М ЕТОДИ ПОШУКУ ТОЧКИ ЕКСТРЕМУМУ З ВИКОРИСТАННЯМ ПОХІДНИХ[1]

Загальні положення

Усі розглянуті в попередніх розділах методи пошуку ґрунтуються на припущеннях про унімодальність й у ряді випадків про безперервність досліджуваної цільової функції. Доцільно припустити, що, якщо на додаток до умови безперервності ввести вимогу диференційованості функції, то ефективність пошукових процедур можна істотно підвищити. Нагадаємо, що у розд. 1 установлено необхідну умову існування локального мінімуму функції в деякій точці x^*, відповідно до якого перша похідна функції в точці x^* повинна перевтілюватися в нуль, тобто

$$f'(x^*) = df/dx|_{x=x^*} = 0.$$

Якщо функція $f(x)$ містить члени, що включають x у третій і більш високих ступенях, то безпосереднє вирішення аналітичного розв'язання рівняння $f'(x) = 0$ може виявитися скрутним. У таких випадках використовуються наближені методи послідовного пошуку стаціонарної точки функції f. Насамперед розглянемо класичну пошукову схему, орієнтовану на перебування кореня нелінійного рівняння. Ця схема була розроблена Ньютоном і пізніше уточнена Рафсоном [1].

1 Метод Ньютона-Рафсона

У рамках схеми Ньютона-Рафсона:

1) передбачається, що функція f двічі диференційована.

2) робота алгоритму починається в точці x_1, що являє початкове наближення (або початкову оцінку) координати стаціонарної точки, або кореня рівняння $f'(x) = 0$. Потім будується лінійна апроксимація функції $f'(x)$ в точці x_1, а точка, у якій апроксимуюча лінійна функція перетворюється в нуль, приймається як наступне наближення. Якщо точка x_k прийнята як поточне наближення до стаціонарної точки, то лінійна функція, що апроксимує функцію $f'(x)$ в точці x_k, записується у виді

$$\tilde{f}'(x; x_k) = f'(x_k) + f''(x_k)(x - x_k). \qquad (1)$$

У точці $x_k = x_{k+1}$ $\quad \tilde{f}'(x; x_k) = 0$

Прирівняємо праву частину рівняння (1) нулеві, одержимо наступне наближення:

$$x_{k+1} = x_k - \left[f'(x_k) / f''(x_k)\right].$$

Рис. 1 ілюструє основні кроки реалізації методу Ньютона-Рафсона.

[1] Математичні основи оптимізації телекомунікаційних систем: підручник. За загальною редакцією Захарченко М.В / Захарченко М.В., Горохов С.М., Балан М.М., Гаджієв М.М., Корчинський В.В., Ложковський А. Г. – Одеса: ОНАЗ і м. О.С. Попова, 2010. – 240 с.

В залежності від вибору початкової точки і виду функції алгоритм може як сходитися до істинної стаціонарної точки, так і розходитися, що показано на рис. 2. Якщо початкова точка розташована правіше x_0, то одержані в наслідок послідовних наближень точки віддаляються від стаціонарної точки x_1.

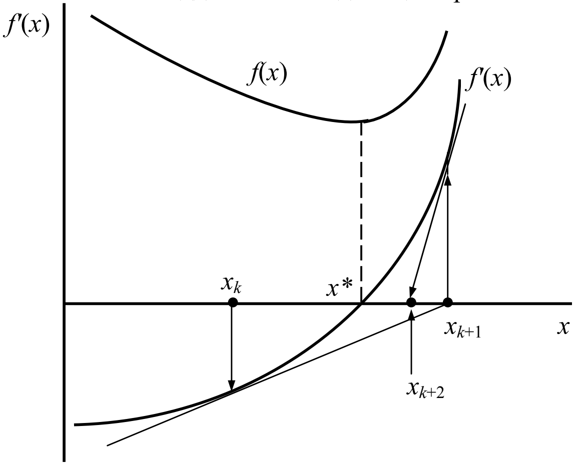

Рисунок 1 – Метод Ньютона-Рафсона (збіжність)

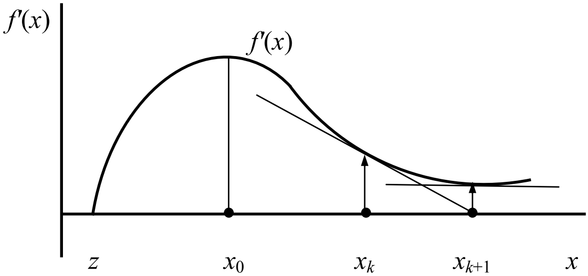

Рисунок 2 – Метод Ньютона-Рафсона (відсутність збіжності)

2 Метод середньої точки

Якщо функція $f(x)$ унімодальна у заданому інтервалі пошуку, то точкою екстремуму є точка, у якій $f'(x)=0$. Якщо при цьому є можливість обчислювати як значення функції, так і її похідну, то для знаходження кореня рівняння $f'(x)=0$ можна скористатися ефективним алгоритмом вилучення інтервалів, на кожній ітерації якого розглядається лише одна спробна точка. Наприклад, якщо в точці x_1 виконується нерівність $f'(x_1)>0$, то з урахуванням припущення про унімодальність природно стверджувати, що точка мінімуму не може знаходитися правіше x_1 і інтервал $x \ge x_1$ можна вилучити. Приведені міркування лежать в основі логічної структури методу середньої точки, що іноді називають *пошуком Больцано* [1].

