ЧИСЕЛЬНІ МЕТОДИ БЕЗУМОВНОЇ ОПТИМІЗАЦІЇ

Лекція

Мета: Познайомитись з ітераційними методами розв`язання задач безумовної оптимізації.

Постановка задачі

Розв'язати задачу безумовної оптимізації:

$$f\left(x\right) \to min,\ x \in E^n. \tag{5.1}$$

1. Знайти точку мінімуму функції $f\left(x\right)$ класичним методом.
2. Зробити кілька кроків (не менше двох) методом найшвидшого спуску з розв'язанням задачі одновимірної оптимізації класичним методом.
3. Розробити програму знаходження оптимального розв'язку задачі безумовної оптимізації градієнтним методом з дробленням кроку. Застосувати її для знаходження оптимального розв'язку для заданих індивідуальним варіантом функцій із заданою точністю ε.
4. Розробити програму знаходження оптимального розв'язку задачі безумовної оптимізації методом Ньютона. Застосувати її для знаходження оптимального розв'язку для заданих індивідуальним варіантом функцій із заданою точністю ε.
5. Виконати геометричну інтерпретацію отриманих результатів за трьома методами. Для цього побудувати на площині лінії рівня, траекторії наближення до точки мінімуму.

Теоретичні відомості

Загальна схема ітераційних методів для розв'язання задачі безумовної мінімізації має вигляд

$$x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)},\ k = 0,1,2,\ldots, \tag{5.2}$$

де $p^{(k)}$ – напрямок спадання функції $f\left(x\right)$ (напрямок спуску) в точці $x^{(k)}$, α_k – параметр, який регулює довжину кроку вздовж $p^{(k)}$.

Методи монотонного спуску

$$f\left(x^{(k+1)}\right) < f\left(x^{(k)}\right) \tag{5.3}$$

називають *релаксаційними* методами. Відповідна послідовність $\left\{x^{(k)}\right\}_0^\infty$ називається також релаксаційною. Якщо функція $f\left(x\right)$ диференційована в точці $x^{(k)} \in E^n$, то релаксаційність методу (5.1) забезпечується тоді, коли напрямок $p^{(k)}$ складає нетупий кут з напрямком градієнта $f'\left(x^{(k)}\right)$, тобто $\left(f'\left(x^{(k)}\right), p^{(k)}\right) \geq 0$.

Основна властивість градієнту. Нехай $f(x)$ диференційовна в точці $x^{(k)}$ і $f'\left(x^{(k)}\right) \neq 0$. Напрямок найшвидшого зростання функції $f(x)$ в точці $x^{(k)}$ співпадає з напрямком градієнта $f'\left(x^{(k)}\right)$ у цій точці, а напрямок *найшвидшого спадання* – з напрямком антиградієнта $\left(-f'\left(x^{(k)}\right)\right)$.

Критерії закінчення ітераційного процесу (5.2)

На практиці часто використовуються такі умови закінчення розрахунків:

$$\left\| x^{(k+1)} - x^{(k)} \right\| \leq \varepsilon_1, \tag{5.4}$$

$$\left| f\left(x^{(k+1)}\right) - f\left(x^{(k)}\right) \right| \leq \varepsilon_2, \tag{5.5}$$

$$\left\| f'\left(x^{(k+1)}\right) \right\| \leq \varepsilon_3. \tag{5.6}$$

До початку обчислень вибирається одна з умов (5.4)-(5.6) і відповідне їй мале довільне число $\varepsilon_i > 0$, $i = \overline{1,3}$. Обчислення закінчуються після $(k+1)$-го кроку, якщо вперше виявляється виконаною обрана умова зупинки. На практиці також використовуються критерії, які складаються в одночасному виконанні двох з умов (5.4)-(5.6) або всіх трьох.

Градієнтні методи

Постановка задачі і умова застосування градієнтних методів:

$$f(x) \to min,\ x \in E^n,\ f(x) \in C^1\left(E^n\right). \tag{5.7}$$

У градієнтних методах за напрямок спуску $p^{(k)}$ з точки $x^{(k)}$ вибирається антиградієнт функції $f(x)$ у точці $x^{(k)}$, тобто

$$x^{(k+1)} = x^{(k)} - \alpha_k f'\left(x^{(k)}\right),\ k = 0, 1, 2, \ldots, \tag{5.8}$$

або в координатній формі

$$x_i^{(k+1)} = x_i^{(k)} - \alpha_k \frac{\partial f\left(x^{(k)}\right)}{\partial x_i},\ i = 1, 2, \ldots, n,\ k = 0, 1, 2, \ldots.$$

Різні способи вибору величини α_k ($\alpha_k > 0$) у методі (5.8) визначають різні варіанти градієнтних методів.

1. Метод найшвидшого спуску

На промені $\left\{ x \in E^n : x = x^{(k)} - \alpha f'\left(x^{(k)}\right),\ \alpha \geq 0 \right\}$, який направлений за антиградієнтом з точки $x^{(k)}$, введемо функцію $g_k(\alpha) = f\left(x^{(k)} - \alpha f'\left(x^{(k)}\right)\right)$, $\alpha \geq 0$ і визначимо α_k з умови

$$g_k\left(\alpha_k\right) = \min_{\alpha \geq 0} g_k(\alpha),\quad \alpha_k > 0. \tag{5.9}$$

Метод (5.8), (5.9), в якому кроковий множник α_k вибирається з умови мінімізації функції $f(x)$ вздовж напрямку антиградієнта, носить назву методу найшвидшого спуску.

