

ЛАБОРАТОРНА РОБОТА 4. Регресійний аналіз

Мета: Розглянути проблему об'єднання результатів різних вимірювань. Вивчити поняття зважених середніх. Розглянути принцип максимальної правдоподібності та метод найменших квадратів і навчитися їх застосовувати.

Теоретичні відомості до завдань 1,2 роботи наведені у лекції 5.

Curve Fitting Toolbox – це набір користувацьких графічних інтерфейсів (GUIs) та М-функцій, створених в обчислювальному середовищі MatLab.

Пакет дозволяє використовувати:

- ✓ Первинну обробку даних, таких як розбиття на області та згладжування.
- ✓ Параметричну і непараметричну апроксимацію: для параметричної апроксимації можна використовувати вбудовану бібліотеку функцій або створити власну. У бібліотеку включені поліноми, експоненти, дрібно-раціональні функції, гаусові суми та ін. Для непараметричної апроксимації використовуються згладжувальні сплайни та різні інтерполяції.

- ✓ Стандартний лінійний метод найменших квадратів, нелінійний метод найменших квадратів, зважений метод найменших квадратів, метод найменших квадратів з обмеженнями та стійкі методи апроксимації.

- ✓ Можливість оцінки якості проведеної апроксимації.

- ✓ Можливості аналізу, такі як екстраполяція, диференціювання та інтегрування.

Curve Fitting Toolbox дає можливість:

1. Працювати з даними, заданими за допомогою векторів в робочому середовищі MatLab, та, за необхідності, забезпечувати дані вагами, також задаючи вектор їх значень.

2. Графічно відображати досліджувані дані.

3. Здійснювати попередню обробку даних, виключаючи частину даних за деяким правилом, по точках в таблиці, або за допомогою миші на графіку, або навпаки залишаючи частину даних для подальшої обробки.

4. Згладжувати та фільтрувати дані різними способами.

5. Наближати дані за допомогою параметричних моделей, в яких шукані параметри можуть входити як лінійно, так і нелінійно. Моделі вибираються з бібліотеки моделей, або задаються користувачем. Цільова функція помилки та методи, що застосовуються для її мінімізації, також можуть бути різними. Допускається задання обмежень на шукані значення параметрів. Крім параметричних моделей, можлива інтерполяція даних сплайнами та наближення згладжуваними сплайнами.

6. Відображати побудовані наближення графічно, формувати графіки та зберігати результати в окремих графічних вікнах.

7. Обчислювати різні критерії придатності отриманого наближення.

8. Проводити ряд операцій з отриманими наближеннями (обчислювати в заданих точках, диференціювати, інтегрувати, проводити екстраполяцію), графічно відображати результат.

До складу Curve Fitting Toolbox входить додаток `cftool` з графічним інтерфейсом користувача, який дозволяє проводити всі названі вище дії, та функції, призначені для визначення параметричної моделі, підбору параметрів, аналізу придатності наближення, операцій з ним і графічного відображення результату.

Додаток `cftool` дає можливість користувачу:

1) імпортувати дані та вагові коефіцієнти, задані в векторах робочого середовища MatLab;

2) експортувати отримане наближення та аналіз результатів у робоче середовище MatLab;

3) зберігати сесію у файлі для продовження роботи;

4) генерувати `m`-файл з файл-функцією, яка може бути використана для наближення інших даних обраним в `cftool` способом та для графічного відображення результату без використання програми `cftool`.

В одній сесії `cftool` допускається робота одночасно:

- з кількома наборами даних;

- з різними наближеннями для кожного набору;

- з підмножинами даних, отриманих після виключення з основної підмножини за різними правилами.

Для роботи з додатком `cftool` практично не потрібне знання MatLab, оскільки всі основні операції виконуються за допомогою графічного інтерфейсу користувача. Функції, що входять до складу Curve Fitting Toolbox, можуть бути використані, наприклад, при розробці власних програм, в яких потрібно реалізувати розв'язання задачі про наближення даних параметричною моделлю.

Всі функції Curve Fitting Toolbox можуть бути умовно розділені на п'ять категорій.

1. *Задання параметричних моделей та різних опцій, підбір параметрів та обчислення критеріїв придатності наближення:*

`fittype` – вибір однієї зі стандартних параметричних і непараметричних моделей або створення довільної параметричної моделі.

