13. МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ

13.1. Загальна схема ітераційних методів безумовної оптимізації

Розглянемо нелінійну неперервну функцію $f(x)$ декількох змінних $x \in R^n$. Задача безумовної оптимізації полягає в знаходженні точки $x^* \in R^n$, в якій функція $f(x)$ приймає мінімальне значення. Стисло задача безумовної оптимізації записується у вигляді

$$f(x) \to \min_{x \in R^n} . \qquad (13.1)$$

У загальному випадку функція $f(x)$ може мати декілька локальних точок мінімуму. Чисельні методи розв'язання задачі багатовимірної оптимізації, що розглядаються нижче, дозволяють знаходити одну з локальних точок мінімуму (залежить від задання початкового наближення розв'язку x^0).

Розглянемо декілька чисельних методів пошуку точки мінімуму функції $f(x)$, які є ітераційними методами і в принципі мають одну загальну схему.

Узагальнена схема чисельних методів розв'язання задачі безумовної оптимізації

При пошуку точки мінімуму функції $f(x)$ чисельним методом задаються початкове наближення розв'язку $x^0 \in R^n$ і бажана точність $\varepsilon > 0$ розв'язку "по градієнту".

На k-й ($k = 0,1,2,\ldots$) ітерації методу маємо поточне наближення розв'язку $x^k \in R^n$. Наступне наближення розв'язку $x^{k+1} \in R^n$ обчислюється за формулою

$$x^{k+1} = x^k + \alpha_k h^k , \qquad (13.2)$$

де **напрям пошуку** $h^k \in R^n$ визначається конкретним чисельним методом, а **кроковий множник** $\alpha_k > 0$ ($\alpha_k \in R^1$) зазвичай визначається одним із способів:

а) як точка мінімуму функції однієї змінної $\phi_k(\alpha) = f(x^k + \alpha\, h^k)$, $\alpha > 0$ (повний пошук по напряму);

б) за алгоритмом дроблення кроку (описаний нижче).

Обчислення продовжується до тих пір, поки не виконається нерівність $\|f'(x^k)\| \le \varepsilon$ (**критерій останову "по градієнту"**).

Зазначимо, що вибір крокового множника α_k за способом а) (повний пошук по напряму) можна проводити одним з методів одновимірної мінімізації (див. розділ 9.4), але ефективніше при цьому скористатися методом золотого перерізу.

Так само треба додати, що незалежно від того, яким способом проводиться вибір крокового множника α_k: за способом а) чи б), виконується таке співвідношення

$$f(x^{k+1}) < f(x^k) \text{ для всіх } k = 0,1,2,\ldots$$

Виконання цієї умови на кожній ітерації методу називають **релаксацією**, тобто значення цільової функції на кожній ітерації зменшується.

Зазначимо, що критерій останову "по градієнту" є сенс використовувати тільки в методах, в яких на кожній ітерації обчислюється градієнт $f'(x)$ функції $f(x)$ у поточній точці x^k. В противному випадку (тобто в методах 0-го порядку) краще скористатися іншим критерієм останову, наприклад "**по аргументу**" ($\|x^{k+1} - x^k\| \le \varepsilon$).

Графічна інтерпретація процесу пошуку точки безумовного мінімуму за схемою (13.2) наведена на рис. 11.1.

