

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Запорізький національний технічний університет

МЕТОДИЧНІ ВКАЗІВКИ

до практичних занять з дисципліни
«Методи обробки багатовимірних даних»
для студентів спеціальності
121 «Інженерія програмного забезпечення»
всіх форм навчання

2017

Методичні вказівки до практичних занять з дисципліни
«Методи обробки багатовимірних даних» для студентів спеціальності
121 «Інженерія програмного забезпечення» всіх форм навчання /
Укладач: Л.Ю. Дейнега. – Запоріжжя: ЗНТУ, 2017. – 17 с.

Укладач: Л. Ю. Дейнега

Рецензент: В.І. Дубровін

Відповідальний за випуск: Л. Ю. Дейнега

Затверджено на
засіданні кафедри ПЗ
Протокол № 5 від 17.11.16

ЗМІСТ

ЗМІСТ	3
1 ЗАГАЛЬНІ ВІДОМОСТІ	4
2 МЕТОДИ БАГАТОВИМІРНОЇ ОПТИМІЗАЦІЇ	5
2.1 Метод Коші	5
2.2 Метод Ньютона	8
2.3 Метод Марквардта	8
2.4 Метод Флетчера – Рівса	10
2.5 Алгоритм Девідона - Флетчера - Пауелла	12
2.6 Контрольні питання	15
3 ПРАКТИЧНЕ ЗАВДАННЯ	16
ЛІТЕРАТУРА	17

1 ЗАГАЛЬНІ ВІДОМОСТІ

Переважна кількість реальних задач оптимізації, що являє практичний інтерес, є багатовимірними: у них цільова функція залежить від декількох аргументів, при цьому в деяких випадках їхня кількість може бути дуже великою.

Математична постановка таких завдань аналогічна їхній постановці в одновимірному випадку: проводиться пошук найменшого (найбільшого) значення цільової функції, заданої на деякій множині G можливих значень її аргументів.

Як і в одновимірному випадку, характер завдання і відповідно можливі методи вирішення істотно залежать від тієї інформації про цільову функцію, яка доступна в процесі її дослідження. В одних випадках цільова функція задається аналітичною формулою, будучи при цьому диференційованою функцією. Тоді можна обчислити її часткові похідні, отримати явний вираз для градієнта, що визначає в кожній точці напрямки зростання та спадання функції, і використати цю інформацію для вирішення завдання. В інших випадках ніякої формули для цільової функції немає, а є лише можливість визначити її значення в будь-якій точці даної області (за допомогою розрахунків, у результаті експерименту тощо).

Квазіньютонівські методи – методи оптимізації, засновані на накопиченні інформації про кривизну цільової функції із використанням спостережень за зміною градієнта. Цим вони принципово відрізняються від ньютонівських методів. Клас квазіньютонівських методів виключає явне формування матриці Гессе, замінюючи її деяким наближенням.

Мета практичних занять з дисципліни «Методи обробки багатовимірних даних» - порівняти на практиці багатовимірні методи оптимізації та зробити аналіз отриманих результатів.

2 МЕТОДИ БАГАТОВИМІРНОЇ ОПТИМІЗАЦІЇ

2.1 Метод Коші

Приклад використання методу Коші.

Розглянемо функцію $f(x) = 8x_1^2 + 4x_1x_2 + 5x_2^2$, використовуючи метод Коші для розв'язку задачі її мінімізації.

Рішення: Обчислимо компоненти градієнту:

$$\frac{\partial f}{\partial x_1} = 16x_1 + 4x_2; \quad \frac{\partial f}{\partial x_2} = 10x_2 + 4x_1.$$

Для того, щоб застосувати метод найшвидшого спуску, задамо початкове наближення $x^{[0]} = [10; 10]^T$ та за допомогою формули (1.5) побудуємо нове наближення.