`fitoptions` – задання різних опцій, що стосуються правил виключення, вагових коефіцієнтів, допустимих меж значень шуканих параметрів, початкового наближення до параметрів, опцій обчислювальних алгоритмів, застосовуваних для мінімізації цільової функції (помилки наближення).

`fit` – основна функція, що виконує підбір параметрів, в якій задаються вихідні дані, параметричну модель, опції (див. функцію `fitoptions`); вона повертає: отриману параметричну модель зі знайденими значеннями параметрів, та різні критерії придатності отриманого наближення.

`cfit` – створення параметричної моделі з заданими значеннями параметрів (без їх визначення за допомогою функції `fit`).

2. Інформація про моделі й об'єкти, створювані функціями Curve Fitting Toolbox:

`cflibhelp` – отримання інформації про стандартні параметричні моделі, що входять до Curve Fitting Toolbox.

`disp` – отримання інформації про об'єкти, що створюються функціями Curve Fitting Toolbox.

3. Задання та отримання значень властивостей об'єктів, що створюються функціями Curve Fitting Toolbox:

`get` – отримання властивостей об'єктів.

`set` – задання властивостей об'єктів.

4. Отримання статистичної інформації про дані та їх попередня обробка:

`excludedata` – пошук даних, які повинні бути виключені з вихідного набору згідно з обраним правилом.

`smooth` – згладжування даних з використанням різних способів.

`datastats` – визначення мінімального та максимального значень, середнього, стандартного відхилення.

5. Робота з отриманим наближенням:

`confint` – визначення довірчих інтервалів для обчислених параметрів моделі.

`predint` – визначення інтервалів прогнозів досліджених значень із заданою імовірністю.

`differentiate` – диференціювання побудованої параметричної моделі.

`integrate` – інтегрування побудованої параметричної моделі.

`feval` – обчислення значень параметричних моделей.

`plot` – побудова графіків даних, параметричних моделей, помилок.

Детальну інформацію про Curve Fitting Toolbox (демонстрація роботи, пробна версія тощо) можна отримати на офіційному сайті MathWorks: <https://nl.mathworks.com/products/curvefitting.html>.

Порядок виконання роботи

Робота складається з трьох завдань:

Завдання 1: До власної або запропонованої вибірки згідно з номером свого варіанту (див. відповідний текстовий файл у теках 'ФФ-91-93-Апроксимація степеневим поліномом' та 'ФФ-91-93-Апроксимація

тригонометричним поліномом³) застосувати МНК для функції однієї змінної за допомогою пакету MatLab, а саме:

1. Провести апроксимацію степеневими поліномами з рівнем значущості 0.05 та максимально допустимим степенем полінома 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
2. Провести апроксимацію степеневими поліномами з рівнем значущості 0.01 та максимально допустимим степенем полінома 20. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
3. Провести апроксимацію степеневими поліномами з рівнем значущості 0.01 та максимально допустимим степенем полінома 20. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
4. Провести тригонометричну апроксимацію з рівнем значущості 0.01 та максимальним номером гармоніки 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
5. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
6. Провести апроксимацію степеневими поліномами з рівнем значущості 0.05 та максимально допустимим степенем полінома 20. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
7. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
8. Провести тригонометричну апроксимацію з рівнем значущості 0.01 та максимальним номером гармоніки 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
9. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальними номерами гармоніки 7 та 10. Порівняти результати. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
10. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 10. Виводити кількість

експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.

11. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 5. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
12. Провести апроксимацію степеневими поліномами з рівнем значущості 0.05 та максимально допустимим степенем полінома 20. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
13. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 5. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
14. Провести апроксимацію степеневими поліномами з рівнем значущості 0.01 та максимально допустимим степенем полінома 20. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.
15. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 5. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
16. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 5. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
17. Провести тригонометричну апроксимацію з рівнем значущості 0.05 та максимальним номером гармоніки 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном.
18. Провести апроксимацію степеневими поліномами з рівнем значущості 0.1 та максимально допустимим степенем полінома 10. Виводити кількість експериментальних точок, довірчі інтервали для параметрів моделі та апроксимуючий поліном оптимального степеня.

Завдання 2: Виконати наступне завдання згідно з номером свого варіанту (взяти власні дані або див. відповідний текстовий файл) для функції кількох змінних за допомогою пакету MatLab (дані для варіантів 1, 3, 7–9, 11, 12, 15–18 розміщуються у теці ‘ФФ-91-93-Апроксимація степеневим поліномом’, для варіантів 2, 4, 13 – у теці ‘ФФ-91-93-Крива насичення’, для варіантів 5, 6, 10, 14 – у теці ‘ФФ-91-93-Лінійна функція двох змінних’):

1. Провести мінімізацію суми модулів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня

за МНК та мінімумом суми модулів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів відхилень.