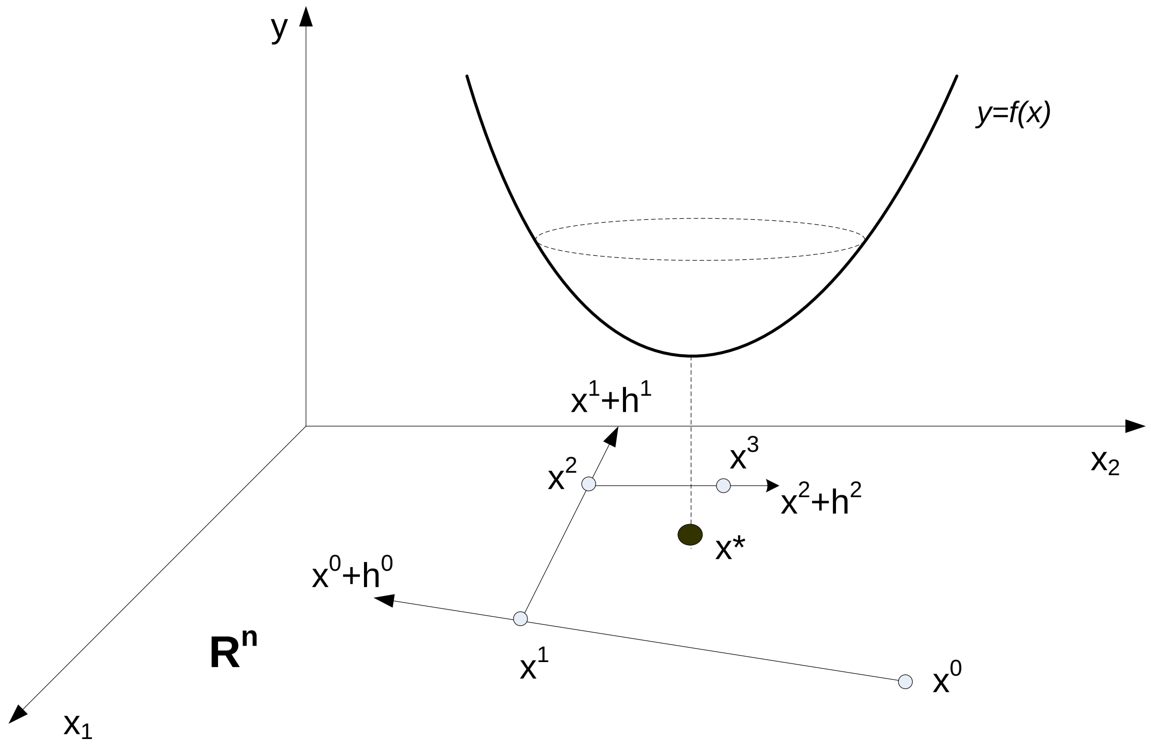

Рис. 13.1. **Пошук точки безумовного мінімуму за схемою (13.2)**

13.2. Алгоритм дроблення кроку для визначення крокового множника

Алгоритм дроблення кроку призначений для визначення крокового множника в чисельних методах 1-го порядку розв'язання задачі безумовної оптимізації, тобто передбачається, що функція $f(x)$ диференційована на R^n.

В алгоритмі дроблення кроку при визначенні крокового множника α_k в чисельних методах розв'язання задачі безумовної оптимізації, що підпадають під схему (13.2), задаються наступні параметри:

а) максимальне значення кроку, $\rho > 0$;

б) коефіцієнт рівня убування функції, $\gamma \in (0, \frac{1}{2})$;

в) коефіцієнт дроблення кроку, $\lambda \in (0,1)$.

При цьому вектор пошуку h^k повинен бути напрямом убування функції $f(x)$ з точки x^k, тобто він повинен задовольняти умові $\left(f'(x^k), h^k\right) < 0$ [28].

Спочатку вибирається $\alpha = \rho$ і перевіряється нерівність

$$f(x^k+\alpha\, h^k)\le f(x^k)+\alpha\,\gamma\big(f'(x^k),h^k\big). \qquad (13.3)$$

Якщо нерівність (13.3) не виконується, то α зменшується шляхом множення на λ, тобто $\alpha=\lambda\times\alpha$, і знову перевіряється (13.3). Якщо нерівність виконується, то кроковий множник α_k береться рівним поточному значенню α.

Зазначимо, якщо вектор h^k є напрямом убування функції $f(x)$ (тобто виконується умова $\big(f'(x^k),h^k\big)<0$), то описаний процес дроблення кроку буде скінченним, тобто нерівність (13.3) виконається через скінчене число кроків [28]. Очевидно, що при цьому виконуватиметься умова релаксації $f(x^{k+1})<f(x^k)$.

13.3. Градієнтні методи

Градієнтними називаються чисельні методи розв'язання задач безумовної оптимізації (13.1), що підпадають під схему (13.2), в яких напрям пошуку h^k визначається за загальною формулою

$$h^k=-B_k f'(x^k), \qquad (13.4)$$

де B_k – деяка додатньо визначена матриця $n\times n$, тобто $B_k>0$ [28]. Як представника градієнтних методів розглянемо метод найшвидшого спуску.