$x^{(1)} = x^{(0)} - \alpha^{(0)} * \nabla f(x^{(0)})$. Оберемо $\alpha^{(0)}$ таким чином, щоб $f(x^{(1)}) \rightarrow \min$. $\alpha^{(0)} = 0.056$.

$$\nabla f(x^{(0)}) = [16x_1 + 4x_2; 10x_2 + 4x_1] = [200; 140]$$

$$\alpha^{(0)} \nabla f(x^{(0)}) = 0.056 * [200; 140] = [11.2; 7.84]$$

$$x^{(1)} = [10; 10] - [11.2; 7.84] = [-1.20; 2.16]$$

Далі знаходимо точку $x^{(2)} = x^{(1)} - \alpha^{(1)} * \nabla f(x^{(1)})$, обчислюємо градієнт в точці $x^{(1)}$ та проводимо пошук вздовж прямої

$$\nabla f(x^{(1)}) = [16x_1 + 4x_2; 10x_2 + 4x_1] = [-19.2 + 8.64; 21.6 - 4.8] = [-10.56; 16.8].$$

В таблиці 2.1 надані дані, отримані при проведенні ітерацій на основі одномірного пошуку методом кубічної апроксимації.

Таблиця 2.1 – Результати обчислень по методу Коші.

K	$X_1^{(k)}$	$X_2^{(k)}$	$f(x^{(k)})$
1	-1,2403	2,1181	24,2300
2	0,1441	0,1447	0,3540
3	-0,0181	0,0309	0,0052
4	0,0021	0,0021	0,0000

Не зважаючи на те, що метод Коші має відносно невелике практичне застосування, він реалізує найважливіші кроки більшості градієнтних методів.

Схема алгоритму Коші представлена на рисунку 2.1.

Робота алгоритму завершується, коли модуль градієнту або модуль вектора Δx стають достатньо малим.