2. Градієнтний метод з дробленням кроку

Вибираються певні константи $\beta > 0$, $0 < \lambda < 1$ (часто $\lambda = \frac{1}{2}$). Для коефіцієнта $\alpha = \beta$ перевіряється умова монотонності

$$f\left(x^{(k)} - \alpha f'\left(x^{(k)}\right)\right) < f\left(x^{(k)}\right). \tag{5.10}$$

Якщо (5.10) виконана, то покладають $\alpha_k = \alpha$. Якщо ні, то проводиться дроблення кроку, тобто приймається $\alpha = \lambda \cdot \beta$, і знову перевіряється виконання умови (5.10). Процес дроблення продовжується доти, поки умова (5.10) не виявиться виконаною. Цей процес не може бути нескінченим, так як $\left(-f'\left(x^{(k)}\right)\right)$ – напрямок спадання. Перше α, для якого умова (5.10) виявиться виконаною, приймається за α_k.

3. Градієнтний метод з адаптивним вибором кроку

Виберемо α_k з умови:

$$f\left(x^{(k)} - \alpha f'\left(x^{(k)}\right)\right) - f\left(x^{(k)}\right) \le -\varepsilon\, \alpha_k \left\| f'\left(x^{(k)}\right) \right\|^2, \tag{5.11}$$

де $0 < \varepsilon < 1$ – довільна стала (однакова на всіх ітераціях), яка називається параметром методу. Якщо ε дуже мале, то метод може збігатися повільно, а якщо ε дуже велике, то ускладняється вибір методу α_k з умови (5.11). Наведемо алгоритм вибору α_k на k-й ітерації:

1) вибираємо деяке довільне α (однакове на всіх ітераціях) і визначаємо точку $x = x^{(k)} - \alpha f'\left(x^{(k)}\right)$,
2) обчислимо $f(x) = f\left(x^{(k)} - \alpha f'\left(x^{(k)}\right)\right)$,
3) перевіримо нерівність (5.11) при $\alpha_k = \alpha$,
4) якщо нерівність (5.11) виконано, то значення α вибираємо за α_k і переходимо до наступної ітерації. У протилежному випадку проводимо дроблення α (помножуючи його на довільне число $0 < \lambda < 1$) поки нерівність (5.11) не буде виконаною. Тоді переходимо до наступної ітерації.

Відзначимо, що з виконання нерівності (5.11) випливає виконання нерівності (5.10). Цей градієнтний метод найбільш часто використовується на практиці.

Збіжність градієнтних методів

Розглянуті варіанти градієнтних методів за своїми властивостями близькі один до одного. Вони можуть використовуватися для мінімізації функцій, які належать одному класу функцій, при цьому швидкість збіжності (в тих випадках, коли її вдається оцінити) приблизно однакова.

Тому сформулюємо теореми про збіжність тільки для методу найшвидшого спуску, який визначається формулами (5.8), (5.9).

Теорема 1. Нехай

1) функція $f(x)$ диференційовна на E^n;
2) функція $f(x)$ обмежена знизу на E^n, тобто $inf\ f(x) = f_* > -\infty$;
3) її градієнт задовольняє умову Ліпшиця:

$$\left\| f'(\overline{x}) - f'(\overline{z}) \right\| \le L \left\| \overline{x} - \overline{z} \right\|, \text{ для } \forall\, \overline{x}, \overline{z} \in E^n, \tag{5.12}$$

де L – константа Ліпшиця.

Тоді для послідовності $\left\{ x^{(k)} \right\}_0^\infty$, яку отримано за формулами (5.8), (5.9), має місце $\left\| f'\left(x^{(k)}\right) \right\| \to 0$, $k \to \infty$ для будь-якої початкової точки $x^{(0)}$.

4) Якщо при цьому множина $M\left(x^{(0)}\right) = \left\{ x \in E^n : f(x) \le f\left(x^{(0)}\right) \right\}$ – обмежена, то

$$\lim_{k\to\infty} \rho\left(x^{(k)}, S_*\right) = 0,$$

де $S_* = \left\{ x \in M\left(x^{(0)}\right) : f'(x) = 0 \right\}$ – множина стаціонарних точок функції $f(x)$ на $M\left(x^{(0)}\right)$, $\rho\left(x^{(k)}, S_*\right)$ – відстань від точки $x^{(k)}$ до множини S_* стаціонарних точок функції $f(x)$.

У випадку виконання умов теореми 1 метод найшвидшого спуску забезпечує тільки збіжність до множини стаціонарних точок S_*, але при таких умовах неможливо оцінити швидкість збіжності і стверджувати збіжність до множини точок мінімуму X_*.

Теорема 2. Нехай виконані умови теореми 1 і, окрім того, функція $f(x)$ опукла на E^n, тоді для послідовності $\left\{ x^{(k)} \right\}_0^\infty$, яку отримано за формулами (5.8)-(5.9), мають місце такі твердження:

1) $\lim\limits_{k\to\infty} f\left(x^{(k)}\right) = f_*$, тобто послідовність $\left\{ x^{(k)} \right\}_0^\infty$ є мінімізуючою,
2) $\lim\limits_{k\to\infty} \rho\left(x^{(k)}, X_*\right) = 0$, послідовність $\left\{ x^{(k)} \right\}_0^\infty$ збігається до множини точок мінімуму X_*,
3) справедлива така оцінка за функцією:

$$0 \le f\left(x^{(k)}\right) - f(x_*) \le 4D^2 L \frac{1}{k}, \quad k = 1, 2, \ldots, \tag{5.13}$$

де $D = \sup\limits_{x, x' \in M\left(x^{(0)}\right)} \left\| x - x' \right\|$ – діаметр множини $M\left(x^{(0)}\right)$.

