за критерій оптимальності якої взята мінімальна вартість перевезень усього вантажу. Позначимо через $C_{i,j}$ тарифи перевезення одиниці вантажу з i -го пункту відправлення в j -й пункт призначення, через a_j – запаси вантажу в j -му пункті відправлення, через b_j – потреби у вантажі в j -ому пункті призначення, а через x_{ij} – кількість одиниць вантажу, перевезеного з i -ого пункту відправлення в j -ий пункт призначення. Тоді математична постановка задачі полягає у визначенні мінімального значення функції

$$F = \sum_{i=1}^m \sum_{j=1}^n c_{ij} x_{ij} \quad (5.2.1)$$

за умов

$$\sum_{i=1}^m x_{ij} = b_j \quad (j = \overline{1, n}); \quad (5.2.2)$$

$$\sum_{i=1}^m x_{ij} = a_j \quad (j = \overline{1, m}); \quad (5.2.3)$$

$$x_{ij} \geq 0 \quad (i = \overline{1, m}; j = \overline{1, n}). \quad (5.2.4)$$

Оскільки змінні x_{ij} ($i = \overline{1, m}; j = \overline{1, n}$) задовольняють системи лінійних рівнянь (5.2.2) і (5.2.3) і умову незаперечності (5.2.4), забезпечується доставка необхідної кількості вантажу в кожний з пунктів призначення, вивезення наявного вантажу з усіх пунктів відправлення, а також включаються зворотні перевезення.

Означення 1. Будь-який негативний розв'язок систем лінійних рівнянь (5.2.2) і (5.2.3), обумовлений матрицею $X = (x_{ij}) (i = \overline{1, m}; j = \overline{1, n})$, називається *планом* транспортної задачі.

Означення 2. План $X^* = (x^*_{ij}) (i = \overline{1, m}; j = \overline{1, n})$, при якому функція (5.2.1) приймає своє мінімальне значення, називається *оптимальним планом* транспортної задачі.

Звичайно вихідні дані транспортної задачі записують у вигляді табл.5.8.

Очевидно, загальна наявність вантажу в постачальників дорівнює $\sum_{i=1}^m a_i$,

а загальна потреба у вантажі в пунктах призначення дорівнює $\sum_{i=1}^n b_i$ оди-

ниць. Якщо загальна потреба у вантажі в пунктах призначення дорівнює запасу вантажу в пунктах відправлення, тобто

$$\sum_{i=1}^m a_i = \sum_{j=1}^n b_j, \quad (5.2.5)$$

то модель такої транспортної задачі називається *закритою*. Якщо ж зазначена умова не виконується, то модель транспортної задачі називається *відкритою*.

Таблиця 5.8.

Пункти відправлення	Пункти призначення					Запаси
	B_1	...	B_j	...	B_n	
A_1	c_{11} X_{11}	...	c_{1j} X_{1j}	...	c_{1n} X_{1n}	a_1
...
A_i	c_{i1} X_{i1}	...	c_{ij} X_{ij}	...	c_{in} X_{in}	a_i

A_m	c_{m1} X_{m1}	...	c_{mj} X_{mj}	...	c_{mn} X_{mn}	a_m
Потреби	b_1	...	b_j	...	b_n	

Теорема 1. Для розв'язання транспортної задачі необхідно і достатньо, щоб запаси вантажу в пунктах відправлення були рівні потребам у вантажі в пунктах призначення, тобто, щоб виконувалась рівність (5.2.5).

У випадку перевищення запасу над потребою, тобто $\sum_{i=1}^m a_i > \sum_{j=1}^n b_j$, вводиться фіктивний $(n+1)$ -ий пункт призначення з потребою $b_{n+1} = \sum_{i=1}^m a_i - \sum_{j=1}^n b_j$ і відповідні тарифи вважаються рівними нулю: $c_{in+1} = 0 (i = \overline{1, m})$. Отримана задача є транспортною задачею, для якої виконується рівність (5.2.5).

Аналогічно, при $\sum_{i=1}^m a_i < \sum_{j=1}^n b_j$ вводиться фіктивний $(m+1)$ -ий пункт відправлення з запасом вантажу $a_{m+1} = \sum_{j=1}^n b_j - \sum_{i=1}^m a_i$ і тарифи покладаються рівними нулю: $c_{m+1,j} = 0 (j = \overline{1, n})$. Цим задача зводиться до звичайної транспортної задачі, з оптимального плану якої виходить оптимальний план вихідної задачі. Надалі будемо розглядати закрити модель транспортної задачі.

Якщо ж модель конкретної задачі є відкритою, то, виходячи із сказаного вище, перепишемо таблицю умов задачі так, щоб виконувалась рівність (5.2.5).

Число змінних x_{ij} у транспортній задачі з m пунктами відправлення і n пунктами призначення дорівнює nm , а число рівнянь у системах (5.2.2) і (5.2.3) дорівнює $n+m$. Оскільки ми припускаємо, що виконується умова (5.2.5), то число лінійно незалежних рівнянь дорівнює $n+m-1$. Отже, опорний план транспортної задачі може мати не більше $n+m-1$ відмінних від нуля невідомих.

Якщо в опорному плані число відмінних від нуля компонентів дорівнює в точності $n+m-1$, то план є невикористаним, а якщо менше – то виродженим.

Для визначення опорного плану існує кілька методів. Три з них – метод північно-західного кута, метод мінімального елемента і метод апроксимації Фогеля – розглядаються нижче.

Як і для всякої задачі лінійного програмування, оптимальний план транспортної задачі є й опорним планом.

Для визначення оптимального плану транспортної задачі можна використовувати викладені вище методи. Однак через виняткову практичну важливість цієї задачі і специфіку її обмежень (кожна невідома входить лише в два рівняння систем (5.2.2) і (5.2.3) і коефіцієнти при невідомих дорівнюють одиниці) для визначення оптимального плану транспортної задачі розроблені спеціальні методи. Один з них – метод потенціалів – розглянутий нижче.