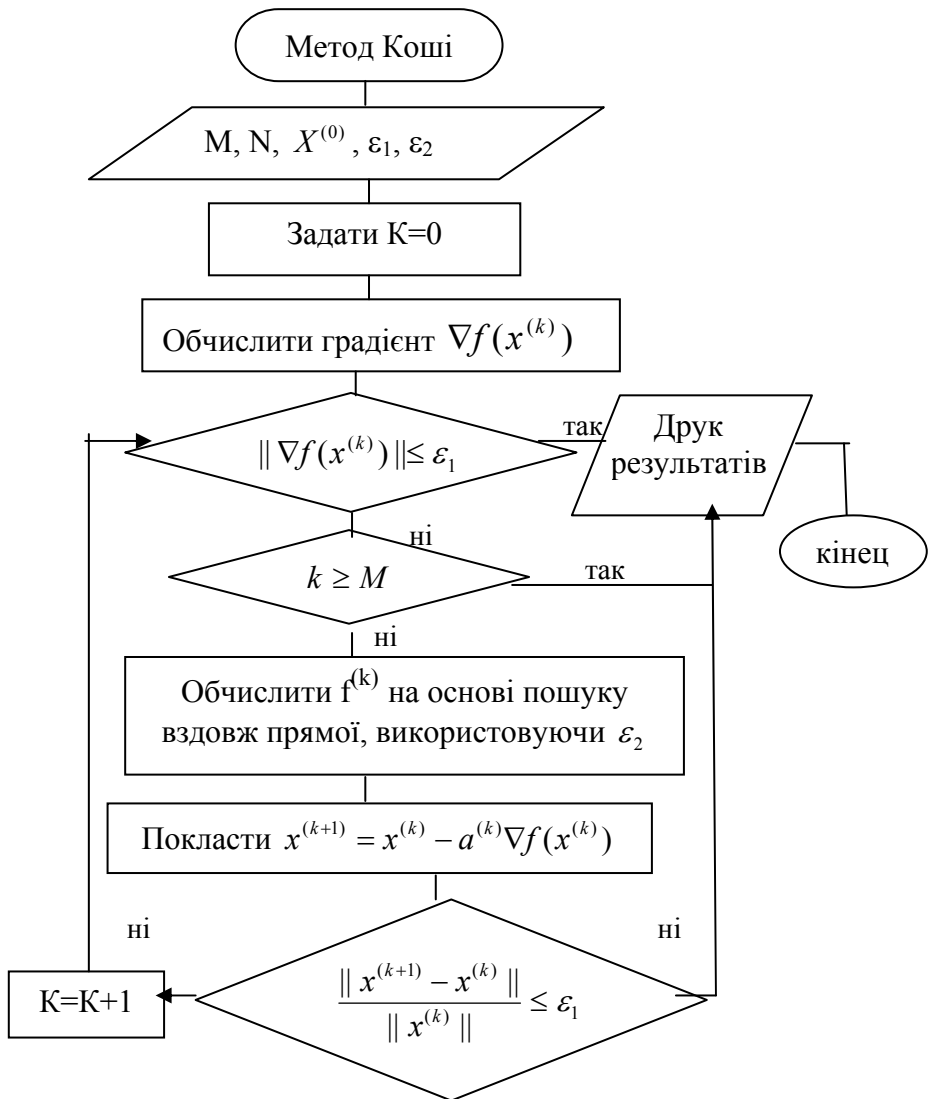


Рисунок 2.1 – Схема методу Коші

2.2 Метод Ньютона

У методі Коші застосовується найкраща локальна стратегія пошуку з використанням градієнту. Однак, рух у напрямку, протилежному градієнту, приводить до точки мінімуму лише в тому випадку, коли лінії рівня функції f являють собою кола. Таким чином, напрямок, протилежний градієнту, загалом кажучи, не може бути придатним глобальним напрямком пошуку точок оптимуму нелінійних функцій.

Послідовне використання схеми квадратичної апроксимації приводить до реалізації оптимізаційного методу Ньютона за формулою:

$$x^{(k+1)} = x^{(k)} - \nabla^2 f(x^{(k)})^{-1} \cdot \nabla f(x^{(k)})$$

Доведено, що метод Ньютона виявляє квадратичну швидкість збіжності, тобто виконується умова

$$\|\varepsilon^{(k+1)}\| \leq c \|\varepsilon^{(k)}\|^2,$$

де константа c зв'язана з обумовленістю матриці Гессе $\nabla^2 f$.

Метод Ньютона збігається всякий раз, коли вибір $x^{(0)}$ здійснюється згідно умові

$$\|\varepsilon x^0\| < \frac{1}{c}.$$

Квадратична швидкість збіжності пояснюється тією обставиною, що метод заснований на квадратичній апроксимації.

2.3 Метод Марквардта

Даний метод є комбінацією методів Коші та Ньютона, в якому добре сполучаються позитивні властивості обох методів. При

використанні методу Марквардта необхідна інформація про значення других похідних цільової функції.

Алгоритм Марквардта

Крок 1. Задати $x^{(0)}$ початкове наближення до x^* , M – максимально припустима кількість ітерацій, E – параметр збіжності.

Крок 2. Покласти $K = 0$, $\lambda^{(0)} = 10^4$.

Крок 3. Обчислити $\nabla f(x^{(K)})$

Крок 4. Перевірити чи виконується нерівність $\|\nabla f(x^{(K)})\| < E$.

Якщо так – перейти до кроку 11, ні – перейти до наступного кроку.

Крок 5. Перевірити чи виконується нерівність $K \geq M$.

Так - перейти до кроку 11, ні - перейти до наступного кроку.

Крок 6. Обчислити $S(x^{(K)})$ за формулою (1.16).

Крок 7. Покласти $x^{(K+1)} = x^{(K)} + S(x^{(K)})$.

Крок 8. Чи виконується нерівність $f(x^{(K+1)}) < f(x^{(K)})$.

Так - перейти до кроку 9, ні - перейти до кроку 10.

Крок 9. Покласти $\lambda^{(K+1)} = 1/2 \lambda^{(K)}$, $K = K + 1$. Перейти до кроку 3.

Крок 10. Покласти $\lambda^{(K+1)} = 2 \lambda^{(K)}$. Перейти до кроку 6.

Крок 11. Друк результатів і зупинка.

2.4 Метод Флетчера – Рівса

Метод дозволяє знайти мінімум нелінійної цільової функції багатьох змінних вигляду

$$M = F(x_1, x_2, \dots, x_N)$$

при відсутності обмежень. Метод заснований на застосуванні часткових похідних цільової функції по незалежним змінним і перевизначений для дослідження унімодальних функцій. За його допомогою можна досліджувати і мультимодальні функції, однак в цьому випадку слід брати декілька вхідних точок і перевіряти, чи в усіх випадках однакове рішення. Схема алгоритму метода Флетчера – Рівса представлена на рис. 2.2.