Ідея методу найшвидшого спуску (як і градієнтних методів в цілому) полягає в тому, що градієнт $f'(x^k)$ функції $f(x)$, обчислений в деякій точці x^k, указує в просторі R^n напрям (з точки x^k) найшвидшого (найбільш крутого) убування функції $f(x)$. Тому пошук наступного наближення x^{k+1} логічно проводити в цьому напрямі.

Метод найшвидшого спуску будує ітераційну послідовність $\{x^k\}$, $k=0,1,2,..$ наближень розв'язку $x^*\in R^n$ за схемою (13.2), де напрям пошуку h^k визначається за формулою

$$h^k=-f'(x^k), \qquad (13.5)$$

(тобто в (13.4) $B_k = E$), кроковий множник $\alpha_k > 0$ визначається як точка мінімуму функції однієї змінної $\phi_k(\alpha) = f(x^k - \alpha f'(x^k))$ (перший спосіб – повний пошук по напряму). Обчислення продовжується до тих пір, поки не виконається умова $\|f'(x^k)\| \le \varepsilon$.

Метод найшвидшого спуску є методом 1-го порядку, оскільки в ньому використовуються значення першої похідної цільової функції $f(x)$.

Недоліком методу найшвидшого спуску є його достатньо низька швидкість (лінійна) збіжності, навіть для функцій "не яружних" (добре обумовлених). Для "яружних" же функцій метод скочується на дно "яру" і дуже повільно рухається до розв'язку зигзагами упоперек дна яру (практично "застряє" там).

13.4. Метод спряжених градієнтів

Метод спряжених градієнтів є модифікацією градієнтних методів, спрямованою на збільшення швидкості збіжності при пошуку точок мінімуму "яружних" функцій.

Метод будує ітераційну послідовність $\{x^k\}$, $k = 0,1,2,..$ наближень розв'язку $x^* \in R^n$ за схемою (13.2), де напрям пошуку h^k визначається за формулою:

$$h^k = -f'(x^k) \text{ при } k \in \{0, n, 2n, 3n, ...\}; \tag{13.6}$$

$$h^k = -f'(x^k) + \beta_{k-1} h^{k-1} \text{ при } k \notin \{0, n, 2n, 3n, ...\}. \tag{13.7}$$

Коефіцієнт β_{k-1} визначається за однією з формул:

$$\text{а) } \beta_{k-1} = \frac{\left(f'(x^k), f'(x^k) - f'(x^{k-1})\right)}{\|f'(x^{k-1})\|^2}; \tag{13.8}$$

$$\text{б) } \beta_{k-1} = \frac{\|f'(x^k)\|^2}{\|f'(x^{k-1})\|^2}. \tag{13.9}$$

Кроковий множник $\alpha_k > 0$ у (13.2) може бути визначений одним із способів:

а) як точка мінімуму функції однієї змінної $\phi_k(\alpha) = f(x^k + \alpha\, h^k)$;

б) за алгоритмом дроблення кроку (13.3).

Обчислення продовжується до тих пір, поки не виконається умова $\left\| f'(x^k) \right\| \le \varepsilon$.

Відмінність методу спряжених градієнтів від методу найшвидшого спуску полягає в тому, що після того, як був зроблений крок по напряму антиградієнта (13.6), наступні $(n-1)$ кроків виконуються в напрямі, спряженому (майже ортогональному) попередньому напряму пошуку. Саме цим обумовлено визначення значення коефіцієнта β_{k-1}, обчислюване за формулами (13.8) або (13.9). Тим самим просування до розв'язку уздовж дна "яру" досягається швидше, ніж у методу найшвидшого спуску.

Метод спряжених градієнтів є методом 1-го порядку, оскільки в ньому використовуються значення першої похідної цільової функції $f(x)$.

Достоїнством методу спряжених градієнтів є його висока швидкість збіжності для нормальних ("не яружних") функцій.

Недоліком методу спряжених градієнтів є його не дуже висока (хоча і цілком допустима) швидкість збіжності для "яружних" (погано обумовлених) функцій .