Задача 1.

Чотири підприємства даного економічного району для виробництва продукції використовують три види сировини. Потреби в сировині кожного з підприємств відповідно рівні 120, 50, 190 і 110 од. Сировина зосереджена в трьох місцях, а запаси відповідно рівні 160, 140, 170 од. На кожне з підприємств сировина може завозитися з будь-якого пункту її отримання. Тарифи перевезень є відомими величинами і задаються матрицею:

$$C = \begin{vmatrix} 7 & 8 & 1 & 2 \\ 4 & 5 & 9 & 8 \\ 9 & 2 & 3 & 6 \end{vmatrix}.$$

Скласти такий план перевезень, при якому загальна вартість перевезень є мінімальною.

Розв'язання.

Позначимо через x_{ij} кількість одиниць сировини, перевезеної з 1-ого пункту її отримання на j -е підприємство. Тоді умови доставки і вивезення необхідної і наявної сировини забезпечуються за рахунок виконання таких рівностей:

$$\begin{aligned}
x_{11} + x_{12} + x_{13} + x_{14} &= 160, \\
x_{21} + x_{22} + x_{23} + x_{24} &= 140, \\
x_{31} + x_{32} + x_{33} + x_{34} &= 170, \\
x_{11} + x_{21} + x_{31} &= 120, \\
x_{12} + x_{22} + x_{32} &= 50, \\
x_{13} + x_{23} + x_{33} &= 190.
\end{aligned}
\tag{5.2.6}$$

При даному плані $X=(x_{ij})(i=1,3; j=1,4)$ перевезень загальна її вартість складе:

$$F=7x_{11}+8x_{12}+x_{13}+2x_{14}+4x_{21}+5x_{22}+9x_{23}+8x_{24}+9x_{31}+2x_{32}+3x_{33}+6x_{34}. \tag{5.2.7}$$

Таким чином, математична постановка даної транспортної задачі полягає в знаходженні такого невід'ємного розв'язку системи лінійних рівнянь (5.2.6), при якому цільова функція (5.2.7) приймає мінімальне значення.

5.2.1 Метод північно-західного кута

При знаходженні опорного плану транспортної задачі методом північно-західного кута на кожному кроці розглядаємо перший з пунктів відправлення, що залишилися, та перший з пунктів призначення, що залишилися. Заповнення клітинок таблиці умови починається з лівої верхньої клітинки для невідомого x_{11} ("північно-західний кут") і закінчується клітинкою для невідомого x_{mn} , тобто наче йде по діагоналі таблиці.

Приклад

На три бази A_1, A_2, A_3 надійшов однорідний вантаж в кількостях, відповідно рівних 140, 180 і 160 одиниць. Цей вантаж треба перевезти в п'ять пунктів призначення B_1, B_2, B_3, B_4, B_5 відповідно в кількостях 60, 70, 120, 130 і 100 одиниць. Тарифи перевезення одиниць вантажу з кожного з пунктів відправлення у відповідні пункти призначення вказані в таблиці:

Таблиця 5.9

Пункти відправлення	Пункти призначення					Запас
	B_1	B_2	B_3	B_4	B_5	
A_1	2	3	4	2	4	140
A_2	8	4	1	4	1	180
A_3	9	7	3	7	2	160
Потреби	60	70	120	130	100	480

Знайти план перевезень даної транспортної задачі методом північно-західного кута.

Розв'язання.

Тут кількість пунктів відправлення $m = 3$, а кількість пунктів призначення $n = 5$. Отже, опорний план задачі визначається числами, які знаходяться в $5+3-1=7$ заповнених клітинках.

Заповнення таблиці почнемо з клітинки для невідомого x_{11} , тобто будемо намагатися задовольнити потреби першого пункту призначення за рахунок запасів першого пункту відправлення. Оскільки запаси пункту A_1 більші, ніж потреби пункту B_1 то вважаємо $x_{11} = 60$, записуємо це значення у відповідній клітинці таблиці 5.11 і тимчасово виключаємо з розгляду стовпець B_1 , вважаючи при цьому запаси пункту A_1 рівними 80.

Таблиця 5.10

Пункти відправлення	Пункти призначення					Запас
	B_1	B_2	B_3	B_4	B_5	
A_1	² 60	³ 70	⁴ 10	²	⁴	140
A_2	⁸	⁴	¹ 110	⁴ 70	¹	180
A_3	⁹	⁷	³	⁷ 60	² 100	160
Потреби	60	70	120	130	100	480

Розглянемо перший з пунктів відправлення A_1 та пункт призначення B_2 . Запаси пункту A_1 більше потреб пункту B_2 . Покладемо $x_{12} = 70$, запишемо це значення у відповідній клітинці таблиці 5.10 і тимчасово виключимо з розгляду стовпець B_2 . В пункті A_1 запаси вважаємо рівними 10 одиницям. Знову розглянемо перші з пунктів відправлення A_1 та призначення B_3 , що залишилися. Потреби пункту B_3 більші запасів, що залишилися в пункті A_1 . Покладемо $x_{13} = 10$ та виключимо з розгляду стовпець A_1 . Значення $x_{13} = 10$ запишемо у відповідну клітинку таблиці 5.10 і будемо вважати потреби пункту B_3 рівними 110 одиницям.