2. Провести нелінійну апроксимацію вибірки за МНК. Виводити кількість експериментальних точок, аналітичний вигляд кривої насичення, довірчі інтервали для параметрів моделі, а також (в одне графічне вікно) експериментальні точки та теоретичну криву насичення.
3. Провести мінімізацію суми модулів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів відхилень.
4. Провести нелінійну апроксимацію вибірки за МНК. Виводити кількість експериментальних точок, аналітичний вигляд кривої насичення, довірчі інтервали для параметрів моделі, а також (в одне графічне вікно) експериментальні точки та теоретичну криву насичення.
5. Побудувати лінійну модель для 2-факторного експерименту з рівнем значущості 0.05 за МНК. Виводити кількість рівнів факторів, довірчі інтервали для параметрів моделі на заданому рівні значущості, аналітичний вигляд лінійної моделі, а також (в одному графічному вікні) теоретичну поверхню (площину) та експериментальні точки.
6. Побудувати лінійну модель для 2-факторного експерименту з рівнем значущості 0.05 за МНК. Виводити кількість рівнів факторів, довірчі інтервали для параметрів моделі на заданому рівні значущості, аналітичний вигляд лінійної моделі, а також (в одному графічному вікні) теоретичну поверхню (площину) та експериментальні точки.
7. Провести мінімізацію суми модулів кубів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів кубів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів кубів відхилень.
8. Провести мінімізацію суми модулів кубів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів кубів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів кубів відхилень.
9. Мінімізувати суму модулів відхилень, максимальне за модулем відхилення та суму модулів кубів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд відповідних поліномів 2-го степеня та

за МНК, а також (в одне графічне вікно) експериментальні точки та всі теоретичні криві.

10. Побудувати лінійну модель для 2-факторного експерименту з рівнем значущості 0.05 за МНК. Виводити кількість рівнів факторів, довірчі інтервали для параметрів моделі на заданому рівні значущості, аналітичний вигляд лінійної моделі, а також (в одному графічному вікні) теоретичну поверхню (площину) та експериментальні точки.
11. Провести мінімізацію суми модулів кубів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів кубів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів кубів відхилень.
12. Провести мінімізацію суми модулів кубів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів кубів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів кубів відхилень.
13. Провести нелінійну апроксимацію вибірки за МНК. Виводити кількість експериментальних точок, аналітичний вигляд кривої насичення, довірчі інтервали для параметрів моделі, а також (в одне графічне вікно) експериментальні точки та теоретичну криву насичення.
14. Побудувати лінійну модель для 2-факторного експерименту з рівнем значущості 0.05 за МНК. Виводити кількість рівнів факторів, довірчі інтервали для параметрів моделі на заданому рівні значущості, аналітичний вигляд лінійної моделі, а також (в одному графічному вікні) теоретичну поверхню (площину) та експериментальні точки.
15. Провести мінімізацію суми модулів відхилень експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом суми модулів відхилень, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та мінімумом суми модулів відхилень.
16. Провести мінімізацію максимального за модулем відхилення експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом максимального за модулем відхилення, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та максимального за модулем відхилення.
17. Провести мінімізацію максимального за модулем відхилення експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом максимального за

модулем відхилення, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та максимального за модулем відхилення.

18. Провести мінімізацію максимального за модулем відхилення експериментальних точок від теоретичних значень. Виводити аналітичний вигляд поліномів 2-го степеня за МНК та мінімумом максимального за модулем відхилення, а також (в одне графічне вікно) експериментальні точки та теоретичні криві за МНК та максимального за модулем відхилення.

Завдання 3. Ознайомитися з основними функціями Curve Fitting Toolbox MatLab, застосовуючи їх або до власної вибірки результатів (або див. текстовий файл з номером, відповідним до номеру у списку групи).

Контрольні питання

1. Проблема об'єднання результатів різних вимірювань.
2. Зважене середнє.
3. Принцип максимальної правдоподібності.
4. Апроксимація методом найменших квадратів.
5. Реалізація задач регресійного аналізу в системі MatLab.
6. Основи роботи у Curve Fitting Toolbox MatLab.