13.5. Метод Ньютона

Метод Ньютона також відноситься до класу градієнтних методів.

Метод будує ітераційну послідовність $\{x^k\}$, $k = 0,1,2,..$ наближень розв'язку $x^* \in R^n$ за схемою (13.2), де напрям пошуку h^k визначається за формулою

$$h^k = -[f''(x^k)]^{-1} f'(x^k) \qquad (13.10)$$

(тобто в (13.4) $B_k = [f''(x^k)]^{-1}$), кроковий множник $\alpha_k > 0$ у (13.2) завжди (**у класичному методі Ньютона**) береться рівним 1.

Ідея методу Ньютона полягає в тому, що на k-й ітерації (x^k – поточне наближення розв'язку) наступне наближення розв'язку x^{k+1} шукається, як точка мінімуму функції

$$f_k(x) \equiv f(x^k) + f'(x^k)(x - x^k) + \frac{1}{2} f''(x^k)(x - x^k) \qquad (13.11)$$

(розкладання в ряд Тейлора [28], тобто функція $f_k(x)$ є квадратичним наближенням цільової функції $f(x)$ в околі точки x^k). Але точка мінімуму функції $f_k(x)$ з (13.11) може бути знайдена аналітично з необхідної умови мінімуму $f_k'(x) = 0$:

$$f_k'(x) = f'(x^k) + f''(x^k)(x - x^k) = 0$$

або

$$f''(x^k)(x - x^k) = -f'(x^k).$$

Тому (через обернену матрицю)

$$(x - x^k) = -[f''(x^k)]^{-1} f'(x^k),$$

звідки і отримуємо точку мінімуму функції (13.11)

$$x = x^k - [f''(x^k)]^{-1} f'(x^k).$$

Таким чином, напрям пошуку h^k для обчислення наступного наближення x^{k+1} визначається за формулою (13.10), а кроковий множник α_k в (13.2) дорівнює 1.

Якщо послідовність $\{x^k\}$, $k = 0,1,2,.$, побудована згідно (13.2), збігається, то за достатньо загальних умов [28] швидкість її збіжності буде квадратичною, тобто $\left\| x^{k+1} - x^* \right\| \le M \left\| x^k - x^* \right\|^2$ починаючи з деякого k, де M – деяка додатня константа.

Зазначимо, що метод Ньютона є методом 2-го порядку, оскільки в ньому використовуються значення першої і другої похідної цільової функції $f(x)$.

Основними недоліками методу Ньютона є:

збіжність тільки для достатньо близьких до розв'язку початкових наближень x^0;

висока трудомісткість методу, оскільки на кожній ітерації необхідно обчислювати матриці $f''(x^k)$ і $\left[f''(x^k)\right]^{-1}$.

Основним достоїнствам методу Ньютона є те, що він має високу (квадратичну) швидкість збіжності навіть для функцій "ярів".

Зазначимо, що при практичній реалізації методу Ньютона вектор h^k ефективніше обчислювати не за формулою (13.10), а як розв'язок системи лінійних рівнянь виду $F'(x^k) \times h = -F(x^k)$.

Метод Ньютона з регулюванням кроку

Для усунення першого недоліку методу Ньютона (див. вище) на практиці зазвичай використовують його варіант, який носить назву метод Ньютона з регулюванням кроку. Суть його полягає в тому, що кроковий множник α_k не фіксується, а визначається за алгоритмом дроблення кроку (13.3), в якому його початкове значення $\rho = 1$.

13.6. Квазі-ньютонівські методи

Як наголошувалося вище, одним з недоліків методу Ньютона є його висока трудомісткість, зважаючи на необхідність обчислення на кожній ітерації матриць $f''(x^k)$ і $\left[f''(x^k)\right]^{-1}$. Для усунення цього недоліку різними авторами були запропоновані модифікації методу, засновані на послідовному (рекурентному) наближенні матриці $f''(x^k)$ або $\left[f''(x^k)\right]^{-1}$ в процесі обчислень [3; 11; 28]. Сімейство цих методів носить назву **квазі-ньютонівські методи.**

Ідея квазі-ньютонівських методів полягає в апроксимації матриць $f''(x^k)$ або $\left[f''(x^k)\right]^{-1}$ на основі інформації про значення градієнтів $f'(x^k)$, $f'(x^{k-1})$, вже обчислених в точках x^k і x^{k-1}.

