УДК 004.021

ЯРУЖНИЙ КРОК АЛГОРИТМІВ СПУСКУ

В.М. Яхно, к.т.н., доцент
Київський національний університет технологій та дизайну

Ключові слова: яружні функції, крок алгоритмів спуску, яружний крок алгоритмів спуску, швидкість збіжності алгоритмів спуску.

Методи нелінійної оптимізації є однією з базових технологій мехатроніки [1]. Один з найвідоміших методів опуклої оптимізації – метод градієнтного спуску. На кожній ітерації розрахунки відбуваються за наступною схемою

$x_{k+1} = x_k - h_k f'(xk)$, $k = 0, 1 \ldots$ (1)

Скалярний - множник h_k перед градієнтом називають довжиною кроку. Очевидно, він повинен бути позитивним. Існує багато різновидів цього методу, які відрізняються один від одного стратегією вибору довжини кроку на кожній (к-тій) ітерації.

Найбільш важливі з них.

1.Постійний крок. $h_k = h>0$, Для функцій, градієнт яких задовольняє умові Ліпшиця - існує таке число L що нерівність

$\|f'(x) - f'(y)\| < L\|x - y\|$

має місце для всіх x, y ∈ R^n. В цьому випадку найкращим є крок $h = 1/L$[1].

2. З дробленням кроку. Крок зменшується с заданим коефіцієнтом якщо не досягається зменшення функції за формулою (1).

3. Найшвидший спуск (або метод повної релаксації)

$h_k = \arg\min f(x_k - hf'(x_k))$, мінімум визначається по $h>0$.

4.Крок знаходиться відповідно до правила Голдштейна - Армійо: знайти таке h що

$x_{k+1} = x_k - hf'(x_k)$, та мають місце нерівності

$\alpha f'(x_k), x_k - x_{k+1}) \le f(x_k) - f(x_{k+1})$,

$\beta f'(x_k), x_k - x_{k+1}) \ge f(x_k) - f(x_{k+1})$,

де α, β, $0<\alpha<\beta<1$, — деякі фіксовані параметри.

В останньому випадку якість алгоритму суттєво залежить від вибору фіксованих значень α, β.

Важливішим параметром будь якого релаксаційного алгоритму є швидкість збіжності. Вважається що для будь якої модифікації алгоритму і для будь якому розмірі простору аргументів швидкість збіжності (кількість ітерацій, що потрібна для розв'язання задачі) залежить від яружності функції, що мінімізується [2]. Для функції $f(x,y) = x^2+y^2$ необхідна одна ітерація алгоритму градієнтного спуску з мінімізацією функції вздовж напрямку градієнту, а для подібної функції $f(x,y) = x^2+5y^2$ кількість ітерацій може наближатися до нескінченності. Ситуацію ілюструє малюнок 1, що наведений нижче. На малюнку зображено траєкторію спуску найбільш поширених модифікацій градієнтного спуску - з постійним кроком, з дробленням кроку і найшвидший спуск для функції $f(x,y) = x^2+5y^2$. Для $f(x,y) = x^2+10y^2$ кількість

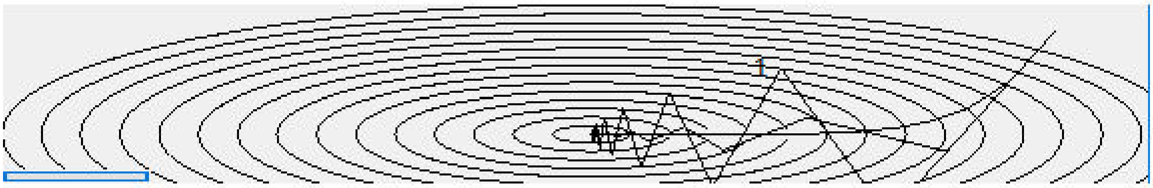

Рисунок 1.

ітерацій ще більше. Цю інформацію можна отримати в будь якому підручнику з програмування. Але існує стратегія яка дозволяє для подібних функцій отримати розв'язок за дві ітерації градієнтного спуску. Результат роботи і ідею алгоритму, що пропонується ілюструє малюнок 2. На цьому малюнку в точці 2 напрямок антиградієнту збігається з ньютонівським напрямком спуску $-f'(x) f''(x)^{-1}$ і тому задовольняє рівнянню $f'(x) = f'(x) f''(x)^{-1}$. Це значить, що антиградієнт в цій точці є власним вектором матриці $f''(x)^{-1}$. Матриці $f''(x)^{-1}$ і $f''(x)$ мають однакові власні вектори і це можна використати для визначення кроку – крок h повинен бути таким, щоб кут між векторами $f'(x_k - hf'(x_k))$ і $f'(x_k - hf'(x_k)) f''(x)$ був мінімальним. Результат застосування такої стратегії наведений на малюнку 2.

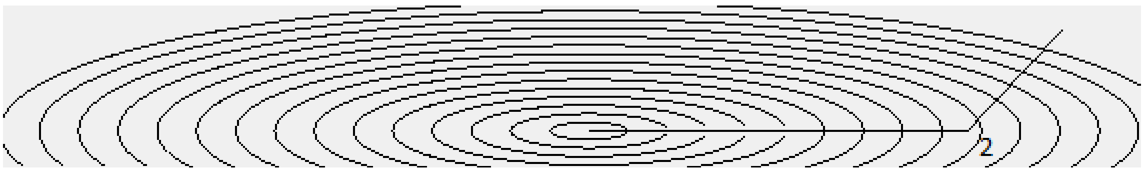

Рисунок 2.

Для використання такої стратегії знаходити обернену матрицю других похідних або розв'язувати систему лінійних рівнянь не потрібно. Алгоритм пошуку необхідного кроку э типовим. Під час пошуку необхідно перевіряти умови Голдштейна – Армійо. Порівняно із стратегіями 1-4 ускладнює алгоритм необхідність вираховувати матрицю других похідних (або якусь оцінку цієї матриці). Автор пропонує ще одну стратегію – шукати крок що мінімізує не функцію цілі а норму градієнта цієї функції. Нижче рисунок 3 дозволяє порівняти стратегію, що базується на мінімізації функції цілі (стратегія 3) з стратегією мінімізації норми градієнта цієї функції.

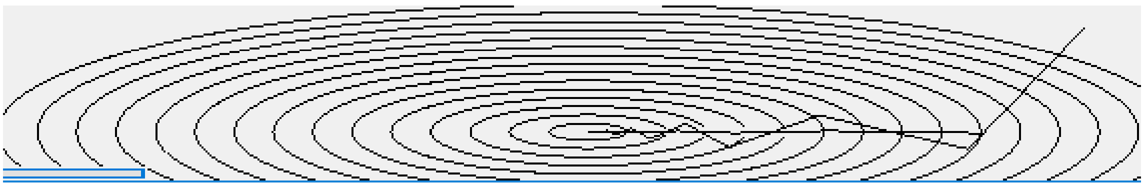

Рисунок 3.

Список використаних джерел

1. Егоров О.Д. Конструирование мехатронных модулей: Учебник. / О.Д. Егоров, Ю.В. Подураев. - М.: ИЦ МГТУ "СТАНКИН", 2004.- 360 с.

2. Нестеров Ю.Е. Методы выпуклой оптимизации.- М. Издательство МЦНМО, 2010, 241 с.