У теоремі 2 оцінюється швидкість збіжності за функцією. Оцінка швидкості збіжності за змінною можлива лише при більш жорстких припущеннях відносно властивостей функції $f\left(x\right)$.

Теорема 3. Нехай $f\left(x\right)$ – двічі неперервно-диференційовна функція, її матриця других частинних похідних $f''(x)$ задовольняє умову:

$$m\| z\|^2 \le \left(f''\left(x\right) z, z\right) \le M \| z\|^2, \; 0 < m \le M, \;\; \forall x, z \in E^n. \qquad (5.14)$$

Тоді послідовність $\left\{x^{(k)}\right\}_0^\infty$, яку отримано за формулами (5.8)-(5.9), збігається до точки мінімуму зі швидкістю геометричної прогресії із знаменником $q = \left(M - m\right) / \left(M + m\right)$ (з лінійною швидкістю), тобто при достатньо великих k виконується нерівність

$$\left\|x^{(k+1)} - x_*\right\| \le \frac{M - m}{M + m} \left\|x^{(k)} - x_*\right\|. \qquad (5.15)$$

Функція $f\left(x\right)$, яка задовольняє умові (5.14), називається *сильно опуклою*.

Геометричний зміст градієнтних методів

1. У точці $x^{(k+1)}$, яка визначається за формулами (5.8)-(5.9), градієнт $f'\left(x^{(k+1)}\right)$ є ортогональним лінії рівня $\Gamma^{(k+1)} = \left\{x \in E^n \colon f\left(x\right) = f\left(x^{(k+1)}\right)\right\}$.
2. Для методу найшвидшого спуску напрямки спуску $\left(-f'\left(x^{(k)}\right)\right)$ та $\left(-f'\left(x^{(k+1)}\right)\right)$ на двох послідовних ітераціях взаємно ортогональні.
3. Градієнтний метод (відмінний від методу найшвидшого спуску) генерує послідовність точок $\left\{x^{(k)}\right\}_0^\infty$, які утворюють зигзагоподібну траєкторію $x^{(0)}, x^{(1)}, x^{(2)}, \ldots, x^{(k)}$.

Лекція

Метод Ньютона

Постановка задачі і умова застосування методу Ньютона:

$$f(x) \to min,\ x \in E^n,\ f(x) \in C^2(E^n). \tag{6.1}$$

Метод Ньютона є методом другого порядку, тобто використовує обчислення других похідних функції $f(x)$, яка мінімізується.

У методі Ньютона послідовність $\left\{x^{(k)}\right\}_0^{\infty}$ генерується виходячи з квадратичної апроксимації цільової функції.

Ітераційна формула методу Ньютона має вигляд:

$$x^{(k+1)} = x^{(k)} - \left[f''\left(x^{(k)}\right)\right]^{-1} f'\left(x^{(k)}\right), k = 0,1,2,... \tag{6.2}$$

Теорема про збіжність методу Ньютона. Нехай

1. Функція $f(x)$ є двічі неперервно-диференційована на E^n.
2. Функція $f(x)$ є сильно опуклою на E^n:
$$m\|z\|^2 \le \left(f''(x)\,z, z\right) \le M\|z\|^2,\ 0 < m \le M,\ \ \forall x, z \in E^n.$$
3. Матриця других похідних (матриця Гессе) задовольняє умові Ліпшиця:
$$\left\|f''(x) - f''(z)\right\| \le L\|x - z\|, \text{ для } \forall\ x, z \in E^n,\quad L > 0.$$
4. Початкове наближення $x^{(0)}$ таке, що
$$\left\|f'\left(x^{(0)}\right)\right\| \le \frac{8M^2}{L} \text{ або } \left\|f'\left(x^{(0)}\right)\right\| = \frac{8M^2 q}{L}, \text{ де } q \in (0,1). \tag{6.3}$$

Тоді послідовність (5.17) збігається до точки мінімуму з квадратичною швидкістю і має місце така оцінка:

$$\left\|x^{(k)} - x_*\right\| \le \frac{4M}{L} q^{2^k},\quad k = 0,1,\ldots. \tag{6.4}$$

Збіжність методу Ньютона доведено лише для достатньо близького до точки мінімуму початкового наближення $x^{(0)}$. Для розташованих далеко від точки мінімуму початкових наближень метод може розбігатися. При цьому умову (6.3), яка гарантує збіжність методу для даного початкового наближення, на практиці важко перевірити у силу того, що константи M і L, як правило, невідомі.

Залежність збіжності методу від початкового наближення є недоліком методу Ньютона. Ще одним недоліком є висока трудомісткість, яка обумовлена необхідністю обчислення і обертання на кожному кроці матриці Гессе.

У зв'язку з названими вище обставинами застосування класичного методу Ньютона далеко не завжди призводить до успіху. Модифікації його направлені на те, щоб зберігаючи основну перевагу методу Ньютона – швидку збіжність – зменшити трудомісткість та ослабити вимоги до вибору початкового наближення.

Модифікація методу Ньютона (метод Ньютона з регулюванням кроку)

Розглянемо метод

$$x^{(k+1)} = x^{(k)} - \alpha_k \left[f''\left(x^{(k)}\right) \right]^{-1} f'\left(x^{(k)}\right), \alpha_k > 0,, k = 0,1,2,\dots, \qquad (6.5)$$

який називається методом Ньютона з регулюванням кроку. При $\alpha_k = 1$ він співпадає з класичним методом Ньютона.

Вибір коефіцієнтів α_k в (6.5) здійснюється аналогічно градієнтним методам.