Квазі-ньютонівські методи також відносяться до класу градієнтних методів.

Методи будують ітераційну послідовність $\{x^k\}$, $k = 0,1,2,..$ наближень розв'язку $x^* \in R^n$ за схемою (13.2), де напрям пошуку h^k **визначається за формулою**

$$h^k = -H_k f'(x^k), \tag{13.12}$$

(тобто в (13.4) $B_k = H_k$), кроковий множник $\alpha_k > 0$ в (13.2) визначається так само, як в методі Ньютона з регулюванням кроку.

Розглянемо варіанти квазі-ньютонівських методів, в яких матриця H_k вибирається так, щоб вона апроксимувала матрицю $[f''(x^k)]^{-1}$. Зазначимо, що з розкладання в ряд Тейлора [28] для векторної функції $f'(x)$ в околі точки x^{k-1}

$$f'(x) \approx f'(x^{k-1}) + f''(x^{k-1})(x - x^{k-1}),$$

а значить,

$$f'(x^k) - f'(x^{k-1}) \approx f''(x^{k-1})(x^k - x^{k-1}).$$

Введемо позначення

$$\Delta y^k = f'(x^k) - f'(x^{k-1}),\ \Delta x^k = x^k - x^{k-1}.$$

Тоді природно зажадати, щоб матриця H_k задовольняла умові

$$\Delta y^k = H_k \Delta x^k. \tag{13.13}$$

Визначення. Умова (13.13) називається квазі-ньютонівською.

У квазі-ньютонівських методах на k-й ітерації рекурентно перераховуються матриці H_k за формулою

$$H_{k+1} = H_k + \Delta H_k, \tag{13.14}$$

де добавка ΔH_k у різних варіантах методів обчислюється за різними формулами, але загальне в них те, що забезпечується виконання умови (13.13).

Наведемо найбільш відомі варіанти [3; 11; 28]:

$$\text{а) } \Delta H_k = \frac{\left(\Delta x^k - H_k \Delta y^k\right)\left(\Delta x^k - H_k \Delta y^k\right)^T}{\left\langle \Delta x^k - H_k \Delta y^k, \Delta y^k \right\rangle};$$

б) (метод Девідона – Флетчера – Пауелла)

$$\Delta H_k = \frac{\left(\Delta x^k\right)\left(\Delta x^k\right)^T}{\left\langle \Delta x^k, \Delta y^k \right\rangle} - \frac{\left(H_k \Delta y^k\right)\left(H_k \Delta y^k\right)^T}{\left\langle H_k \Delta y^k, \Delta y^k \right\rangle}.$$

Розглянемо квадратичну цільову функцію

$$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle,$$

де A – деяка квадратна симетрична матриця $b \in R^n$. Для таких функцій квазі-ньютонівські методи генерують послідовність $\{x^k\}$, що збігається до точки мінімуму x^* при $k \to \infty$, причому $H_k \to \left[f''(x^*)\right]^{-1}$.

Квазі-ньютонівські методи мають надлінійну швидкість збіжності. Це дуже ефективні методи розв'язання задач безумовної оптимізації. Методи не дуже трудомісткі, хоча в процесі обчислень необхідно зберігати дві матриці H_k, H_{k-1}, що при великій розмірності задачі (числі n змінних x) може утруднити його застосування.

13.7. Метод Гауса – Ньютона для задачі найменших квадратів

Розглянемо випадок, коли цільова функція в задачі безумовної оптимізації має вигляд

$$\Phi(x)=\frac{1}{2}\sum_{i=1}^{m} f_i^2(x)=\frac{1}{2}\|F(x)\|^2, \qquad (13.15)$$

де $f_i(x)$, $i=\overline{1,m}$ – деякі нелінійні функції від $x\in R^n$, $F(x)=\begin{pmatrix} f_1(x) \\ \cdots \\ f_m(x) \end{pmatrix}$ – векторна функція від $x\in R^n$ розмірності m.

