МОДИФІКАЦІЯ r-АЛГОРИТМУ ДЛЯ ЗАДАЧІ КВАНТИЛЬНОЇ РЕГРЕСІЇ[1]

Супрун А.А.

anton_s2007@ukr.net

Інститут кібернетики імені В.М. Глушкова НАН України

Задача знаходження m-вимірного вектора β параметрів моделі квантильної регресії за вибіркою $(y_i, x_i)_{i=1}^n$ зазвичай формулюється як задача лінійного програмування і для її розв’язання застосовуються варіанти симплекс метода або метода внутрішніх точок. В [1] показано, що цю задачу можна сформулювати як задачу мінімізації штрафної функції такого вигляду:

$$\sum_{i=1}^{n} \rho_\tau (y_i - x_i'\beta) = \sum_{i=1}^{n} \max(\tau(y_i - x_i'\beta), -(1-\tau)(y_i - x_i'\beta)), \tag{1}$$

де τ - значення квантилю. При $\tau = 0.5$ отримуємо модель медіанної регресії.

Для мінімізації функції (1) можна застосовувати r-алгоритм [1]. Оскільки розмірність m вектора β може бути досить великою, розглянемо модифікацію r-алгоритму [2], яку можна використати для подібних задач.

На кожній ітерації r-алгоритму використовується матриця B_i, яка перераховується за наступною формулою:

$$B_{i+1} = B_i R_\beta(\xi_i) = B_i (I + (\beta - 1)\xi_i \xi_i^T), \tag{2}$$

де ξ_i – нормалізований вектор різниці 2 послідовних субградієнтів, а матриця $R_\beta(\xi_i)$ задає оператор розтягу простору в напрямку ξ_i, що направлений на зменшення ступеню яружності функції. В [2] показано, що кількість арифметичних операцій в (2) можна зменшити, якщо у ξ_i залишити лише $k < m$ найбільших за модулем елементів. Таким чином, оскільки нульові елементи в ξ_i не впливають на перерахунок B_i, кількість операцій множення в (2) зменшується з $2m^2 + m$ до $2km + m$. Значення k може бути вибраним наперед, або визначатись адаптивно під час роботи алгоритму [2].

---

[1] Робота підтримана CRDF Global (грант G-202102-68020)

Якщо при цьому матриця B_0 буде одиничною матрицею і буде вибиратися один і той же фіксований набір k змінних, за якими визначаються ненульові елементи ξ_i, то отримаємо алгоритм, в якому оператор розтягу $R_\beta(\xi_i)$ діє лише на відповідному k-вимірному підпросторі всього простору змінних. Це означає, що на кожній ітерації для матриці B_i в (2) будуть перераховуватись тільки одні і ті самі k^2 елементів, тому, замість зберігання матриці B_i розміром m^2 достатньо зберігати відповідну підматрицю розміром k^2.

На основі описаної модифікації можна створювати ефективні гібридні субградієнтні алгоритми, в яких на деякому підпросторі змінних застосовується r-алгоритм, а на доповненні до нього можна застосовувати, наприклад, один із варіантів алгоритму субградієнтного спуску з спеціальним вибором кроку у напрямку спуску.

Цей підхід може бути ефективним, якщо яружність функції визначається деяким підпростором змінних. Такі змінні можна спробувати ідентифікувати, наприклад, шляхом попереднього застосування алгоритму узагальненого градієнтного спуску на декількох ітераціях, і на основі чого вибрати ті змінні, відповідні компоненти субградієнтів яких змінювалися найбільше (мали найбільшу варіативність).

Наразі проводяться дослідження із застосування описаної модифікації *r*-алгоритму для мінімізації яружних штрафних функцій (1) з великими *m*.

Список літератури

1. **Suprun A.** Computational aspects of quantile regression. – Тези доповідей XVIII міжнародної науково-практичної конференції «Математичне та програмне забезпечення інтелектуальних систем (MPZIS-2020)», м. Дніпро, 18-20 листопада 2020 р. Дніпро: ДНУ, 2020. С. 241–242.

2. **Стецюк П.І., Жмуд О.О.** Про прискорену реалізацію оператору розтягу простору // Питання прикладної математики і математичного моделювання. – Там же. С. 239–240.