Наприклад, адаптивний спосіб визначення α_k. Виберемо α_k з умови:

$$f\left(x^{(k)} - \alpha_k \left[f''\left(x^{(k)}\right) \right]^{-1} f'\left(x^{(k)}\right) \right) - f\left(x^{(k)}\right) \le \varepsilon\, \alpha_k \left(f'\left(x^{(k)}\right), p^{(k)} \right), \quad (6.6)$$

де $0 < \varepsilon < 1$ – довільна стала (однакова на всіх ітераціях), яка називається параметром методу.

Знаходження α_k з умови мінімізації цільової функції уздовж заданого напрямку $p^{(k)}$:

$$f\left(x^{(k)} - \alpha_k \left[f''\left(x^{(k)}\right) \right]^{-1} f'\left(x^{(k)}\right) \right) = \min_{\alpha \ge 0} f\left(x^{(k)} - \alpha \left[f''\left(x^{(k)}\right) \right]^{-1} f'\left(x^{(k)}\right) \right). \quad (6.7)$$

Можна показати, що вказані варіанти методу (6.5) збігаються для будь якого початкового наближення $x^{(0)} \in E^n$. При цьому швидкість збіжності буде або надлінійною, або квадратичною залежно від вимог, яким буде задовольняти цільова функція $f(x)$.

Зменшити трудомісткість методу можна, якщо обчислювати матрицю Гессе не на кожному кроці, як це робиться в (6.5), а один раз через кожні s кроків. Підбираючи емпіричним шляхом s, іноді вдається отримати за допомогою цього методу непогані результати. Однак, істотного зменшення трудомісткості методу Ньютона цей прийом не дає.

Приклади чисельної реалізації

1. Розглянемо задачу мінімізації квадратичної функції

$$f(x) = 7x_1^2 + 4x_1x_2 + 2x_2^2 + 10x_1$$

$$7x_1^2 + 4x_1x_2 + 2x_2^2 + 10x_1 = const - \text{лінії}\ \text{рівня},$$

$x = (x_1, x_2) \in E^2$.

Оптимальний розв’язок цієї задачі є точка $x_* = (-1.00, 1.00)$, $f(x_*) = -5$. Лінії рівня цієї функції наведено на рис. 1.

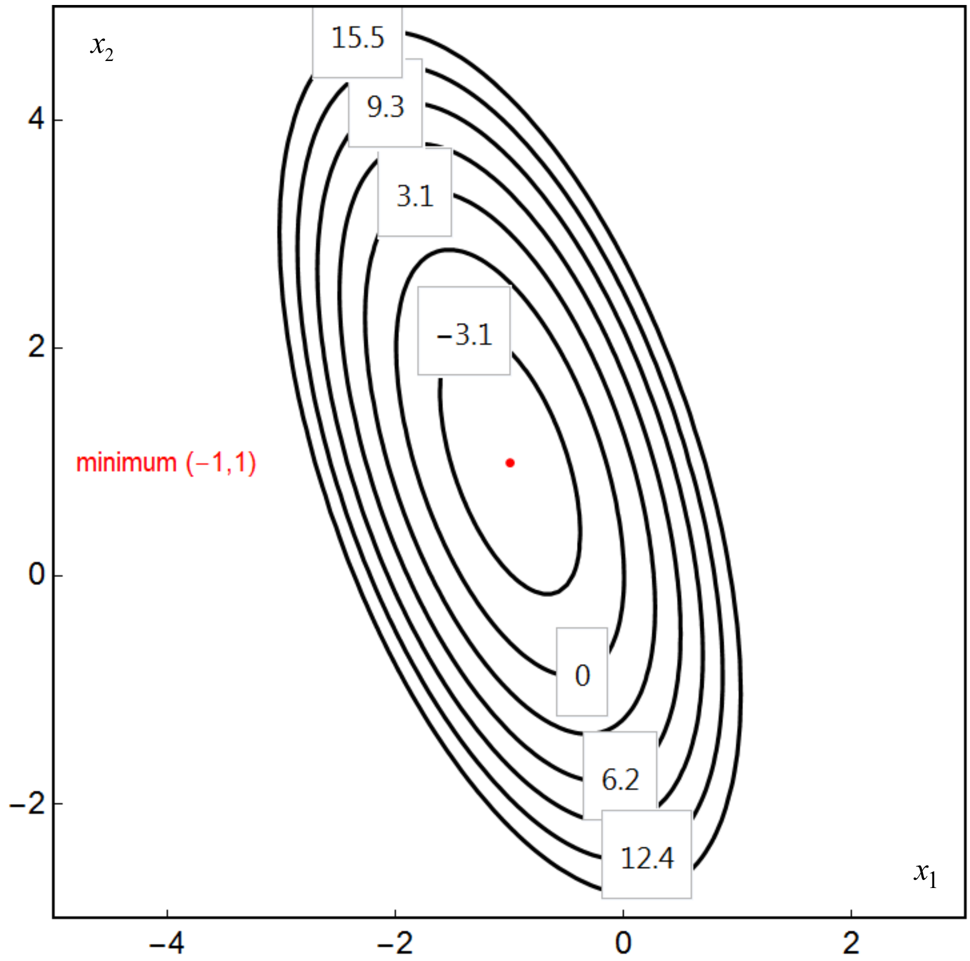


Рис. 1

Знайдемо точку мінімуму функції $f(x)$ класичним методом. Градієнт цієї функції має вигляд:

$$f'(x) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 14x_1 + 4x_2 + 10 \\ 4x_1 + 4x_2 \end{pmatrix}.$$

За необхідної умови мінімуму маємо:

$$\begin{cases} 14x_1 + 4x_2 + 10 = 0, \\ 4x_1 + 4x_2 = 0. \end{cases} \Rightarrow \begin{cases} x_1 = -1, \\ x_2 = 1. \end{cases}$$

Будуємо матрицю других похідних (гесіан):

$$f''(x) = \begin{pmatrix} 14 & 4 \\ 4 & 4 \end{pmatrix}.$$

Кутові мінори $M_1 = 14 > 0$, $M_2 = \begin{vmatrix} 14 & 4 \\ 4 & 4 \end{vmatrix} = 40 > 0$. Отже, гесіан є додатно визначений, тобто точка $x^* = (-1,1)$ є точкою мінімуму і $f(x^*) = -5$.