Таким чином, розв'язується задача

$$\Phi(x)\to \min_{x\in R^n}, \qquad (13.16)$$

яка носить назву "**Задача найменших квадратів"**. Як правило в задачі (13.15), на практиці $m\geq n$. З цією задачею ми вже стикалися в розділі 7.4.

Метод Гауса – Ньютона будує ітераційну послідовність $\{x^k\}$, $k=0,1,2,\ldots$ наближень розв'язку $x^*\in R^n$ за схемою (13.2), де напрям пошуку h^k визначається за формулою

$$h^k=\left\{[F'(x^k)]^T[F'(x^k)]\right\}^{-1}[F'(x^k)]^T[F'(x^k)], \qquad (13.17)$$

а кроковий множник $\alpha_k>0$ в (13.2) визначається за алгоритмом дроблення кроку (13.3), в якому $\rho=1$.

Ідея методу Гауса – Ньютона полягає в тому, що на k-й ітерації (x^k – поточне наближення розв'язку) наступне наближення розв'язку x^{k+1} шукається, як точка мінімуму функції

$$\Phi_k(x)=\frac{1}{2}\|F_k(x)\|^2. \qquad (13.18)$$

де $F_k(x)\equiv F(x^k)+F'(x^k)(x-x^k)$ (розкладання в ряд Тейлора [28] для векторної функції $F(x)$).

Але точка мінімуму функції $\Phi_k(x) = \dfrac{1}{2}\|F_k(x)\|^2$ може бути знайдена аналітично з необхідної умови мінімуму $\Phi_k'(x) = 0$. А саме:

$$\Phi_k'(x) = [F'(x^k)]^T \times [F(x^k) + F'(x^k)(x - x^k)] = 0,$$

звідки

$$[F'(x^k)]^T F'(x^k)(x - x^k) = -[F'(x^k)]^T F(x^k).$$

Тому (через обернену матрицю)

$$(x - x^k) = -\left\{[F'(x^k)]^T F'(x^k)\right\}^{-1} [F'(x^k)]^T F(x^k),$$

звідки і отримуємо точку мінімуму функції (11.15)

$$x = x^k - \left\{[F'(x^k)]^T [F'(x^k)]\right\}^{-1} [F'(x^k)]^T [F'(x^k)].$$

Таким чином, напрям пошуку h^k для обчислення наступного наближення x^{k+1} визначається за формулою (13.17), а в алгоритмі дроблення кроку (13.3) початкове значення крокового множника α_k логічно брати рівним 1.

Метод Гауса – Ньютона для розв'язання задачі найменших квадратів (13.15) та (13.16) займає проміжне місце між методом спряжених градієнтів і методом Ньютона, як за швидкістю збіжності, так і за трудомісткістю. Тому саме його частіше всього і застосовують на практиці при розв'язання задачі найменших квадратів.

13.8. Методи випадкового пошуку глобального мінімуму

Описані вище чисельні методи безумовної оптимізації, по-перше, є методами 1-го порядку і вище, а по-друге, дозволяють визначити тільки точку локального мінімуму. Пошук глобальної точки мінімуму, в цьому випадку, вимагає додаткових процедур вибору початкових точок для багатократного пошуку з них [28].

Існує клас методів **глобальної оптимізації**, що носять загальну назву "**Методи випадкового пошуку**" [23]. Вони є методами 0-го порядку, тобто використовують тільки обчислення значень цільової функції $f(x)$.

Розглянемо алгоритм простого методу випадкового пошуку.

Метод також будує ітераційну послідовність $\{x^k\}$, $k = 0,1,2,...$ наближень розв'язку $x^* \in R^n$ за схемою (13.2), проте напрям пошуку h^k на k-й ітерації визначається не детермінованим чином, як в описаних вище методах (тобто за деякими формулами), а випадково.