Тепер перейдемо до заповнення клітинок для невідомого x_{23} і т.д. Через шість кроків залишається один пункт відправлення A_3 з запасом вантажу 100 одиниць і один пункт призначення B_5 з потребою в 100 одиниць. Відповідно є одна вільна клітинка, яку і заповнюємо, вважаючи $x_{35} = 100$ (таблиця 5.10). В результаті отримуємо опорний план:

$$X = \begin{pmatrix} 60 & 70 & 10 & 0 & 0 \\ 0 & 0 & 110 & 70 & 0 \\ 0 & 0 & 0 & 60 & 100 \end{pmatrix}.$$

Згідно з даним планом перевезень, загальна вартість перевезень всього вантажу складає:

$$F = 2 \cdot 60 + 3 \cdot 70 + 4 \cdot 10 + 1 \cdot 110 + 4 \cdot 70 + 7 \cdot 60 + 2 \cdot 100 = 1380.$$

5.2.2 Метод мінімального елемента

Суть методу мінімального елемента полягає у виборі клітинки з мінімальним тарифом. Слід відмітити, що цей метод, як правило, дозволяє знайти опорний план транспортної задачі, при якому загальна вартість перевезень вантажу менша, ніж загальна вартість перевезень при план, знайденому для даної задачі за допомогою методу північно-західного кута. Тому найбільш доцільно опорний план транспортної задачі знаходити методом мінімального елемента.

Приклад.

Вхідні дані задачі запишемо у вигляді таблиці 5.11.

Таблиця 5.11

Пункти відправлення	Пункти призначення					Запас
	B ₁	B ₂	B ₃	B ₄	B ₅	
A ₁	2	3	4	2	4	140
A ₂	8	4	1	4	1	180
A ₃	9	7	3	7	2	160
Потреби	60	70	120	130	100	480

Покладемо $x_{23} = 120$ і запишемо це значення у відповідну клітинку таблиці 5.12. Виключимо тимчасово з розгляду стовпець B₃.

В частині таблиці, що залишилася з трьома рядками і з чотирма стовпцями B₁, B₂, B₄ і B₅, клітинка з найменшим значенням тарифу C₂₅ знаходиться на перетині рядка A₂ і стовпця B₅. Покладемо $x_{25} = 60$ і внесемо це значення у відповідну клітинку таблиці 5.12.

Таблиця 5.12

Пункти відправлення	Пункти призначення					Запас
	B ₁	B ₂	B ₃	B ₄	B ₅	
A ₁	2 10	3	4	2 130	4	140
A ₂	8	4	1 120	4	1 60	180
A ₃	9 20	7 90	3	7	2 40	160
Потреби	60	70	120	130	100	480

Тимчасово виключимо з розгляду рядок A₂. Оскільки тариф 2 зустрічається у клітинках тричі, то доцільно вибрати тариф C₁₄. Покладемо $x_{14} =$

130, стовпець B_4 виключаємо з розгляду. Після цього з врахуванням мінімальних тарифів заповнюємо стовпець B_5 : $x_{35} = 40$ і стовпець B_1 , $x_{11} = 10$.

Знову розглядаємо частину таблиці, що залишилася. В ній мінімальний тариф C_{ij} знаходиться в клітинці на перетині рядка A_3 і стовпця B_2 , що дорівнює 7.

Заповнимо наведеним вище способом цю клітинку і аналогічно заповнюємо (у відповідній послідовності) клітинку, що знаходиться на перетині рядка A_3 і стовпця B_1 . В результаті отримаємо опорний план:

$$X = \begin{pmatrix} 10 & 0 & 0 & 130 & 0 \\ 0 & 0 & 120 & 0 & 60 \\ 20 & 90 & 0 & 0 & 40 \end{pmatrix}.$$

При даному плані перевезень загальна вартість транспортування складає:

$$F = 2 \cdot 10 + 9 \cdot 20 + 7 \cdot 90 + 1 \cdot 120 + 2 \cdot 130 + 1 \cdot 60 + 2 \cdot 40 = 1170.$$

5.2.3 Метод апроксимації Фогеля

При визначенні оптимального плану транспортної задачі методом апроксимації Фогеля на кожній ітерації по всіх стовпцях і по всіх рядках знаходять різницю між двома записаними в них мінімальними тарифами. Ці різниці записують в спеціально відведених для цього рядку та стовпці в таблиці умови задачі. В рядку (чи в стовпці), якому дана різниця відповідає, визначають мінімальний тариф. Клітинку, в якій він записаний, заповнюють на даній ітерації.

Якщо мінімальний тариф однаковий для декількох клітинок даного рядка (стовпця), то для заповнення вибирають ту клітинку, яка знаходиться в стовпці (рядку), що відповідає найбільшій різниці між двома мінімальними тарифами, які знаходяться в даному стовпці (рядку).

Приклад.

Використовуючи метод апроксимації Фогеля, знайдемо опорний план попередньої задачі.

Таблиця 5.13

Пункти відправлення	Пункти призначення					Запас
	B_1	B_2	B_3	B_4	B_5	
A_1	2	3	4	2	4	140
A_2	8	4	1	4	1	180
A_3	9	7	3	7	2	160
Потреби	60	70	120	130	100	480

Розв'язання.

Для кожного рядка і стовпця таблиці умови знайдемо різницю між двома мінімальними тарифами, записаними в даному рядку або стовпці та помістимо їх у відповідному додатковому стовпці або в додатковому рядку таблиці. Так, в рядку A_1 мінімальний тариф, який дорівнює 2, зустрічається двічі, тому різниця дорівнює 0.

Аналогічна ситуація і в рядку A_2 з мінімальними тарифами 1. В рядку A_3 мінімальні тарифи 3 і 2. Отже, різниця між ними: $3-2=1$. Аналогічно отримуємо різницю в стовпці B_1 : $8-2=6$. Вирахувавши всі ці різниці, бачимо, що найбільша з них відповідає стовпцю B_1 . Таким чином, необхідно заповнити клітинку, що знаходиться на перетині рядка A_1 та B_1 , оскільки в ній найменша ціна. Заповнивши її, ми тим самим задовольнили потреби пункту B_1 , тому виключаємо з розгляду стовпець B_1 і будемо вважати запаси пункту A_1 рівними $140-60=80$ одиниць. Після цього визначимо наступну клітинку для заповнення. Знову знаходимо різниці між мінімальними тарифами в кожному з рядків і стовпців, що залишились. Запишемо їх, відповідно, в другому додатковому стовпці і рядку таблиці 5.14. Як видно з цієї таблиці, найбільша різниця (в даному випадку однакова) буде в стовпцях B_3 і B_4 . Вибираємо той стовпець, в якому найменший мінімальний тариф. Це буде стовпець B_3 . На перетині цього стовпця з рядком A_2 визначаємо клітинку з мінімальним тарифом і заповнюємо її з врахуванням потреб пункту B_3 та запасів пункту A_2 , виключаємо з розгляду стовпець B_3 і будемо вважати запаси пункту A_2 рівними $180-120=60$ одиниць.

