Лабораторна робота № 5

Розробка програми для апроксимації (наближення) табличних даних поліномом за методом найменших квадратів

Мета роботи: оволодіти практичними навичками використання методу найменших квадратів для апроксимації експериментальних даних.

Теоретичні відомості

Задачі апроксимації експериментальних даних функціями заданого вигляду досить часто зустрічаються у інженерній та науковій практиці. При побудові інтерполяційних формул (лабораторна робота 4.) шукають поліном, значення якого у вузлових точках точно дорівнюють табличним значенням. Іншим типом задачі обробки експериментальних табличних даних є апроксимація – наближення функцією заданого вигляду, в частковому випадку - поліномом заданої степені, при якому функція не обов’язково проходить через усі табличні точки, але описує весь їх набір найкращим способом. Застосування такої процедури доцільне у цілому ряді випадків, в першу чергу, коли вигляд функції відомий або табличні точки знайдено (виміряно) з деякою похибкою.

Процедура побудови функції заданого вигляду за табличними даними аналогічна (у випадку функції однієї змінної) процедурі побудови графіка, наприклад, прямої за набором точок. На рисунку 5.1 показано побудову графіка за набором точок. Суцільною лінією показана лінійна залежність, а штрихованою – квадратична.

Розглянемо процес побудови апроксимуючої функції. Для простоти обмежимо вигляд функції поліномом, загальний вигляд якого:

$$f = A_o + A_1 x + A_2 x^2 + \ldots + A_n x^n \qquad (5.1)$$

Для побудови заданого поліному використовується k точок. Очевидно, що повинна виконуватись строга нерівність: $n < k$.

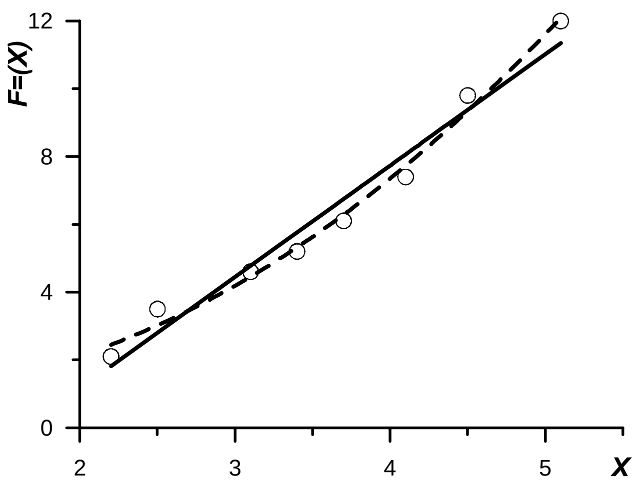

Рисунок 5.1 Побудова графіків заданого вигляду
за набором точок

Для кожної i-ї точки можна записати рівняння:

$$f\left(x_i\right) = y_i = A_0 + A_1 x_i + A_2 x_i^2 + \ldots + A_n x_i^n$$

Зібравши такі рівняння по усіх k точках, ми одержимо систему k рівнянь з $n+1$ невідомими. Невідомими є коефіцієнти $A_0, A_1, \ldots, A_n$, а рівняння – лінійними відносно невідомих.

Відомо, що система лінійних рівнянь має розв'язок тільки у випадку, якщо кількість рівнянь дорівнює кількості невідомих і при цьому відсутні лінійно залежні рівняння. Якщо $k > n+1$, а це є найбільш розповсюдженим випадком, то система рівнянь є перевизначеною і може не мати розв'язку взагалі. В таких випадках говорять про наближений розв'язок системи лінійних рівнянь.

Перепишемо систему у вигляді:

$$\begin{cases} A_0 + A_1 x_1 + A_2 x_1^2 + \ldots + A_n x_1^n - y_1 = q_1 \\ A_0 + A_1 x_2 + A_2 x_2^2 + \ldots + A_n x_2^n - y_2 = q_2 \\ \ldots\ldots \\ A_0 + A_1 x_k + A_2 x_k^2 + \ldots + A_n x_k^n - y_k = q_k \end{cases}, \qquad (5.2)$$

вводячи до розгляду нові змінні q_i, які називають розбігами. Запишемо функцію сумарного розбігу системи (5.2):

$$Q = \sum_{i=1}^{k} q_i^2 \qquad (5.3)$$

Функція сумарного розбігу при постійній таблиці даних залежить виключно від набору коефіцієнтів $A_0, A_1, \ldots, A_n$, причому завдяки своєму вигляду вона має мінімум – точку що дозволить знайти найкращі можливі значення коефіцієнтів A.

Необхідною умовою існування мінімуму функції Q у деякій точці є рівність нулю усіх частинних похідних. Таким чином ми одержимо нову систему рівнянь:

$$\begin{cases} \dfrac{\partial Q}{\partial A_0} = 2\sum\limits_{i=1}^{k}\left(A_0 + A_1 x_i + A_2 x_i^2 + \ldots + A_n x_i^n - y_i\right) = 0 \\ \dfrac{\partial Q}{\partial A_1} = 2\sum\limits_{i=1}^{k}\left(A_0 + A_1 x_i + A_2 x_i^2 + \ldots + A_n x_i^n - y_i\right)\cdot x_i = 0 \\ \ldots\ldots \\ \dfrac{\partial Q}{\partial A_n} = 2\sum\limits_{i=1}^{k}\left(A_0 + A_1 x_i + A_2 x_i^2 + \ldots + A_n x_i^n - y_i\right)\cdot x_i^n = 0 \end{cases} \qquad (5.4)$$

Система (5.4) є системою $n+1$ лінійних рівняння з $n+1$ невідомими, що має єдиний розв’язок – набір значень $A_0, A_1, \ldots, A_n$ - коефіцієнтів, що забезпечують найкраще наближення табличних даних. Кожне рівняння системи (5.4) включає постійний вираз в дужках – фактично розбіг на i-й точці, та змінну величину, що є фактично виразом частинної похідної по відповідному A_j від виразу в дужках.