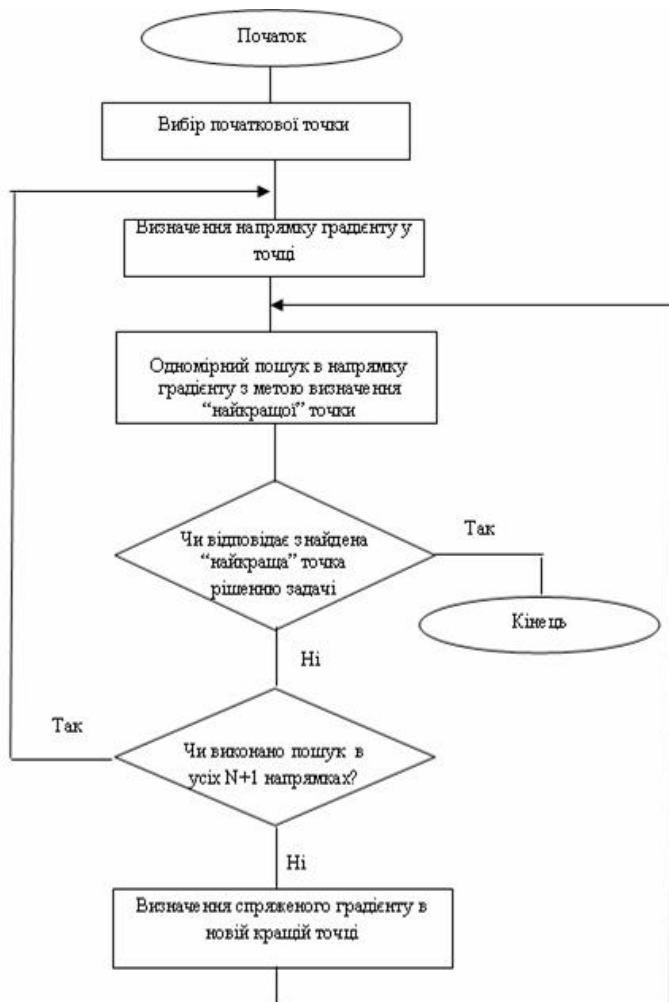


Рисунок 2.2 - Схема алгоритму метода Флетчера - Рівса

2.5 Алгоритм Девідона - Флетчера - Пауелла

Розглянемо алгоритм методу Девідона - Флетчера – Пауелла - алгоритм мінімізації функції декількох змінних.

Початковий етап:

Нехай $\varepsilon > 0$ - константа для зупинки. Вибрати точку x_1 і початкову симетричну позитивно визначену матрицю D_1 . Покласти $y_1 = x_1$, $k = j = 1$ і перейти до основного етапу.

Основний етап:

Крок 1. Якщо $\|\nabla f(y_j)\| < \varepsilon$, то зупинитися, інакше покласти $d_j = -D_j \nabla f(y_j)$ та взяти в якості λ_j оптимальне рішення задачі мінімізації $f(y_j + \lambda d_j)$ при $\lambda \geq 0$. Покласти $y_{j+1} = y_j + \lambda_j d_j$.

Якщо $j < n$, то перейти до кроку 2. Якщо $j = n$, то покласти $y_1 = x_{k+1} = y_{n+1}$, замінити k на $k+1$, покласти $j = 1$ і повторити крок 1.

Крок 2. Побудувати D_{j+1} наступним чином:

$$D_{j+1} = D_j + \frac{p_j p_j^T}{p_j^T p_j} - \frac{D_j q_j q_j^T D_j}{q_j^T D_j q_j},$$

де

$$p_j = \lambda_j d_j,$$

$$q_j = \nabla f(y_{j+1}) - \nabla f(y_j).$$

Замінити j на $j+1$ та перейти до кроку 1.