2. Знайдемо два перші наближення до точки мінімуму функції $f\left(x\right)$ за методом найшвидшого градієнтного спуску.

Нехай $x^{(0)}=\left(0{,}0\right)^T$ – початкова точка, $f\left(x^{(0)}\right)=0$. Як початкову можна брати будь-яку точку простору E^2.

Для визначення крокового множника α_k побудуємо функцію $g\left(\alpha\right)=f\left(x^{(k)}-\alpha\, f'\left(x^{(k)}\right)\right)$, мінімум якої при $\alpha\geq 0$ необхідно визначити. Ця функція є функцією однієї змінної, таким чином, для знаходження мінімуму можна використати класичний метод або, при програмній реалізації, один з методів одновимірної оптимізації.

Перша ітерація. Обчислимо

$$x^{(1)}=x^{(0)}-\alpha_0\, f'\left(x^{(0)}\right),$$

де α_0 – шуканий кроковий множник, $f'\left(0{,}0\right)=\begin{pmatrix}14\cdot 0+4\cdot 0+10\\ 4\cdot 0+4\cdot 0\end{pmatrix}=$ $f'\left(x^{(0)}\right)=\begin{pmatrix}10\\ 0\end{pmatrix}$. Тоді отримаємо

$$x^{(1)}=\begin{pmatrix}0\\ 0\end{pmatrix}-\alpha_0\begin{pmatrix}10\\ 0\end{pmatrix}=\begin{pmatrix}-10\alpha_0\\ 0\end{pmatrix}.$$

Для визначення α_0 побудуємо функцію $g\left(\alpha_0\right)=f\left(x^{(0)}-\alpha_0\, f'\left(x^{(0)}\right)\right)$, $f\left(x\right)=7x_1^2+4x_1x_2+2x_2^2+10x_1$ і знайдемо мінімум цієї функції, тобто

$$g\left(\alpha_0\right)=7\cdot 100\alpha_0^2+4\cdot\left(-10\alpha_0\right)\cdot 0+2\cdot 0^2+10\cdot\left(-10\alpha_0\right)=700\alpha_0^2-100\alpha_0\to min,$$

$$\alpha_0\geq 0.$$

Знайдемо розв’язок одновимірної задачі оптимізації класичним методом.

$$g'\left(\alpha_0\right)=0:\ 1400\alpha_0-100=0,\ \Rightarrow\alpha_0=\frac{1}{14}.$$

Таким чином, $x^{(1)}=\begin{pmatrix}-5/7\\ 0\end{pmatrix}$, $f\left(x^{(1)}\right)=-25/7\approx-3{,}57$. Умова монотонності виконується, так як $f\left(x^{(1)}\right)<f\left(x^{(0)}\right)$.

Перевіримо критерій завершення ітераційного процесу, тобто умову $\left\|x^{(k+1)}-x^{(k)}\right\|<\varepsilon$, де ε – задана точність. Нехай $\varepsilon=10^{-3}$.

$$\left\|x^{(1)}-x^{(0)}\right\|=\sqrt{\left(x_1^{(1)}-x_1^{(0)}\right)^2+\left(x_2^{(1)}-x_2^{(0)}\right)^2}=\sqrt{\left(5/7-0\right)^2+\left(0-0\right)^2}=5/7>\varepsilon.$$

Друга ітерація. Обчислимо наступну точку $x^{(2)}$:

$$f'\left(-\frac{5}{7},0\right)=\begin{pmatrix}14\cdot(-\frac{5}{7})+4\cdot 0+10\\ 4\cdot(-\frac{5}{7})+4\cdot 0\end{pmatrix}=\begin{pmatrix}0\\ 4\cdot(-\frac{5}{7})\end{pmatrix}$$

$$x^{(2)}=\begin{pmatrix}-5/7\\ 0\end{pmatrix}-\alpha_1\begin{pmatrix}0\\ -20/7\end{pmatrix}=\begin{pmatrix}-5/7\\ 20\alpha_1/7\end{pmatrix}.$$

$$f(x)=7x_1^2+4x_1x_2+2x_2^2+10x_1$$

$$g(\alpha_1)=f\left(x^{(1)}-\alpha_1 f'\left(x^{(1)}\right)\right)=\frac{800}{49}\alpha_1^2-\frac{400}{49}\alpha_1-\frac{25}{7}\to min$$

Отже, $\begin{pmatrix}-5/7\\ 20\alpha_1/7\end{pmatrix}$,

$$f(x)=7\cdot\frac{25}{49}+4(\frac{-5}{7})\frac{20\alpha_1}{7}+2\frac{400\alpha_1^2}{49}+10(-5/7)=$$

$\alpha_1\geq 0$, звідки отримаємо $g'(\alpha_1)=\frac{1600}{49}\alpha_1-\frac{400}{49}=0$ і $\alpha_1=\frac{1}{4}$.

