Sprawozdanie - laboratorium nr 10
Poszukiwanie minimum wartości funkcji w dwóch wymiarach metodą Newtona.

Damian Płóciennik

8 maja 2019

1 Wstęp teoretyczny

1.1 Optymalizacja funkcji

Zadaniem optymalizacji jest poszukiwanie minimum lub maksimum funkcji (wielu zmiennych).

W praktyce problem sprowadza się do poszukiwania minimum, czyli takiego punktu dla którego zachodzi:

$$f : R^n \rightarrow R, \tag{1}$$

$$\min f(\vec{x}) = f(\vec{x}^*) \Leftrightarrow \bigwedge_{\vec{x} \in R^n} f(\vec{x}^*) < f(\vec{x}), \tag{2}$$

$$\vec{x} = [x_1, x_2, \ldots, x_n]^T. \tag{3}$$

1.2 Metoda Newtona poszukiwania minimum funkcji kwadratowej w $\mathbf{R^n}$

Funkcję kwadratową definiujemy następująco:

$$f(\vec{x}) = \frac{1}{2}\vec{x}^T \mathbf{A}\vec{x} + \vec{x}^T \vec{b} + c \tag{4}$$

gdzie: $\mathbf{A}$ jest macierzą kwadratową oraz

$$\vec{x}, \vec{b} \in R^n, c \in R. \tag{5}$$

Jeśli macierz $\mathbf{A}$ jest symetryczna to wówczas zachodzi:

$$\nabla f(\vec{x}) = \mathbf{A}\vec{x} + \vec{b} \tag{6}$$

oraz

$$\nabla^2 f(\vec{x}) = \mathbf{A} \Leftrightarrow H(\vec{x}) = \mathbf{A}. \tag{7}$$

Jeśli $\mathbf{A}$ jest dodatniookreślona to rozwiązanie można łatwo znaleźć, ponieważ:

$$\nabla f(\vec{x}) = \mathbf{A}\vec{x} + \vec{b} = 0, \tag{8}$$

$$\vec{x}^* = -\mathbf{A}^{-1}\vec{b}. \tag{9}$$

W metodzie Newtona zakładamy:

$$\vec{x}^* = \vec{x}^i + \delta, \tag{10}$$

gdzie: $\vec{x}^i$ to przybliżenie rozwiązania w i-tej iteracji.

Korzystając z rozwinięcia funkcji w szereg Taylora można zapisać:

$$0 = \nabla f(\vec{x}^*) = \nabla f(\vec{x}^i + \delta) = \nabla f(\vec{x}^i) + H(\vec{x}^i)\delta + O(||\delta||^2). \tag{11}$$

Jeśli pominiemy wyrazy rzędu $||\delta||^2$ to:

$$\nabla f(\vec{x}^i) + H(\vec{x}^i)\delta = 0. \tag{12}$$

W i-tej iteracji poprawiamy rozwiązanie tj.:

$$\vec{x}^{i+1} = \vec{x}^i + \delta \tag{13}$$

i ostatecznie:

$$\vec{x}^{i+1} = \vec{x}^i - H^{-1}(\vec{x}^i)\nabla f(\vec{x}^i). \tag{14}$$

Oczekujemy, że metoda Newtona będzie pracować również dla innych funkcji niż kwadratowe tj. gdy badaną funkcję celu można lokalnie przybliżyć funkcją kwadratową.

2 Zadanie do wykonania

2.1 Opis problemu

Zadaniem było znalezienie miejsca zerowego funkcji:

$$f(x, y) = x^2 - 4x + 8 + y^2 - 4y + xy. \tag{15}$$

Zauważono, że liczba 8 przesuwa jedynie wszystkie wartości o stały poziom, więc została ona pominięta w dalszych rozważaniach. Szukano więc minimum funkcji:

$$g(x, y) = f(x, y) - 8 = x^2 - 4x + y^2 - 4y + xy. \tag{16}$$

Minimum funkcji poszukiwano przy użyciu metody Newtona jak dla funkcji kwadratowej:

$$g(\vec{r}) = \frac{1}{2}\vec{r}^T A\vec{r} + \vec{r}^T\vec{b}, \tag{17}$$

gdzie: $\vec{r}^T = [x, y]$, **A** jest macierzą Hessego, $\vec{b} = [b_1, b_2]$ jest wektorem wyrazów wolnych.

Znaleziono macierz **A**:

$$a_{11} = \frac{\partial^2 g}{\partial x^2} = 2 \tag{18}$$

$$a_{12} = \frac{\partial^2 g}{\partial x \partial y} = 1 \tag{19}$$

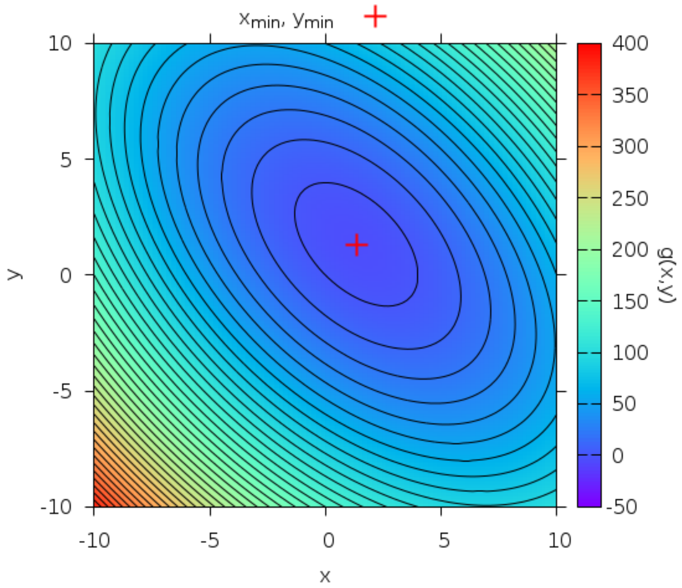
$$a_{21} = \frac{\partial^2 g}{\partial x \partial y} = 1 \tag{20}$$

$$a_{22} = \frac{\partial^2 g}{\partial y^2} = 2 \tag{21}$$

A wektor $\vec{b}$ miał postać $\vec{b} = [-4, -4]$.

