Лабораторна робота №6

Оптимізація функції декількох змінних методом Ньютона

Мета роботи: здобуття практичних навичок використання методу Ньютона для оптимізації функції декількох змінних.

Теоретичні відомості.

Метод Ньютона є одним з методів оптимізації функції декількох змінних, що носить найбільш загальну стратегію пошуку. Цього досягають, залучаючи інформацію не тільки про перші, але й про другі похідні функції, таким чином реалізуючи алгоритм другого порядку. Метод Ньютона можна представити ітерацією:

$$x^{(k+1)} = x^{(k)} - \nabla^2 Q\left(x^{(k)}\right)^{-1} \cdot \nabla Q\left(x^{(k)}\right) \tag{6.1}$$

де $x^{(i)}$ - вектор координат *i*-ї точки наближення - $\left\lfloor x_1^{(i)},\ \ x_2^{(i)},\ \ \ldots \right\rfloor$;

$\nabla^2 Q(x)$ - матриця Гьоссе.

Додаткова складність методу полягає в необхідності знаходження компонентів матриці Гьоссе за загальною формулою:

$$\Gamma_{ij} = {\partial^2 Q}\Big/{\partial x_i \partial x_j} \tag{6.2}$$

Чисельне знаходження других похідних можна проводити за модифікованими формулами. Приклад формул для компонентів матриці Гьоссе для функції двох змінних наведено нижче:

$$\Gamma_{11} = \frac{\partial^2 Q}{\partial x_1^2} \approx \frac{Q\left(x_1^k + g, x_2^k\right) - 2 \cdot Q\left(x_1^k, x_2^k\right) + Q\left(x_1^k - g, x_2^k\right)}{g^2};$$

$$\Gamma_{22} = \frac{\partial^2 Q}{\partial x_2^2} \approx \frac{Q\left(x_1^k, x_2^k + g\right) - 2 \cdot Q\left(x_1^k, x_2^k\right) + Q\left(x_1^k, x_2^k - g\right)}{g^2};$$

$$\Gamma_{12} = \Gamma_{21} = \frac{\partial^2 Q}{\partial x_1 \partial x_2} \approx \tag{6.3}$$

$$\approx \frac{Q\left(x_1^k + g, x_2^k + g\right) - Q\left(x_1^k - g, x_2^k + g\right) - Q\left(x_1^k + g, x_2^k - g\right) - Q\left(x_1^k - g, x_2^k - g\right)}{4 \cdot g^2}.$$

Розглянемо роботу метода на прикладі:

Знайти мінімум функції $\mathbf{Q=100\cdot(x_2-0.5\cdot x_1^2)^2+3\cdot(1-x_1)^2}$, починаючи з точки **[-1, 1]**.

$x_1^{(0)}=-1$ $x_2^{(0)}=1$ $Q(-1,\ 1) = 37.00$

Знайдемо компоненти вектора градієнта при базі оцінювання похідних $g=0.1$:

$\nabla Q(X^{(0)}) = [87,\ 100]^T$ літера T свідчить, що матриця чи вектор транспоновані, іншими словами градієнт являє собою вектор-стовбчик.

Застосувавши формули (6.3) знайдемо коефіцієнти матриці Гьоссе.

$\Gamma_{11}=106.5$ $\Gamma_{12}= \Gamma_{21}=200$ $\Gamma_{22}=200$

Таким чином ітерацію (4-38) можна записати в вигляді:

$$X^{(1)} = [-1,1]^T - \begin{bmatrix} 106.5 & 200 \\ 200 & 200 \end{bmatrix}^{-1} \cdot [87,100]^T$$

звідки $X^{(1)}=[-1,\ 1] - [0.139,\ 0.361] = [-1.139,\ 0.639]$

$Q(X^{(1)})= Q(-1.139,\ 0.639) = 13.74$

Повторимо проведені операції зменшивши базу точкової оцінки похідних до $g=0.05$.

$\nabla Q(X^{(1)}) = [-15.32,\ -1.93]^T$

$\Gamma_{11}=267.54$ $\Gamma_{12}= \Gamma_{21}=227.81$ $\Gamma_{22}=200$

$X^{(2)}=[-1.139,\ 0.639] - [-1.628,\ 1.845] = [0.49,\ -1.2]$

$Q(X^{(2)})= Q(0.49,\ -1.2) = 176.41$ – що набагато гірше, ніж попереднє наближення.

$\nabla Q(X^{(2)}) = [126.67,\ -265.05]^T$

$\Gamma_{11}=319.00$ $\Gamma_{12}= \Gamma_{21}=-97,8$ $\Gamma_{22}=200$

$X^{(3)}=[\ 0.49,\ -1.2] -[-0.01,\ -1.3] = [0.5,\ 0.12]$

$Q(X^{(3)})= Q(0.5,\ 0.12) = 0.75$

Геометрично реалізація методу Ньютона є непоказовою, оскільки відповідне зображення матриці – неможливе.

Основні переваги методу Ньютона реалізуються у випадках, коли:

- поверхня функції близька до квадратичної, не містить сідловин точок, точок перегину, тошо;
- існує спосіб знаходження компонентів градієнта функції та матриці Гьоссе з достатньою точністю;
- доступні алгоритми швидкого знаходження оберненої матриці.

Рекомендації з розрахунків

Метод Ньютона, у випадку реалізації для функції декількох змінних, є не досить зручним для розрахунків в середовищі електронних таблиць. Причиною цього є значна кількість математичних виразів, яка має бути реалізована у вигляді функцій. Опис функцій значно простіший у мовах програмування високого рівня.

З іншого боку мови програмування, наприклад *pascal*, на містять вбудованих можливостей розрахунку оберненої матриці. Тому, для реалізації методу Ньютона доцільно застосовувати гібридні середовища – такі, що містять можливості опису функцій і достатньо розвинуті бібліотеки лінійної алгебри. Такими середовищами можуть бути математичні моделюючі системи типу Matlab або Scilab. Також можна застосувати систему електронних таблиць, яке підтримує програмування, наприклад таблицю MS Excel з програмою на VBA.

В останньому випадку зручно в середовищі VBA створити модуль, в якому описати функцію (вихідну, яку необхідно оптимізувати), та функції, що реалізують відповідні перші та другі частинні похідні, а безпесереднє знаходження оберненої матриці та множення матриць виконувати в середовищі електронної таблиці.

Необхідно враховувати, що для знаходження комонентів градієнта (перших частинних похідних) потрібні значення функції в двох точках – всього 2n точок, а для знаходження компонентів матриці Гьоссе, додатково в 1 (середній) точці для діагональних елементів, та в чотирьох точках для недаоганальних елементів. Враховуючи кількість недіагональних елементів матриці $k = n^2 - n$, а токож симетричність матриці, загальна кількість значень функції для однієї ітерації складе:

$m = 2 \cdot n + 2 \cdot (n^2 - n) + 1 = 2 \cdot n^2 + 1$, або в перерахунку на функцію трьох змінних: 19 точок.

Завдання на лабораторну роботу

Використовуючи метод Ньютона, знайти мінімум функції 3-х змінних за завданням викладача.

Контрольні запитання:

1. Чому метод Ньютона невигідно застосовувати для експериментальної оптимізації ?
2. Які обмеження в використанні методу Ньютона ?
3. Як визначити точність роботи методу Ньютона?

Рекомендована література: [1], [4], [5]