Тоді $x^{(2)}=\begin{pmatrix}-5/7\\ 5/7\end{pmatrix}$. $f\left(x^{(2)}\right)=-225/49\approx-4{,}59$. Умова монотонності виконується, так як $f\left(x^{(2)}\right)<f\left(x^{(1)}\right)$.

Перевіримо критерій завершення ітераційного процесу. Наприклад, умову $\left\|x^{(k+1)}-x^{(k)}\right\|<\varepsilon$, де ε – задана точність. Нехай $\varepsilon=10^{-3}$.

$$\left\|x^{(2)}-x^{(1)}\right\|=\sqrt{\left(x_1^{(2)}-x_1^{(1)}\right)^2+\left(x_2^{(2)}-x_2^{(1)}\right)^2}=5/7>\varepsilon.$$

Наближене значення мінімуму за дві ітерації $x^*=x^{(2)}=\begin{pmatrix}-5/7\\ 5/7\end{pmatrix}$, $f\left(x^*\right)\approx-4{,}59$.

3. Знайдемо наближення до точки мінімуму функції $f(x)$ за градієнтним методом с дробленням кроку. Результати чисельного експерименту наведено у табл. 1 при початковому наближенні $x^{(0)}=(0,0)^T$, $f\left(x^{(0)}\right)=0$, $\alpha_0=0{,}1$. Ітераційна процедура була зупинена при виконанні умови $\left\|x^{(k+1)}-x^{(k)}\right\|\leq\varepsilon$ ($\varepsilon=10^{-2}$).

Таблиця 1

k	$x^{(k)}$	$f\left(x^{(k)}\right)$	$-f'\left(x^{(k)}\right)$	$\left\Vert x^{(k+1)}-x^{(k)}\right\Vert$
1	$\begin{pmatrix}-1\\0\end{pmatrix}$	-3	$\begin{pmatrix}10\\0\end{pmatrix}$	1
2	$\begin{pmatrix}-0,6\\0,4\end{pmatrix}$	-4,12	$\begin{pmatrix}-4\\-4\end{pmatrix}$	0,5657
3	$\begin{pmatrix}-0,92\\0,48\end{pmatrix}$	-4,58	$\begin{pmatrix}3,2\\-0,8\end{pmatrix}$	0,3299
4	$\begin{pmatrix}-0,824\\0,656\end{pmatrix}$	-4,789	$\begin{pmatrix}-0,96\\-1,76\end{pmatrix}$	0,2005
5	$\begin{pmatrix}-0,9328\\0,7232\end{pmatrix}$	-4,8896	$\begin{pmatrix}1,088\\-0,672\end{pmatrix}$	0,1279
6	$\begin{pmatrix}-0,91616\\0,80704\end{pmatrix}$	-4,94104	$\begin{pmatrix}-0,1664\\-0,8384\end{pmatrix}$	0,0855
7	$\begin{pmatrix}-0,956352\\0,850688\end{pmatrix}$	-4,96814	$\begin{pmatrix}0,40192\\-0,43648\end{pmatrix}$	0,0593
8	$\begin{pmatrix}-0,957734\\0,892954\end{pmatrix}$	-4,98267	$\begin{pmatrix}0,013824\\-0,422656\end{pmatrix}$	0,0423
9	$\begin{pmatrix}-0,974088\\0,918866\end{pmatrix}$	-4,99054	$\begin{pmatrix}0,163533\\-0,259123\end{pmatrix}$	0,0306
10	$\begin{pmatrix}-0,977911\\0,940955\end{pmatrix}$	-4,99483	$\begin{pmatrix}0,0382362\\-0,220887\end{pmatrix}$	0,0224
11	$\begin{pmatrix}-0,985217\\0,955737\end{pmatrix}$	-4,99717	$\begin{pmatrix}0,0730604\\-0,147827\end{pmatrix}$	0,0165
12	$\begin{pmatrix}-0,988208\\0,967529\end{pmatrix}$	-4,99845	$\begin{pmatrix}0,0299065\\-0,11792\end{pmatrix}$	0,0122
13	$\begin{pmatrix}-0,991729\\0,975801\end{pmatrix}$	-4,99915	$\begin{pmatrix}0,0352054\\-0,0827147\end{pmatrix}$	0,009

4. Розв’яжемо задачу методом Ньютона. Матриця других похідних для квадратичної функції є постійною і для заданої функції має вигляд:

$$f''(x)=\begin{pmatrix}14 & 4\\ 4 & 4\end{pmatrix}.$$

Обернена до матриці Гессе $f''(x)$ матриця є

$$\left[f''(x)\right]^{-1}=\frac{1}{40}\begin{pmatrix}4 & -4\\ -4 & 14\end{pmatrix}=\frac{1}{20}\begin{pmatrix}2 & -2\\ -2 & 7\end{pmatrix}.$$

За початкове наближення візьмемо точку $x^{(0)}=(0,0)^T$, $f\left(x^{(0)}\right)=0$. За ітераційною формулою (6.2) для першого наближення маємо:

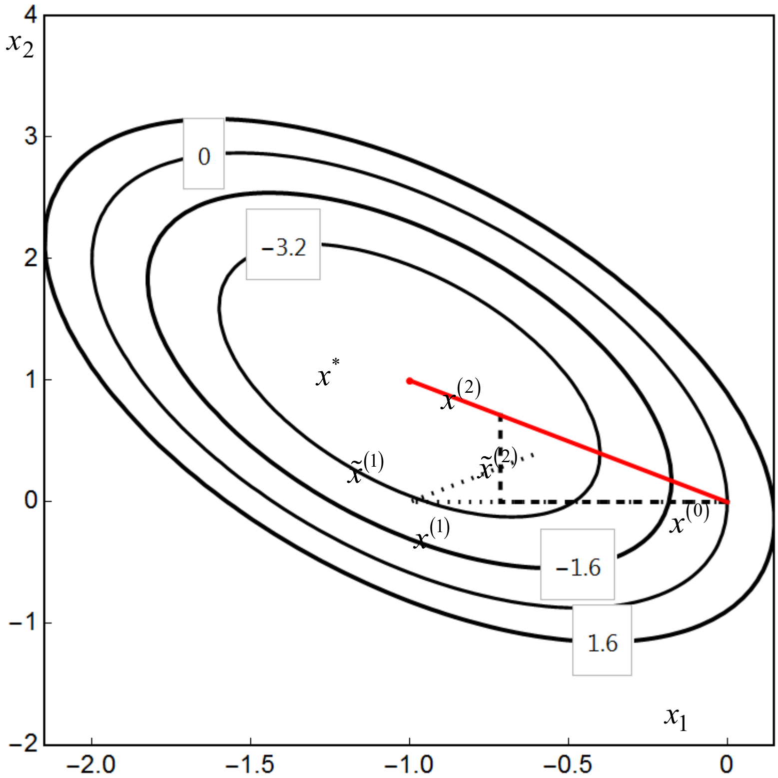
$$x^{(1)} = x^{(0)} - \left[f''\left(x^{(0)}\right) \right]^{-1} f'\left(x^{(0)}\right),$$

$$x^{(1)} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{20} \begin{pmatrix} 2 & -2 \\ -2 & 7 \end{pmatrix} \begin{pmatrix} 10 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \frac{1}{20} \begin{pmatrix} 20 \\ -20 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} - \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

$$x^{(1)} = \begin{pmatrix} 10 \\ 5 \end{pmatrix} - \frac{1}{20} \begin{pmatrix} 2 & -2 \\ -2 & 7 \end{pmatrix} \begin{pmatrix} 170 \\ 60 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 22 \\ 8 \end{pmatrix} = \begin{pmatrix} 10 \\ 5 \end{pmatrix} - \begin{pmatrix} 11 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Як бачимо, $x_* = x^{(1)}$ – точка мінімуму. Тобто, для квадратичної функції, матриця других похідних $f''(x)$ якої є додатно визначеною, метод Ньютона збігається за одну ітерацію до точного значення мінімуму.

Перші дві ітерації градієнтних методів та метод Ньютона проілюстровані графічно (рис. 2).



- - - - – метод найшвидшого градієнтного спуску, точки $x^{(0)}, x^{(1)}, x^{(2)}$;

...... – градієнтний метод с дробленням кроку, точки $x^{(0)}, \tilde{x}^{(1)}, \tilde{x}^{(2)}$;

―― – метод Ньютона, точки $x^{(0)}, x^*$.

Рис. 2

5. Розглянемо функцію $f(x) = 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2$, яка називається функцією Розенброка (Rosenbrock).

Функція Розенброка – невипукла функція, яка використовується для оцінки продуктивності алгоритмів оптимізації, була запропонована Х. Розенброком в 1960 році. Ця функція має надзвичайно пологий вигнутий яр, що сильно ускладнює пошук. Завдяки "неприємним" для методів пошуку екстремуму особливостям функція Розенброка часто використовується для випробування збіжності різних алгоритмів і для їх порівняння.

Функція Розенброка має декілька локальних мінімумів, але тільки один глобальний мінімум, який знаходиться в точці $x_* = (1,\ 1)$, $f(x_*) = 0$.