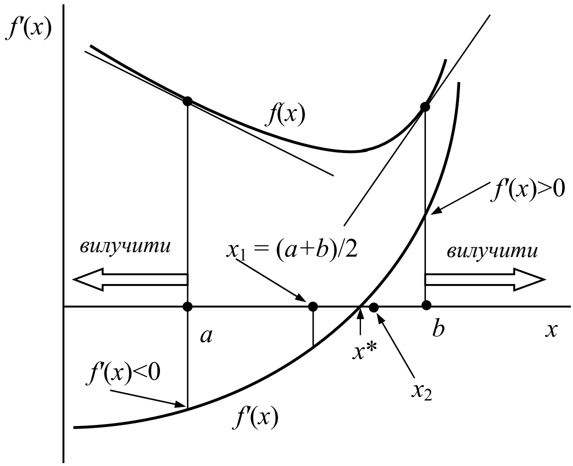

Рисунок 3 – Метод середньої точки

Визначимо дві точки a і b таким чином, що $f'(a) < 0$ і $f'(b) > 0$. Стаціонарна точка розташована між a і b. Обчислимо значення похідної функції в середній точці розглянутого інтервалу $x_1 = (a + b)/2$. Якщо б $f'(x_1) > 0$, то інтервал (x_1, b) слід вилучити з інтервалу пошуку. З іншого боку, якщо $f'(x_1) < 0$ (див. рис. 3), то можна вилучити інтервал (a, x_1). Нижче дається формалізований опис кроків алгоритму.

Нехай маємо обмежений інтервал $a \le x \le b$ і задану заздалегідь установлену величину припустимого відхилення – параметр ε.

Крок 1. Початковий інтервал (a, b); при цьому $f'(a) < 0$ і $f'(b) > 0$.

Крок 2. Обчислити $x_1 = (a + b)/2$ і $f'(x_1)$.

Крок 3. Якщо $|f'(x_1)| < \varepsilon$ – пошук закінчено, можна визначити точку x_1 точкою екстремуму. У протилежному випадку ($|f'(x_1)| > \varepsilon$), якщо $f'(x_1) > 0$, покласти $b = x_1$ і перейти до кроку 2. Якщо $f'(x_1) < 0$, покласти $a = x_1$ і перейти до кроку 2.

Слід зазначити, що логічна структура пошуку відповідно до викладеного методу вилучення інтервалів заснована лише на дослідженні знака похідної незалежно від значень, які ця похідна приймає. У наступному підрозділі розглядається метод січних, при реалізації якого враховуються як знак похідної, так і її значення.

3 Метод січних

Метод січних є комбінацією методу Ньютона-Рафсона і загальної схеми вилучення інтервалів та орієнтований на знаходження кореня рівняння $f(x) = 0$ у інтервалі (a, b), якщо, зрозуміло, такий корінь існує.

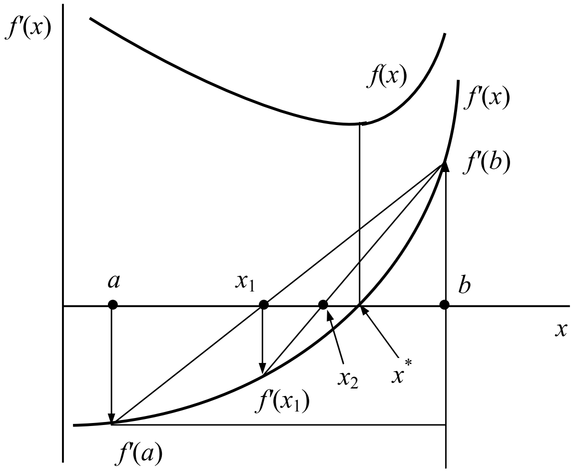

Рис. 4 Метод січних

Припустимо, що в процесі пошуку стаціонарної точки функції $f(x)$ в інтервалі (a, b) виявлені дві точки a і b, у яких знаки похідної різні. У цьому випадку алгоритм методу січних дозволяє апроксимувати функцію $f'(x)$ «січною прямою» (прямою лінією, що з'єднує дві точки) і знайти точку, у якій січна графіка $f'(x)$ перетинає вісь абсцис (рис. 4). Таким чином, наступне наближення до стаціонарної точки x^* визначається із співвідношення

$$\frac{b-a}{f'(b)-f'(a)} = \frac{b-x_1}{f'(b)}$$

за формулою

$$x_1 = b - \frac{(b-a)f'(b)}{f'(b)-f'(a)}.$$

Якщо $|f'(x_1)| \le \varepsilon$, пошук слід закінчити. У протилежному випадку необхідно вибрати одну з точок a або b таким чином, щоб знаки похідної в цій точці і точці x_1 були різні, а потім повторити основний крок алгоритму. Наприклад, у ситуації, що показана на рис. 4, у якості двох наступних точок мають бути обрані точки x_1 і b. Легко бачити, що на відміну від методу середньої точки метод січних заснований на дослідженні не тільки знака, але і значень похідної у спробних точках і тому в ряді випадків дозволяє вилучити більш половини інтервалу пошуку (див. рис. 4).