Приклад

Розглянемо наступну задачу: мінімізувати $(x_1 - 2)^4 + (x_1 - 2x_2)^2$

Результати обчислень методом Девідона - Флетчера - Пауелла наведені в таблиці 2.2.

Таблиця 2.2 - Результати обчислень методом Девідона - Флетчера - Пауелла

k	$\mathbf{x}_k$ $\mathbf{f}(\mathbf{x}_k)$	j	$\mathbf{y}_j$ $\mathbf{f}(\mathbf{y}_j)$	$\nabla \mathbf{f}(\mathbf{y}_j)$	$\ \nabla \mathbf{f}(\mathbf{y}_j)\$	$\mathbf{D}$	$\mathbf{d}_j$	λ_j	$\mathbf{y}_{j+1}$
1	(0.0, 3.0) (52.0)	1	(0.0, 3.0) (52.0)	(-44.0, 24.0)	50.12	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	(44, -24.)	0.0 6	(2.7, 1.5)
		2	(2.7, 1.5) (0.3)	(0.7, 1.3)	1.47	$\begin{bmatrix} 0.25 & 0.38 \\ 0.38 & 0.81 \end{bmatrix}$	(-0.6, -1.3)	0.2	(2.5, 1.2)
2	(2.5, 1.2) (0.103)	1	(2.5, 1.2) (0.10)	(0.9, -0.4)	0.99	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	(-0.9, 0.4)	0.1	(2.4, 1.2)
		2	(2.4, 1.3) (0.05)	(0.2, 0.4)	0.40	$\begin{bmatrix} 0.65 & 0.45 \\ 0.45 & 0.46 \end{bmatrix}$	(-0.25, - 0.3)	0.6	(2.2, 1.1)
3	(2.2, 1.1) (0.01)	1	(2.3, 1.1) (0.01)	(0.18, - 0.20)	0.27	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	(-0.18, 0.2)	0.1	(2.2, 1.1)
		2	(2.2, 1.1) (0.004)	(0.04, 0.04)	0.06	$\begin{bmatrix} 0.80 & 0.38 \\ 0.38 & 0.31 \end{bmatrix}$	(-0.05, 0.0)	2.6	(2.1, 1.0)
4	(2.1, 1.05) (0.0005)	1	(2.1, 1.0) (0.0005)	(0.05, - 0.08)	0.09	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	(-0.05, 0.08)	0.1	(2.1, 1.0)
		2	(2.1, 1.1) (0.00)	(0.04, 0.04)	0.006				

На кожній ітерації вектор d_j для $j = 1, 2$ визначається у вигляді - $D_j \nabla f(y_j)$, де D_1 - одинична матриця, а D_2 обчислюється за формулами (1) - (3). При $k = 1$ маємо $p_1 = (2.7, -1.49)T$, $q_1 = (44.73, -22.72)T$

На другій ітерації $p_1 = (-0.1, 0.05)T$, $q_1 = (-0.7, 0.8)T$ і, нарешті, на третій ітерації $p_1 = (-0.02, 0.02)T$, $q_1 = (-0.14, 0.24)T$, $q_1 = (-0.14, 0.24)T$.

Точка y_{j+1} обчислюється оптимізацією вздовж напрямку d_j при початковій точці y_j для $j = 1, 2$. Процедура зупинена в точці $y_2 = (2.115, 1.058)T$ на четвертій ітерації, тому що норма $\|f(y_2)\| = 0.006$ досить мала.

Траєкторія руху, отримана методом, показана на рис. 2.3.

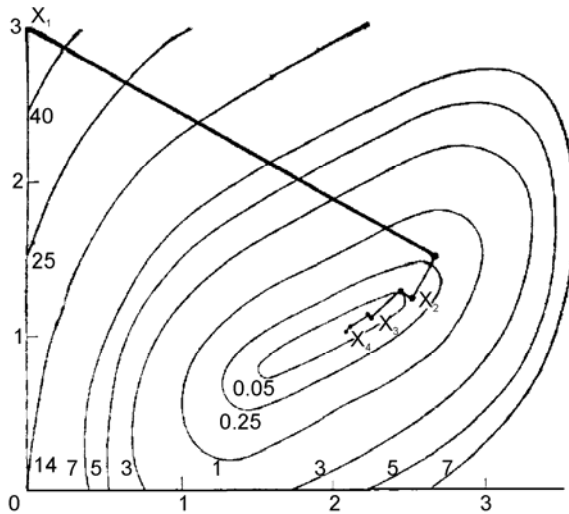


Рисунок 2.3 – Графічне відображення пошуку за методом Девідона - Флетчера - Пауелла.