Продовжимо ітераційний процес, послідовно заповнюючи клітинки, що знаходяться на перетині рядка A_3 і стовпця B_5 , рядка A_1 та стовпця B_4 , рядка A_2 і стовпця B_2 , рядка A_3 і стовпця B_4 , рядка A_3 і стовпця B_2 .

Таблиця 5.14

Пункти відправлення	Пункти призначення					Запаси	Різниці по рядках				
	B_1	B_2	B_3	B_4	B_5						
A_1	² 60	³	⁴	² 80	⁴	140	0	1	1	1	0
A_2	⁸	⁴ 60	¹ 120	⁴	¹	180	0	0	3	0	0
A_3	⁹	⁷ 10	³	⁷ 50	² 100	160	1	1	⁵⁽³⁾	0	0
Потреби	60	70	120	130	100	480					
Різниці по стовпцях	⁶⁽¹⁾	1	2	2	1						
	-	1	²⁽²⁾	2	1						
	-	1	-	2	1						
	-	1	-	²⁽⁴⁾	-						
	-	³⁽⁵⁾	-	3	-						

В результаті отримуємо опорний план:

$$X = \begin{pmatrix} 60 & 0 & 0 & 80 & 0 \\ 0 & 60 & 120 & 0 & 0 \\ 0 & 10 & 0 & 50 & 100 \end{pmatrix}.$$

При цьому плані загальна вартість перевезення така:

$$F = 2 \cdot 60 + 4 \cdot 60 + 7 \cdot 10 + 1 \cdot 120 + 2 \cdot 80 + 7 \cdot 50 + 2 \cdot 100 = 1260.$$

Як правило, використання методу апроксимації Фогеля позволяв отримати опорний план, близький до оптимального, або сам оптимальний план.

5.2.4 Розв'язання транспортної задачі методом потенціалів

Для визначення оптимального плану транспортної задачі розроблено декілька методів. Але найбільш часто використовуються метод потенціалів і метод диференціальних рент.

Загальний принцип визначення оптимального плану транспортної задачі методом потенціалів аналогічний принципу розв'язання задачі лінійного програмування симплексним методом, тобто: спочатку знаходять опорний план транспортної задачі, а потім його послідовно покращують до отримання оптимального плану. Планом транспортної задачі буде матриця

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{pmatrix},$$

в якій спостерігається баланс по рядках і стовпцях. Це відповідає вимогам обмежень задачі. Опорний план транспортної задачі буде містити $(m+n-1)$ базисний компонент. Система обмежень складається з $(m+n)$ обмежень – рівнянь, однак незалежних рівнянь $(m+n-1)$ (вимога $\sum_{i=1}^m a_i = \sum_{j=1}^n b_j$ робить систему залежною).

Якщо число базисних компонент менше, ніж $(m+n-1)$, то опорний план вироджений, тоді деякі нульові компоненти необхідно вважати базисними. Їх кількість така, що базисних усього буде $(m+n-1)$.

Опорний план може бути визначений різними способами. Наприклад, метод північно-західного кута, метод найменших елементів, метод Фогеля. Алгоритмічно визначити опорний план – це значить знайти місця розташування додатних компонентів і їх величини таким чином, щоб був баланс по рядках і стовпцях (умова плану) і їх кількість була $(m+n-1)$ (умова опорності плану).

Метод потенціалів заснований на ідеї знаходження потенціалів (величин α_i і β_j) і виборі компоненти, що вводиться в базис, якщо опорний план не оптимальний.

Алгоритм методу потенціалів являє собою покроковий процес, що скла-

дається з перевірки на оптимальність опорного плану, виродженості і переходу на повний опорний план. Процес керується ознакою оптимальності. Опишемо цей процес.

Крок 1. Звести дані в таблицю і визначити опорний план. Якщо план вироджений, то ввести нульові базисні елементи.

Крок 2. Скласти систему рівнянь для $x_{ij} \neq 0$ і обчислити всі потенціали, відповідно, пунктів відправки і пунктів призначення α_i і β_j . Ці числа знаходяться з системи рівнянь

$$\beta_j - \alpha_i = c_{ij}, \quad (5.2.6)$$

де c_{ij} – тарифи, що стоять в заповнених клітинках таблиці умов транспортної задачі.

Оскільки кількість заповнених клітинок $n+m-1$, то система (5.2.6) з $n+m$ невідомими містить $n+m-1$ рівнянь. Оскільки кількість невідомих перевищує на одиницю кількість рівнянь, то одне з невідомих можна покласти рівним довільному числу, наприклад, $\alpha_i = 0$, і знайти послідовно з рівнянь (5.2.6.) значення інших невідомих. Після того, як всі потенціали визначені, для кожної вільної клітинки визначають числа

$$\alpha_{ij} = \beta_j - \alpha_i - c_{ij}.$$

Якщо серед чисел α_{ij} немає додатних, то знайдений опорний план є оптимальним. В іншому випадку необхідно перейти до нового опорного плану. Для цього розглядають всі вільні клітинки, для яких $\alpha_{ij} > 0$, і серед даних чисел вибирають максимальне. Клітинку, якій це число відповідає, необхідно заповнити.

Заповнюючи вибрану клітинку, необхідно змінити об'єми поставок, записаних в рядку інших зайнятих клітинок.