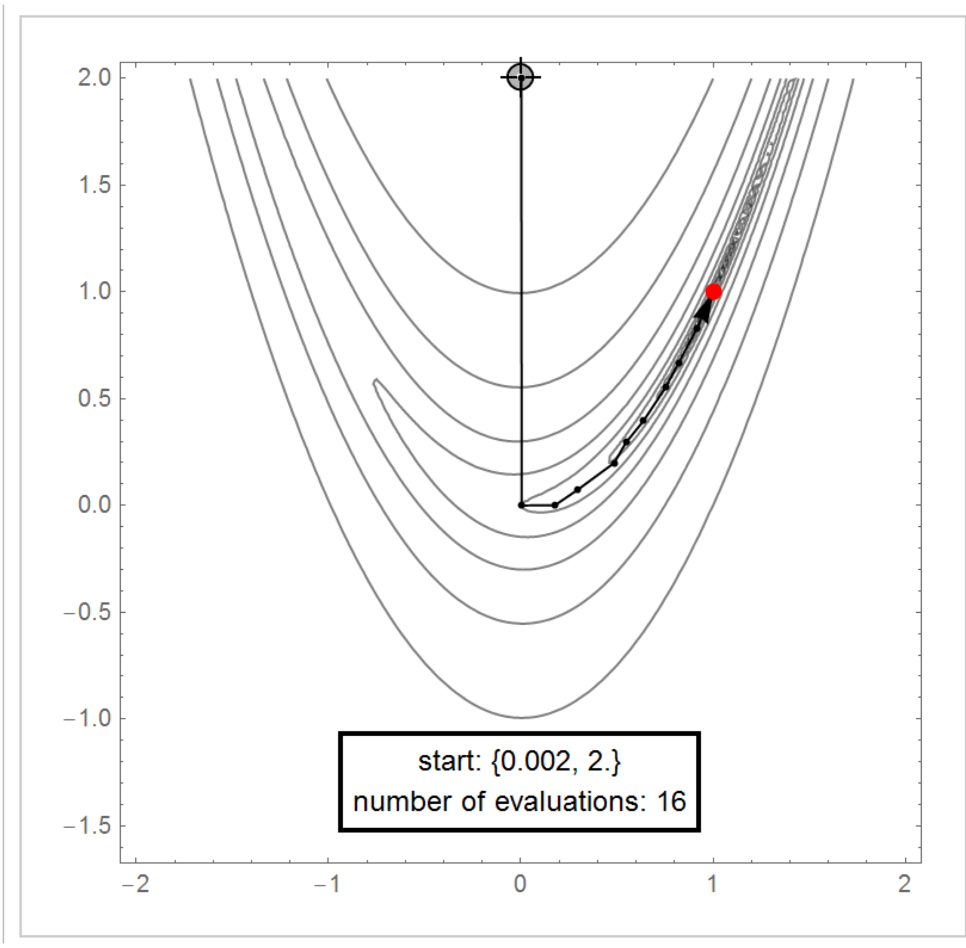


Рис. 3.

Рис. 3 ілюструє збіжність методу Ньютона з початкової точки $x^{(0)} = (0{,}002,\ 2)$ за 16 ітерацій.

$$f(x) = 100\left(x_2 - x_1^2\right)^2 + \left(1 - x_1\right)^2,$$

$$f'(x) = \begin{pmatrix} -400x_1\left(x_2 - x_1^2\right) - 2\left(1 - x_1\right) \\ 200\left(x_2 - x_1^2\right) \end{pmatrix},$$

$$f''(x) = \begin{pmatrix} 2 - 400\left(x_2 - x_1^2\right) + 800x_1^2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}.$$

Контрольні запитання

1. Класифікація методів безумовної мінімізації функцій багатьох змінних.
2. Що таке градієнт та антиградієнт функції?
3. Основна властивість градієнту(антиградієнту).
4. В чому полягає суть градієнтних методів пошуку оптимуму?
5. Чим градієнтний метод з дробленням кроку відрізняється від методу найшвидшого спуску?
6. До методів якого порядку відноситься метод Ньютона мінімізації функції багатьох змінних?
7. Дайте визначення матриці Гессе.
8. У яких випадках послідовність, яку побудовано за методом Ньютона, може розбігатися?
9. Критерії порівняння та порівняльна характеристика градієнтних методів багатовимірної оптимізації.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Акулич И. Л. Математическое программирование в примерах и задачах / И. Л. Акулич. – М.: Высшая школа, 1986. – 319 с.
2. Аоки М. Введение в методы оптимизации / М. Аоки. – М.: Наука, 1977. – 344 с.
3. Ашманов С. А. Линейное программирование / С. А. Ашманов. – М.: Наука, 1981. – 340 с.
4. Базара М. Нелинейное программирование. Теория и алгоритмы / М. Базара, К. Шетти. – М.: Мир, 1982. – 583 с.
5. Бейко И. В. Методы и алгоритмы решения задач оптимизации / И. В. Бейко, Б. Н. Бублик, П. Н. Зинько– Киев: Вища школа, 1983. – 512 с.
6. Васильев Ф. П. Лекции по методам решения экстремальных задач / Ф. П. Васильев. – М.: МГУ, 1974. – 374 с.
7. Васильев Ф. П. Численные методы решения экстремальных задач / Ф. П. Васильев. – М.: Наука, 1988. – 552 с.
8. Жалдак М. І. Основи теорії та методов оптимізації: Навчальний посібник / М. І. Жалдак, Ю. В. Триус – Черкаси: Брама-Україна, 2005. – 608 с.
9. Зайченко Ю. П. Дослідження операцій: Підручник / Ю. П. Зайченко. – К., Видавничий дом «Слово», 2000. – 816 с.
10. Зайченко О. Ю. Дослідження операцій. Збірник задач / О. Ю. Зайченко, Ю. П. Зайченко. –К.: Видавничий дом «Слово», 2007. – 472 с.
11. Капустин В. Ф. Практические занятия по курсу математического программирования / В. Ф. Капустин. – Ленинград, изд-во Ленинградского университета, 1976. – 192 с.
12. Кісельова О. М. Чисельні методи оптимізації. Навч. посібник / О. М. Кісельова, А. Є. Шевельова. – Д.: Вид-во ДНУ, 2008. – 212 с.
13. Ляшенко И. Н. Линейное и нелинейное программирование / И. Н. Ляшенко, Е. А. Карагодова, Н. В. Черникова, Н. З. Шор. – Киев: Вища школа, 1975. – 372 с.
14. Мину М. Математическое программирование / М. Мину. – М.: Наука, 1990. – 488 с.
15. Моисеев Н. Н. Методы оптимизации / Н. Н. Моисеев, Ю. П. Иванилов, Е. М. Столярова. – М.: Наука, 1978. – 351 с.
16. Наконечний С. І. Математичне програмування: Навч. посіб. / С. І. Наконечний, С. С. Савіна. – К.: КНЕУ, 2003. – 452 с.
17. Попов Ю. Д. Методи оптимізації / Ю. Д. Попов, В. І. Тюптя, В. І. Шевченко. – К.: Ел.вид КНУ, 2003. – 215 с.
18. Пшеничный Б. Н. Численные методы в экстремальных задача / Б. Н. Пшеничный, Ю. М. Данилин. – М.: Наука, 1986. – 320 с.
19. Сухарев А. Г. Курс методов оптимизации / А. Г. Сухарев, А. В. Тимохов, В. В. Федоров. – М.,1986.– 328 с.
20. Химмельблау Д. Прикладное нелинейное программирование / Д. Химмельблау. – М.: Мир, 1975. – 536 с.