Спростивши вирази, систему (5.4) можна переписати у вигляді:

$$\begin{cases} A_0 k + A_1\sum\limits_{i=1}^{k} x_i + A_2\sum\limits_{i=1}^{k} x_i^2 + \ldots + A_n\sum\limits_{i=1}^{k} x_i^n = \sum\limits_{i=1}^{k} y_i \\ A_0\sum\limits_{i=1}^{k} x_i + A_1\sum\limits_{i=1}^{k} x_i^2 + A_2\sum\limits_{i=1}^{k} x_i^3 + \ldots + A_n\sum\limits_{i=1}^{k} x_i^{n+1} = \sum\limits_{i=1}^{k} y_i x_i \\ \ldots \quad \ldots \quad \ldots \quad \ldots \\ A_0\sum\limits_{i=1}^{k} x_i^n + A_1\sum\limits_{i=1}^{k} x_i^{n+1} + A_2\sum\limits_{i=1}^{k} x_i^{n+2} + \ldots + A_n\sum\limits_{i=1}^{k} x_i^{2n} = \sum\limits_{i=1}^{k} y_i x_i^n \end{cases} \qquad (5.5)$$

Для розв'язку системи (5.5) можна застосувати будь-який із доступних методів розв'язання систем лінійних рівнянь.

Розглянемо описаний вище алгоритм на прикладі – знайдемо коефіцієнти поліному другої степені, що описує набір даних:

X	2.2	2.5	3.1	3.4	3.7	4.1	4.5	5.1
Y	2.1	3.5	4.6	5.2	6.1	7.4	9.8	12.1

Для того, щоб використати систему (5.5), знайдемо відповідні коефіцієнти:

$k=8$ -кількість рівнянь;

$$\sum_{i=1}^{k} x_i = 2.2 + 2.5 + 3.1 + 3.4 + 3.7 + 4.1 + 4.5 + 5.1 = 28.6$$

$$\sum_{i=1}^{k} x_i^2 = 2.2^2 + 2.5^2 + 3.1^2 + 3.7^2 + 4.1^2 + 4.5^2 + 5.1^2 = 109.02$$

$$\sum_{i=1}^{k} y_i = 2.1 + 3.5 + 4.6 + 5.2 + 6.1 + 7.4 + 9.8 + 12 = 50.8$$

Аналогічно:

$$\sum_{i=1}^{k} x_i^3 = 438.718 \qquad \sum_{i=1}^{k} x_i^4 = 1845.049$$

$$\sum_{i=1}^{k} y_i x_i = 204.03 \qquad \sum_{i=1}^{k} y_i x_i^2 = 857.431$$

Таким чином, систему можна записати як:

$$\begin{cases} 8A_0 + 28.6A_1 + 109.02A_2 = 50.8 \\ 28.6A_0 + 109.02A_1 + 438.718A_2 = 204.03 \\ 109.02A_0 + 438.718A_1 + 1845.049A_2 = 857.431 \end{cases}$$

Розв'язавши систему, одержимо:

$$A_0 = 1.523, \quad A_1 = -0.8357, \quad A_2 = 0.5734$$

Розробка програми

Програма апроксимації табличних даних методом найменших квадратів повинна містити наступні послідовні частини:

1. Блок введення таблиці значень незалежного аргументу та функції;
2. Блок розрахунку коефіцієнтів системи лінійних рівнянь;
3. Процедуру розв’язання системи лінійних рівнянь;
4. Вивід результатів обчислень.

Очевидно, що ключовим є блок розрахунку коефіцієнтів системи лінійних рівнянь. Блоки 1 та 4 достатньо прості і не потребують коментарів. Блок 3 можна реалізувати, наприклад методом Гаусса (лабораторна робота №2).

При реалізації блоку 2 фактично необхідно сформулювати матрицю розмірністю $(n+1)\times(n+2)$, кожен з елементів якої є сумою відповідних виразів. Якщо вважати початком нумерації елементів матриці 0 (нуль), то можна помітити, що кожен з елементів B_{ij}, крім стовпчика вільних членів, включає елементи в степені, що рівна сумі $(i+j)$, а степінь x_i в стовпчику вільних членів дорівнює номеру рядка матриці.

Таким чином, для знаходження елементів матриці системи лінійних рівнянь необхідно організувати подвійний цикл для вибору елемента та внутрішній цикл для знаходження суми.

Завдання

Розробити, скласти та відлагодити програму для апроксимації табличних даних поліномом заданої степені за методом найменших квадратів.

Варіанти:

Програма з розрахунком середньоквадратичної похибки апроксимації.

Програма з побудовую графіка одержаного поліному.

Зробити висновки.

Рекомендована література: [3], [7], [8], [9].