Процедура переходу від одного опорного плану до іншого виконується таким чином. Нехай вибраний деякий небазисний елемент x_{ij} , який вводиться в базис. Відмітимо його знаком "+". Далі будується замкнений ланцюжок, вершинами якого будуть базисні елементи опорного плану. Вершини ланцюжка позначаються знаком "+" і "-" по черзі, дотримуючи правило: кожний рядок і стовпець повинні мати стільки елементів, відмічених знаком "-", скільки "+". Серед елементів, відмічених знаком "-", вибирається мінімальний елемент. Позначимо його через θ . θ додається до елементів, відмічених знаком "+", і віднімається від елементів, відмічених знаком "-".

Крок 3. Після цієї процедури отриманий новий опорний план перевіряємо на оптимальність, тобто знову повторюємо всі дії кроку 2.

Зауважимо, що в процесі розв'язання задачі може бути отриманий вироджений опорний план. Щоб уникнути в цьому випадку зациклювання, необхідно відповідні нульові елементи опорного плану замінити малим додатним числом ϵ і розв'язувати задачу як не вироджену. В оптимальному плані такої задачі необхідно вважати ϵ рівним нулеві.

Визначимо оптимальний план задачі 1 методом потенціалів.

В першу чергу, використовуючи метод мінімального елемента, знаходимо опорний план задачі. Цей план записаний в таблицю 5.15.

Таблиця 5.15

Пункти відправлення	Пункти призначення				Запас
	B ₁	B ₂	B ₃	B ₄	
A ₁	⁷	⁸	¹	²	160
A ₂	⁴	⁵	⁹	⁸	140
A ₃	⁹	²	³	⁶	170
Потреби	120	50	190	110	470

Складемо систему рівнянь:

$$\alpha_1 + \beta_3 = 1;$$

$$\alpha_2 + \beta_1 = 4;$$

$$\alpha_2 + \beta_4 = 8;$$

$$\alpha_3 + \beta_2 = 2;$$

$$\alpha_3 + \beta_3 = 3;$$

$$\alpha_3 + \beta_4 = 6.$$

Нехай $\alpha_1 = 0$. Тоді $\beta_3 = 1$, $\alpha_3 = 0$, $\beta_4 = 6$, $\alpha_2 = 2$, $\beta_1 = 2$, $\beta_2 = 2$.

Тепер для кожної вільної клітинки визначимо числа $\alpha_{ij} = \beta_j - \alpha_i - c_{ij}$:

$$\alpha_{11} = \beta_1 - \alpha_1 - c_{11} = 2 - 0 - 7 = -5;$$

$$\alpha_{12} = \beta_2 - \alpha_1 - c_{12} = 2 - 0 - 8 = -6;$$

$$\alpha_{14} = \beta_4 - \alpha_1 - c_{14} = 6 - 0 - 2 = 4;$$

$$\alpha_{22} = \beta_2 - \alpha_2 - c_{22} = 2 - 2 - 5 = -5;$$

$$\alpha_{23} = \beta_3 - \alpha_2 - c_{23} = 1 - 2 - 9 = -10;$$

$$\alpha_{31} = \beta_1 - \alpha_3 - c_{31} = 2 - 0 - 9 = -7.$$

Поміщаємо знайдені числа в рамки і записуємо їх відповідно в кожну вільну клітинку таблиці 5.16.

Таблиця 5.16

Пункти відправлення	Пункти призначення				Запас
	B_1	B_2	B_3	B_4	
A_1	-5	-6	160		160
A_2	120	-5	-10	20	140
A_3	-7	50	30	90	170
Потреби	120	50	190	110	470

План неоптимальний, оскільки $\alpha_{14} = 4 > 0$. Потрібно компонент x_{14} ввести в базис. Побудуємо ланцюжок.

		-160	+
120			20
	50	+30	-90

Найменше з чисел в мінусових клітинках дорівнює 90. Отже, клітинка, в якій знаходиться це число, стає вільною в новій таблиці 5.17. Інші числа в табл.5.1 отримуються так: до числа 30, що стоїть в плюсовій клітинці табл.5.16. додамо 90 і віднімемо 90 від числа 160, що знаходиться в мінусовій клітинці табл. 5.16. Клітинка на перетині рядка A_3 і стовпця B_4 стає вільною.

Після цих перетворень отримуємо новий опорний план (табл. 5.17) .

Таблиця 5.17

Пункти відправлення	Пункти призначення				Запас
	B_1	B_2	B_3	B_4	
A_1	-2	-19	70		160
A_2	120	-16	-16	20	140
A_3	-8	50	120	-8	170
Потреби	120	50	190	110	470

Перевіримо цей план на оптимальність. Знову знаходимо потенціали пунктів відправки і призначення. Для цього складемо таку систему рівнянь:

$$\alpha_1 + \beta_3 = 1;$$

$$\alpha_1 + \beta_4 = 2;$$

$$\alpha_2 + \beta_2 = 5;$$

$$\alpha_2 + \beta_3 = 9;$$

$$\alpha_3 + \beta_1 = 9;$$

$$\alpha_3 + \beta_4 = 6.$$

Покладемо $\alpha_1 = 0$. Отримаємо: $\beta_3 = 1$, $\beta_4 = 2$, $\alpha_3 = 6 - 2 = 4$,
 $\beta_1 = 9 - 4 = 5$, $\alpha_2 = 8$, $\beta_2 = -3$.