Наприклад, випадковим чином згідно з рівномірним розподілом визначається точка $\xi \in R^n$ на кулі радіусу $\rho_k > 0$ до тих пір, поки не виконається нерівність $f(x^k + \xi) < f(x^k)$, тобто напрям ξ виявився вдалим. Тоді вектор h^k береться рівним ξ. Потім проводиться спроба поліпшити пошук у напрямі h^k, тобто визначається $\alpha_k > 0$ (для (13.2)), наприклад, за алгоритмом повного одновимірного пошуку (як точка мінімуму функції однієї змінної $\phi_k(\alpha) = f(x^k + \alpha\, h^k)$). Хоча повний пошук тут, по суті, буде неефективним, а краще застосовувати який-небудь неповний пошук [16], але який не буде використовувати обчислення похідних.

Тоді наступне наближення розв'язку обчислюється як $x^{k+1} = x^k + \alpha_k h^k$.

Зазначимо, що при реалізації методів випадкового пошуку виникають проблеми з тим, що вони не гарантують успіху при пошуку точки мінімуму, і тому критерій останову в них визначається евристичним шляхом. Наприклад, при виборі напряму пошуку h^k, згідно з описаною вище процедурою, багато що залежить від значення радіусу кулі ρ_k. Якщо декілька спроб виявилися невдалими, то його значення потрібно зменшувати. Інакше величину ρ_k можна збільшувати для прискорення процесу пошуку. Процес обчислень можна закінчувати, якщо ρ_k стає дуже малим.

Існує багато версій методу випадкового пошуку, наприклад [23]: випадковий пошук з поверненням; випадковий пошук з адаптацією; пошук з покаранням випадковістю і т. п., мета яких підвищити ефектив-ність процедури пошуку глобального мінімуму. Проте відзначимо, що універсальних підходів немає, а багато що залежить від властивостей цільової функції. Так само потрібно зазначити, що методи випадкового пошуку достатньо просто розповсюджуються на випадок задач умовної оптимізації [4; 23].

Значення методів випадкового пошуку для глобального розв'язання задач оптимізації особливо зростає при збільшенні продуктивності ЕОМ, оскільки вони легко пристосовуються для використання в паралельних комп'ютерних середовищах. Це дозволяє припустити, що в майбутньому саме методи випадкового пошуку можуть стати основними процедурами розв'язання задач нелінійного програмування.

13.9. Висновки

1. Задача пошуку точки мінімуму функції однієї змінної є допоміжною при розв'язанні задач безумовної оптимізації.
2. Методи випадкового пошуку для глобального розв'язання задач оптимізації особливо ефективні при проведенні обчислень в паралельних комп'ютерних мережах.
3. Методи розв'язання задачі безумовної оптимізації в більшості випадків мають загальну схему, що нерідко використовується при розробці діалогових систем оптимізації.
4. Одним з найефективніших методів розв'язання задач безумовної оптимізації є метод спряжених градієнтів.

13.10. Контрольні запитання та завдання

1. Наведіть загальну схема ітераційних методів безумовної оптимізації.
2. У чому полягають способи визначення крокового множника в чисельних методах розв'язання задачі безумовної оптимізації?
3. Наведіть алгоритм дроблення кроку для визначення крокового множника.
4. Що таке градієнт функції в точці та яку властивість він має?
5. Які чисельні методи розв'язання задачі безумовної оптимізації називаються градієнтними?
6. У чому полягає метод найшвидшого спуску для розв'язання задачі безумовної оптимізації? Вкажіть його достоїнства та недоліки.
7. У чому полягає метод спряжених градієнтів для розв'язання задачі безумовної оптимізації? Вкажіть його достоїнства та недоліки.

8. У чому полягає метод Ньютона для розв'язання задачі безумовної оптимізації? Вкажіть його достоїнства та недоліки. Які модифікації методу Ньютона ви знаєте?

9. У чому полягають квазі-ньютонівські методи для розв'язання задачі безумовної оптимізації?

10. У чому полягає метод Гауса – Ньютона для задачі найменших квадратів?

11. Які є методи пошуку глобальної точки мінімуму?

12. Знайдіть точку мінімуму функції кількох змінних $f(x) = 5x_1^2 + 4x_1x_2 + x_2^2 - 16x_1 - 12x_2$ методами найшвидшого спуску, спряжених градієнтів і Ньютона. Початкове наближення $x^0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Точність розв'язку $\varepsilon = 10^{-4}$. Порівняйте трудомісткість і швидкість збіжності методів.