



УДК 519.237.5

РОЗРОБКА МАТЕМАТИЧНОГО І ПРОГРАМНОГО ЗАБЕЗПЕЧЕННЯ ДЛЯ СПРОЩЕННЯ ЕМПІРИЧНИХ МОДЕЛЕЙ СКЛАДНИХ СИСТЕМ

Студ. А.М. Суховерхий

Науковий керівник проф. С.М. Краснитський
Київський національний університет технологій та дизайну

Мета роботи - розробити програмне забезпечення для спрощення емпіричних моделей складних систем на основі результатів математичного моделювання за допомогою спеціальних методів [1] та дослідити вплив різних факторів на процеси спрощення моделей. Показниками зазначеного впливу можуть бути:

1. Визначення ступеня детермінованості варіації критеріальної (залежної) змінної предикатами (незалежними змінними).
2. Прогнозування значення залежної змінної за допомогою незалежної.
3. Визначення внеску окремих незалежних змінних у варіацію залежної.

Завдання роботи полягає у розробці програмного забезпечення для оптимального в певному сенсі рішення поставленої задачі, зокрема створення програмного продукту для спрощення емпіричних моделей складних систем, а також порівняння отриманих результатів для окремих моделей та формулювання висновків про якість роботи розглянутих алгоритмів.

Об'єктом дослідження є методи моделювання процесу спрощення емпіричних моделей складних систем.

Предметом дослідження є показники та результати чисельних експериментів.

Методи та засоби дослідження. При побудові алгоритму тестування методів вибору ознак використовувалися: елементи теорії ймовірностей; багатовимірної статистики; стійкого оцінювання параметрів залежностей; обчислювальної лінійної алгебри. Для програмної реалізації розробленого алгоритму використовувалася середа C++ Visual Studio 2012.

Наукова новизна та практичне значення отриманих результатів. У зв'язку з дослідженням складних систем виникають численні статистичні задачі. Ця обставина пояснюється не в останню чергу тим, що головним джерелом відомостей щодо характеристик складних систем і особливо їх елементів є спостереження над реально функціонуючою системою. Дослідження питань про практичні якості методів адекватного спрощення моделей складних систем (див. перелік, розміщений нижче) ще далеко від свого завершення [1,2]. Зазначені дослідження використовують як методи аналітичних представлень, так і чисельних експериментів. У напрямку реалізації останніх підходів одержано результати функціонування розробленого програмного забезпечення.

Результати дослідження. Вирішення практичних завдань відновлення регресійних залежностей вимагають розгляду великого числа породжуваних ознак. Етапи побудови регресійних моделей можна розділити два кроки. На першому кроці, на основі вільних змінних – результатів вимірювань – породжується набір ознак, на другому кроці – вибір ознак. При виборі ознак виконується налаштування параметрів моделі і оцінюється її якість. Модель з налаштованими параметрами, що доставляє мінімум заданому функціоналу якості, називається моделлю оптимальної структури. Регресійний аналіз використовується в тому випадку, якщо відношення між змінними можуть бути виражені кількісно у виді деякої комбінації цих змінних. Отримана комбінація використовується для передбачення значення, що може приймати цільова

(залежна) змінна, яка обчислюється на заданому наборі значень вхідних (незалежних) змінних. У найпростішому випадку для цього використовуються стандартні статистичні методи, такі як лінійна регресія. На жаль, більшість реальних моделей не вкладаються в рамки лінійної регресії. Тому в сучасній літературі для її розв'язання використовуються різні методи. Зокрема в [1], [2] наводяться наступні методи:

- 1) повний перебір моделей;
- 2) метод головних компонент;
- 3) генетичні алгоритми;
- 4) метод групового врахування аргументів;
- 5) покрокова регресія;
- 6) алгоритм Лассо;
- 7) східчаста регресія;
- 8) послідовне додавання ознак з ортогоналізацією;
- 9) метод найменших кутів LARS;
- 10) оптимальне проріджування;
- 11) метод моделей найбільшої обґрунтованості.

Стандартне програмне забезпечення (програмні комплекси Statistica, Statgraphics та ін.) далеко не завжди надає можливість реального використання зазначених методів. Основним результатом роботи є розробка програмного забезпечення для можливості використання методів 2), 8) з числа наведених вище. Метод 2) (головних компонент) має кілька основних базових версій:

- а) апроксимація даних лінійним многовидами меншої розмірності;
- б) знаходження підпросторів меншої розмірності, в ортогональній проекції на які розкид даних x максимальним;
- в) знаходження підпросторів меншої розмірності, в ортогональній проекції на які середньоквадратичні відстані між точками є максимальним;
- г) для даної багатовимірної випадкової величини побудувати таке перетворення координат, що в результаті кореляція між окремими координатами дорівнюватиме 0.

Метод 8) у нашому виконанні використовує ортогоналізацію стовпців матриці експерименту X за алгоритмом Грама — Шмідта: У матричному вигляді процедура може бути представлена наступним чином [2]:

де Z — матриця вже перетворених стовпців, Z_i — наступний стовпець матриці X , що перетворюється, Z_{iT} — вектор-результат перетворення, ортогональний стовпцям матриці Z .

Вказані методи застосовано для розробки математичних моделей деяких складових процесу створення текстильних виробів, що апріорі вимагають врахування великої кількості чисельних показників різних етапів виробництва.

Висновки. Програмне забезпечення, що реалізує вищеописані кроки, дозволить раціоналізувати роботу дослідника. Знання оптимальних параметрів процесу спрощення емпіричних моделей дозволить ефективно впровадити це на практиці, а саме – дасть змогу ефективно використовувати ці методи в різних галузях.

Ключові слова. Емпіричні моделі, регресійний метод, обчислювальна лінійна алгебра, предикати, варіація.

ЛІТЕРАТУРА:

1. Стрижов В. В. Крымова Е. А. Методы выбора регрессионных моделей/В.В. Стрижов, Е.А. Крымова. – Вычислительный центр РАН Москва, 2010.