Для кожної вільної клітинки обчислимо числа α_{ij} .

$$\alpha_{11} = \beta_1 - \alpha_1 - c_{11} = 5 - 0 - 7 = -2;$$

$$\alpha_{12} = \beta_2 - \alpha_1 - c_{12} = -3 - 8 - 8 = -19;$$

$$\alpha_{22} = \beta_2 - \alpha_2 - c_{22} = -3 - 8 - 5 = -16;$$

$$\alpha_{23} = \beta_3 - \alpha_2 - c_{23} = 1 - 8 - 9 = -16;$$

$$\alpha_{31} = \beta_1 - \alpha_3 - c_{31} = 5 - 4 - 9 = -8;$$

$$\alpha_{34} = \beta_4 - \alpha_3 - c_{34} = 2 - 4 - 6 = -8.$$

В даному випадку всі $\alpha_{ij} < 0$ і опорний план є оптимальний. Сумарні транспортні витрати:

$$S = 1 \cdot 70 + 2 \cdot 90 + 4 \cdot 120 + 8 \cdot 20 + 2 \cdot 50 + 3 \cdot 120 = 1350.$$

Отже, отримуємо такий оптимальний план:

$$X^* = \begin{pmatrix} 0 & 0 & 70 & 90 \\ 120 & 0 & 0 & 20 \\ 0 & 50 & 120 & 0 \end{pmatrix}.$$

VI ГРАДІЄНТНИЙ МЕТОД РОЗВ'ЯЗУВАННЯ ЗАДАЧ НЕЛІНІЙНОГО ПРОГРАМУВАННЯ

В загальному випадку задача нелінійного програмування полягає у визначенні максимального (мінімального) значення функції

$$f(x_1, x_2, \dots, x_n) = f(x) \quad (6.1)$$

за умови, що її змінні задовольняють співвідношення:

$$\begin{cases} g_i(x_1, x_2, \dots, x_n) \leq \vartheta_i & (i = \overline{1, k}), \\ g_i(x_1, x_2, \dots, x_n) = \vartheta_n & (i = \overline{k+1, m}), \end{cases} \quad (6.2)$$

де f і g_i – деякі відомі функції n змінних, а ϑ_i – задані числа.

В результаті розв'язання задачі буде визначена точка $x^* = (x_1^*; x_2^*; \dots; x_n^*)$, координати якої задовольняють (6.2) і така, що для будь-якої іншої точки $x = (x_1, x_2, \dots, x_n)$, що задовольняє умови (6.2), виконується нерівність:

$$f(x_1^*; x_2^*; \dots; x_n^*) \geq f(x_1, x_2, \dots, x_n), \left[f(x_1^*; x_2^*; \dots; x_n^*) \leq f(x_1, x_2, \dots, x_n) \right]$$

Якщо f і g_i – лінійні функції, то задача (6.1), (6.2) є задачею лінійного програмування.

Співвідношення (6.2) утворюють систему обмежень і містять в собі умови невід’ємності змінних, якщо такі умови існують. Умови додатності змінних можуть бути задані окремо.

Для розв’язання вищенаведеної задачі не існує універсальних методів. Але для окремих класів задач, в яких зроблені додаткові обмеження відносно властивостей функцій f і g_i , розроблені ефективні методи їх розв’язання. Так, наприклад, за умови, що f – опукла функція, то область допустимих розв’язків $x_i \geq 0$, що визначається обмеженнями (6.2), – опукла.

Для кращого розуміння вищесказаного наведемо деякі означення. Нехай E – (лінійний простір) дійсна вісь, або комплексна площина. Відрізком, що з’єднує точки $x_1, x_2 \in E$, називається сукупність всіх точок x вигляду:

$$x = (1-t)x_1 + tx_2, \quad t \in [0,1].$$

Множина W в лінійному просторі E називається опуклою, якщо кожний раз з того, що $x_1, x_2 \in W$, випливає: W належить відрізок, який з’єднує x_1 і x_2 .

Введемо тепер поняття опуклої функції (функціонала). Нехай на лінійному просторі E задана функція, що ставить кожному $x \in E$ у відповідність $f(x)$. В цьому випадку говорять, що на E заданий функціонал $f(x)$. Якщо всі значення $f(x)$ дійсні, то функціонал $f(x)$ називається дійсним функціоналом. Дійсний функціонал $f(x)$ називається опуклим на множині $x \in E$, якщо для будь-яких точок x_1 і x_2 із x і $t \in [0,1]$:

$$f[(1-t) \cdot x_1 + tx_2] \leq (1-t)f(x_1) + tf(x_2). \quad (6.3)$$

Якщо нерівність (6.3) виконується навпаки, то функціонал $f(x)$ називається вгнутим.

Задача (6.1), (6.2) називається задачею опуклого програмування, якщо функція $f(x)$ є вгнутою, а функції $g_i(x)$ – опуклі.

Використання градієнтних методів можливе для кожної задачі нелінійного програмування. Але в загальному випадку використання цих методів дозволяє знайти точку локального екстремуму. Тому більш доцільно використовувати градієнтні методи для знаходження розв’язку задач опуклого програмування. Процес знаходження розв’язку задачі полягає в тому, що, починаючи з деякої точки $x^{(k)}$, відбувається послідовний перехід до деяких інших точок до тих пір, поки не виявляється прийнятний розв’язок даної задачі.

Найбільш розповсюдженим методом, при якому точки, що досліджуються, не виходять за межі допустимих розв’язків задачі, є метод Франка-Вульфа.

Процес знаходження розв’язку задачі починають з визначення точки, що належить області допустимих розв’язків задачі. Нехай це точка x^k , тоді в цій точці визначають градієнт функції f

$$\nabla f(x^{(k)}) = \left(\frac{\partial f(x^{(k)})}{\partial x_1}; \frac{\partial f(x^{(k)})}{\partial x_2}; \dots; \frac{\partial f(x^{(k)})}{\partial x_n} \right)$$

і будують лінійну функцію

$$F = \frac{\partial f(x^{(k)})}{\partial x_1} \cdot x_1 + \frac{\partial f(x^{(k)})}{\partial x_2} \cdot x_2 + \dots + \frac{\partial f(x^{(k)})}{\partial x_n} \cdot x_n. \quad (6.4)$$

Потім знаходять максимальне значення цієї функції при обмеженні (6.2) і $x_i \geq 0$. Нехай розв'язок задачі визначають точкою $z^{(k)}$. Тоді за новий допустимий розв'язок приймають координати точки $x^{(k+1)}$:

$$x^{(k+1)} = x^{(k)} + \lambda_k (z^{(k)} - x^{(k)}), \quad (6.5)$$

де λ_k – деяке число ($0 \leq \lambda_k \leq 1$). Це число λ_k беруть довільно або визначають мінімальне значення з умови $df/d\lambda_k = 0$. Якщо його значення більше одиниці, то покладають $\lambda_k = 1$. Після визначення числа λ_k знаходять координати точки $x^{(k+1)}$, визначають значення цільової функції в ній і виявляють необхідність переходу до нової точки $x^{(k+2)}$. Якщо така необхідність є, то визначають в точці $x^{(k+1)}$ градієнт цільової функції, переходять до відповідної задачі лінійного програмування і знаходять її розв'язок.