2.6 Контрольні питання

- 2.6.1 Що називається градієнтом функції?
- 2.6.2 Дайте визначення $s(x^{(k)})$ в методах багатовимірної оптимізації.
- 2.6.3 Дайте визначення матриці Гессе.
- 2.6.4 В чому полягає суть методу Коші?
- 2.6.5 Відмінності методу Ньютона від методу Коші.
- 2.6.6 Наведіть аналіз властивостей збіжності методу Ньютона.
- 2.6.7 Відмінності методу Марквардта від методів Коші та Ньютона.
- 2.6.8 Чим характеризується метод Марквардта?
- 2.6.9 Алгоритм методу Марквардта.
- 2.6.10 В чому полягає суть методу Флетчера-Рівса?
- 2.6.11 В чому полягає суть методу Девідона-Флетчера-Пауелла?
- 2.6.12 Які критерії завершення пошуку використовуються в роботі?
- 2.6.13 Комбінація яких методів забезпечує знаходження найбільш точного значення?
- 2.6.14 На чому заснований метод Флетчера – Рівса?
- 2.6.15 Які функції можна досліджувати методом Флетчера - Рівса?
- 2.6.16 Алгоритм методу Флетчера – Рівса.

3 ПРАКТИЧНЕ ЗАВДАННЯ

Виконати порівняльний аналіз наведених методів багатовимірної оптимізації, в якості методів одновимірного пошуку використовувати методи поділу інтервалу навпіл, золотого перетину, кубічної апроксимації. В результаті порівняння повинні бути отримані наступні дані:

- значення функції на момент завершення пошуку;
- кількість обчислень значення функції;
- час роботи до завершення пошуку.

Використовувати наступні критерії закінчення пошуку:

- а) $K > M$, де K – поточна кількість ітерацій, M – максимальна кількість ітерацій;

б)
$$\frac{\|\Delta x\|}{\|x^{(k)}\|} \leq \varepsilon_1$$
;

в)
$$\|\nabla f(x^{(k)})\| \leq \varepsilon_1$$
;

г)
$$\nabla f(x^{(k)})^T S(x^{(k)}) > 0$$
;

д)
$$f(x^{(k+1)}) > f(x^{(k)})$$
.

В якості досліджуваної функції використовувати функцію Розенброка.

Результати порівняльного аналізу зведені в таблицю.

ЛІТЕРАТУРА

- 1 Банди Б. Методы оптимизации. Вводный курс/ Б. Банди; пер. с англ.. - М.: Радио и связь, 1988.- 128 с.
- 2 Гилл Ф. Практическая оптимизация/ Ф. Гилл, У. Мюррей, М. Райт. -М.: Мир, 1985.-509 с.
- 3 Дубровин В.И. Идентификация и оптимизация сложных технических процессов и объектов / В.И Дубровин. - Запорожье: ЗГТУ, 1997. - 92 с.
- 4 Дубровін В.І. Методи оптимізації та їх застосування в задачах навчання нейронних мереж: навчальний посібник/ В.И Дубровин, С.О. Субботін.- Запоріжжя: ЗНТУ, 2003.- 136 с.
- 5 Исследование операций: Пер. с англ.: В 2-х т.:/под ред. Дж. Моудера, С. Элмаграби. - М.: Мир, 1981. Т.1- 712с, Т.2- 678с.
- 6 Таха. Введение в исследование операций: в 2-х кн./ А. Хзмди; пер. с англ. - М: Мир, 1985. Кн.1- 479с, Кн.2- 496с.
- 7 Химмельблау Д.М. Прикладное нелинейное программирование/ Д.М. Химмельблау. - М. : Мир, 1975. – 534 с.