2.2 Wyniki

Z wykorzystaniem skryptu w programie gnuplot dostarczonego przez prowadzącą zajęcia sporządono wykres konturowy wartości funkcji $g(x, y)$ w zakresie $x \in (-10, 10)$, $y \in (-10, 10)$.



Rysunek 1: Wykres konturowy na tle mapy wartości funkcji $g(x, y)$. Czerwonym krzyżykiem zaznaczono dokładne minimum funkcji: $x_{min} = \frac{4}{3}, y_{min} = \frac{4}{3}$.

Łatwo zauważyć, że przybliżone położenie minimum funkcji przedstawionej na powyższym wykresie znajduje się w miejscu dokładnego teoretycznego położenia oznaczonego czerwonym krzyżykiem i wynoszącego $x_{min} = \frac{4}{3}$, $y_{min} = \frac{4}{3}$.

Kolejno, korzystając z programu napisanego w języku C, na podstawie wzoru (9) również znaleziono minimum funkcji, które wynosiło:

$$\vec{r}_{min} = \begin{pmatrix} 1.33333 \\ 1.33333 \end{pmatrix}.$$

Łatwo zauważyć, ze uzyskany wynik jest zgodny z wynikiem teoretycznym oraz z położeniem odczytanym z wykresu konturowego.

Następnie do znalezienia minimum użyto przepisu iteracyjnego. Jako punkty startowe przyjęto kolejno: $(0, 0)$, $(10, -10)$, $(100, 100)$, $(500, 500)$, a jako warunek zakończenia przyjęto $\varepsilon_{i+1} = |\vec{r_{i+1}} - \vec{r_i}| < 10^{-6}$. Wyniki oraz liczbę iteracji konieczną do uzyskania zbieżności przedstawiono poniżej:

(x_0, y_0)	**Liczba iteracji**	$\mathbf{x_{min}}$	$\mathbf{y_{min}}$
(0, 0)	2	1.33333	1.33333
(10, -10)	2	1.33333	1.33333
(100, 100)	2	1.33333	1.33333
(500, 500)	2	1.33333	1.33333

Tabela 1: Wyniki poszukiwania minimum funkcji (x_{min}, y_{min}) dla różnych punktów startowych (x_0, y_0).

Łatwo zauważyć, że przyjęty punkt startowy nie wpływał ani na wynik końcowy, ani na liczbę iteracji przed uzyskaniem zbieżności. Ponadto bardzo mało obiegów było koniecznych, aby uzyskać oczekiwany wynik z dużą dokładnością.

Kolejno w przepisie iteracyjnym wprowadzono wagę, tj. przepis iteracyjny miał postać:

$$\vec{x}_{i+1} = \vec{r}_i - \omega \cdot H^{-1}(\vec{r^i}) \nabla g(\vec{r}), \tag{22}$$

gdzie: $\omega \in [0, 1]$ jest wagą.

Powtórzono obliczenia dla różnych wartości wagi, przyjmując punkt startowy jako $(x_0, y_0) = (10, 10)$.

ω	**Liczba iteracji**	$\mathbf{x_{min}}$	$\mathbf{y_{min}}$
0.1	135	1.33334	1.33334
0.4	32	1.33333	1.33333
0.7	15	1.33333	1.33333

Tabela 2: Wyniki poszukiwania minimum funkcji (x_{min}, y_{min}) przy różnych wartościach wagi ω dla punktu startowego $(x_0, y_0) = (10, 10)$.

Jak widać w powyższej tabeli, dla wag mniejszych od 1 ilość iteracji koniecznych do uzyskania zbieżności była wyraźnie wyższa. Łatwo zauważyć, że im mniejsza waga tym ilość iteracji znacznie wzrastała, natomiast dokładność ulegała niewielkiemu pogorszeniu.

3 Wnioski

Metoda Newtona pozwoliła na odnalezienie minimum wartości funkcji w dwóch wymiarach. Początkowo skorzystano z przypadku szczególnego, że rozważana funkcja jest kwadratowa, co pozwoliło na uzyskanie rozwiązania z bardzo dużą dokładnością w jednej iteracji, jednak wymagało znalezienia macierzy odwrotnej $\mathbf{A^{-1}}$. Użycie przepisu iteracyjnego pozwoliło uzyskać wyniki z bardzo dużą dokładnością po zaledwie dwóch obiegach. Zauważono także, że przyjęty punkt startowy nie wpływa ani na otrzymany wynik, ani na ilość iteracji. Wprowadzenie do przepisu iteracyjnego wagi mniejszej od 1 powodowało, że ilość obiegów potrzebnych do uzyskania zbieżności była wyraźnie wyższa, a dokładność otrzymanych wyników ulegała niewielkiemu pogorszeniu.

Reasumując, metoda Newtona jest bardzo dobrą metodą poszukiwania minimum wartości funkcji kwadratowej w dwóch wymiarach.

4 Bibliografia

- Chwiej Tomasz: Minimalizacja wartości funkcji. [online]. [dostęp: 14.05.2019]. Dostęp w Internecie:
  `http://galaxy.agh.edu.pl/~chwiej/mn/minimalizacja_1819.pdf`