Визначають координати точки $x^{(k+2)}$ і досліджують необхідність проведення подальших розрахунків.

Приклад

Методом Франка-Вульфа знайти розв'язок задачі, що полягає у визначенні максимального значення функції

$$f = 2x_1 + 4x_2 - x_1^2 - 2x_2^2 \quad (6.6)$$

за умов:

$$\begin{cases} x_1 + 2x_2 \leq 8, \\ 2x_1 - x_2 \leq 12, \\ x_1, x_2 \geq 0. \end{cases} \quad (6.7)$$

Розв'язання

Знайдемо градієнт функції

$$\nabla f = \left(\frac{\partial f}{\partial x_1}; \frac{\partial f}{\partial x_2} \right) = (2 - 2x_1; 4 - 4x_2), \quad (6.8)$$

і за початковий допустимий розв'язок задачі виберемо точку $x^{(0)} = (0; 0)$, а за отримуваний розв'язок – нерівність

$$|f(x^{(k+1)}) - f(x^{(k)})| \leq \varepsilon, \quad \text{де } \varepsilon = 0,01. \quad (6.9)$$

Ітерація

В точці $x^{(0)}$ градієнт $\nabla f(x^{(0)}) = (2; 4)$. Знаходимо максимум функції за формулою (6.4): $F_I = 2x_1 + 4x_2$ за умов (6.5) і (6.6):

$$\begin{cases} x_1 + 2x_2 \leq 8, \\ 2x_1 - x_2 \leq 12, \end{cases} \quad (6.10)$$

$$x_1, x_2 \geq 0. \quad (6.11)$$

Запишемо задачу в канонічній формі і визначимо точку z максимуму функції F_1 :

$$\begin{aligned} F_1 &= 2x_1 + 4x_2, \\ x_1 + 2x_2 + x_3 &= 8, \end{aligned} \quad (6.12)$$

$$2x_1 - x_2 + x_4 = 12. \quad (6.13)$$

Нехай x_3, x_4 – базисні змінні, а x_1, x_2 – вільні змінні.

Тоді маємо:

$$x_3 = 8 - x_1 - 2x_2;$$

$$x_4 = 12 - 2x_1 + x_2.$$

Покладемо $x_1^{(0)} = x_2^{(0)} = 0$, тоді $x_3^{(0)} = 8$, $x_4^{(0)} = 12$, цільова функція $F_1^{(0)} = 0$.

Нехай $x_1^{(1)} = 0$, тоді $F_1^{(1)} = 4x_2$.

Збільшуємо x_2 з врахуванням границь x_3 і x_4 :

$$x_3 = 8 - 2x_2 \geq 0, x_2 = 4, \text{ тоді } x_3 = 0;$$

$$x_4 = 12 + x_2 \geq 0.$$

Отже, при $x_1 = 0$, $x_2 = 4$ маємо:

$$x_1^{(1)} = 0, x_2^{(2)} = 4, x_3^{(1)} = 0, x_4^{(1)} = 16.$$

Значення цільової функції при цьому буде дорівнювати:

$$F^{(1)} = 16.$$

Тепер нехай x_3, x_4 базисні змінні, а x_1, x_2 – вільні змінні. Тоді

$$x_2 = 1 - \frac{x_3}{2} - \frac{x_1}{2},$$

$$x_4 = 12 - 2x_1 + x_2,$$

$$F = 2x_1 + 16 - 2x_3 - 2x_1 = 16 - 2x_3.$$

Значення цільової функції не збільшиться, отже, оптимальний варіант буде в точці $z^{(0)} = (0; 4)$.

Знайдемо нові допустимі значення розв'язку задачі за формулою:

$$x^{(1)} = x^{(0)} + \lambda_1 \cdot (z^{(0)} - x^{(0)}), \text{ де } 0 \leq \lambda_1 \leq 1.$$

Підставимо замість $x^{(0)}$ і $z^{(0)}$ їх значення, отримаємо:

$$\begin{cases} x_1^{(1)} = 0 + \lambda_1 \cdot 0, \\ x_2^{(1)} = 0 + \lambda_1 \cdot 4. \end{cases} \quad (6.14)$$

Визначимо тепер число λ_1 з врахуванням (6.5) і (6.14):

$$f(\lambda_1) = 16\lambda_1 - 32\lambda_1^2.$$

Знайдемо похідну цієї функції за λ_1 і прирівняємо її до нуля:

$$f'(\lambda_1) = 16 - 64\lambda = 0.$$

Маємо $\lambda_1 = 0,25$. Оскільки знайдене значення $\lambda_1 < 1$, то приймаємо його за величину кроку. Таким чином:

$$x^{(1)} = (0;1), f(x^{(1)}) = 2, f(x^{(1)}) - f(x^{(0)}) = 2 \geq \varepsilon = 0,01. \quad (6.15)$$

II ітерація.

Градiєнт цiльової функції задачі в точці $x^{(1)} \in \nabla f(x^{(1)}) = (2;0)$. Знаходимо максимум функції $F_2 = 2x_1$ за умов (6.6) і (6.7).

Тобто, знову розв'язуємо задачу лінійного програмування. Оскільки F_2 залежить тільки від x_1 , то визначимо її з врахуванням того, що $x_3 = x_4 = 0$. В даному випадку, розв'язуючи систему рівнянь (6.16)

$$\begin{cases} x_1 + 2x_2 = 8, \\ 2x_1 - x_2 = 12, \end{cases} \quad (6.16)$$

маємо $x_1 = 6,4, x_2 = 0,8, F_2(x_1) = 12,8$.

Оптимальний план $z^{(1)} = (6,4;0,8)$.

Визначаємо тепер $x^{(2)} = x^{(1)} + \lambda_2(z^{(1)} - x^{(1)})$.

Останню рівність перепишемо так:

$$\begin{aligned} x_1^{(2)} &= 6,4\lambda_2, \\ x_2^{(2)} &= 1 - 0,2\lambda_2. \end{aligned} \quad (6.17)$$

Підставимо тепер в (6.4) значення (6.16), отримаємо:

$$f(\lambda_2) = 2 + 12,8\lambda_2 - 41,76\lambda_2^2,$$

звідки $f'(\lambda_2) = 12,8 - 83,52\lambda_2$. Прирівнюємо $f'(\lambda_2)$ до нуля. Тоді $\lambda_2 = 0,15$. Таким чином,

$$\begin{aligned} x_1^{(2)} &= 0,96, \\ x_2^{(2)} &= 0,97, \end{aligned} \quad (6.18)$$

тобто $x^{(2)} = (0,96;0,97), f(x^{(2)}) = 2,9966$,

$$f(x^{(2)}) - f(x^{(1)}) = 2,9966 - 2 = 0,9966 > \varepsilon.$$

III ітерація

Градiєнт функції f в точці $x^{(2)} \in \nabla f(x^{(2)}) = (0,08;0,12)$. Знаходимо максимум функції $F_2 = 0,08x_1 + 0,12x_2$ за умов (6.6) і (6.7), тобто, розв'язуючи задачу лінійного програмування, знаходимо оптимальний план $z^{(2)} = (6;0)$.

Визначимо $x^{(3)} = x^{(2)} + \lambda_3(z^{(2)} - x^{(2)})$. Маємо:

$$x_1^{(3)} = 0,96 + \lambda_3(6 - 0,96) = 0,96 + 5,04\lambda_3;$$

$$x_2^{(3)} = 0,97 + \lambda_3(0 - 0,97) = 0,97 + 0,97\lambda_3;$$

$$f(\lambda_3) = 2,9384 + 0,4032\lambda_3 - 27,3416\lambda_3^2;$$

$$f'(\lambda_3) = 0,4032 - 54,6832\lambda_3.$$

Розв'язуючи рівняння, $f'(\lambda_3) = 0$, $\lambda_3 = 0,007$.

Отже, $x^{(3)} = (0,99528; 0,96321)$, $f(x^{(3)}) = 2,99957$,

$$f(x^{(3)}) - f(x^{(2)}) = 2,99957 - 2,9966 = 0,00297 > \varepsilon.$$

Таким чином, $x^{(3)} = (0,99528; 0,96321)$ є розв'язком даної задачі згідно з вибраним значенням ε . Зменшуючи значення ε , можна показати, що максимальне значення цільової функції буде в точці $x^* = (1; 1)$.

VII Література

1. Эндрюс Дж., Мак-Лоун Р. Математическое моделирование – М.: Мир, 1979. - 217 с.
2. Вознесенский В.А., Лященко Т.В., Огарков Б.Л. Численные методы решения строительно-технологических задач на ЭВМ. – К.: Вища школа, 1989. - 328 с.
3. Борщ И.М., Вознесенский В.А., Мукин В.З. и др. Процессы и аппараты в технологии строительных материалов. – К.: Вища школа, 1981. - 296 с.
4. Подвальный А.М., Проценко А.М. Исследование проницаемости на математических моделях // Защита строительных конструкций зданий от коррозии. – М.: Стройиздат, 1973. – С. 149-156.
5. Вознесенский В.А., Лященко Т.В., Иванов Я.П., Николов И.И. ЭВМ и оптимизация композиционных материалов. – К.: Будівельник, 1989. - 240 с.
6. Моисеев Н.Н., Иванилов Ю.П., Столярова Е.М. Методы оптимизации. – М.: Наука, 1978. - 272 с.
7. Самарский А.А. Введение в численные методы. – М.: Наука, 1983. - 272 с.
8. Зенкевич О. Метод конечных элементов в технике. – М.: Мир, 1975.- 266 с.
9. Норри Д., Ж. де Фриз. Введение в метод конечных элементов. – М.: Мир, 1981. - 305 с.
10. Румишский Л.З. Математическая обработка результатов эксперимента – М.: Наука, 1971. - 192 с.
11. Акулич И.Л. Математическое программирование в примерах и задачах: Учебное пособие для студентов экономических специальностей вузов. – М.: Высшая школа, 1986. - 319 с.
12. Методичні вказівки до виконання лабораторних робіт з курсу „Чисельні методи”/Укладачі Козиренко С.І., Артюк Л.Ю.- Харків: ХДТУР, 1999. - 44 с.
13. Норри Д., Ж. де Фриз. Введение в метод конечных элементов. – М.: Мир, 1981. - 304 с.
14. Хайрер Э., Нерсетт С., Ваннер Г. Решение обыкновенных дифференциальных уравнений. Нежесткие задачи. – М.: Мир, 1990. – 347 с.
15. Плис А.И., Сливина Н.А. Mathcad. Математический практикум.- М.: Финансы и статистика, 1999. – 254 с.
16. Амосов А.А., Дубинский Ю.А., Копченова Н.В. Вычислительные методы для инженеров. – М.: Высшая школа, 1994. – 318 с.
17. Mathcad PLUS 6.0. Финансовые, инженерные и научные расчеты в среде Windows 95: Пер. с англ.- М.: Информационно-издательский дом «Филинь», 1996.- 284 с.
18. Голуб Дж., Ван Лоун Ч. Матричные вычисления. – М.: Мир, 1999. – 211 с.
19. Митчелл Э., Уайт Р. Метод конечных элементов для уравнений с частными производными. – М.: Мир, 1981.- 244 с.