Minimalizacja wartości funkcji metodą interpolacji kwadratowej Powella

Yuliya Zviarko

06.05.2024

Spis treści

1 Wstęp teoretyczny

1.1 Poszukiwanie Ekstremum: Optymalizacja(minimalizacja) Funkcji

Optymalizacja funkcji to proces znajdowania ekstremum, czyli minimum lub maksimum, funkcji z jedną lub wieloma zmiennymi.
Celem jest znalezienie punktu, dla którego zachodzi:

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

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

$$x = [x_1, x_2, x_3, \dots, x_n]^T \tag{3}$$

z uwzględnieniem warunków:

$$\begin{aligned} g_j(x) \leq 0, \quad j = 1, 2, \dots, m \\ h_j(x) = 0, \quad j = 1, 2, \dots, r \end{aligned} \tag{4}$$

gdzie funkcje f(x),g(x), h(x) są funkcjami skalarnymi.

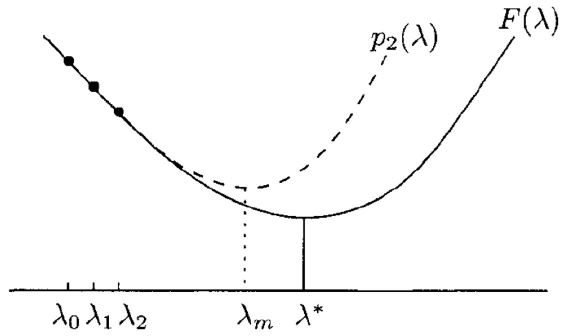
f(x) to **funkcja celu**,
a **g(x)** i **h(x)** to funkcje, które **określają warunki**, jakie musi spełniać rozwiązanie, ograniczając przestrzeń dopuszczalnych rozwiązań.

1.2 Metoda interpolacji kwadratowej Powella

Metoda interpolacji kwadratowej Powella polega na przybliżaniu funkcji za pomocą wielomianu drugiego stopnia prowadzonego przez trzy punkty: $\lambda_1, \lambda_2, \lambda_3$.

$$p_2(\lambda) = F(\lambda_0) + F[\lambda_0, \lambda_1](\lambda - \lambda_0) + F[\lambda_0, \lambda_1, \lambda_2](\lambda - \lambda_0)(\lambda - \lambda_1) \tag{5}$$

gdzie $F(\lambda_0)$ oznacza wartość funkcji, $F[\lambda_0, \lambda_1]$ to iloraz różnicowy pierwszego rzędu, a $F[\lambda_0, \lambda_1, \lambda_2]$ to iloraz różnicowy drugiego rzędu.



Rysunek 1: Rys. Wyznaczanie przybliżonego rozwiązania w metodzie Powell'a.
Źródło: [1].

Następnie, dla przybliżonego minimum, narzucamy warunek zerowania się pochodnej:

$$\frac{dp_2}{d\lambda} = F[\lambda_0, \lambda_1] + 2F[\lambda_0, \lambda_1, \lambda_2] - F[\lambda_0, \lambda_1, \lambda_2](\lambda_0 + \lambda_1) = 0. \tag{6}$$

Rozwiązując to równanie ze względu na λ, otrzymujemy:

$$\lambda_m = \frac{F[\lambda_0, \lambda_1, \lambda_2](\lambda_0 + \lambda_1) - F[\lambda_0, \lambda_1]}{2F[\lambda_0, \lambda_1, \lambda_2]} \approx \lambda^*. \tag{7}$$

Aby znaleziony punkt był rzeczywistym minimum, iloraz ($F[\lambda_0, \lambda_1, \lambda_2]$) musi spełniać warunek:

$$F[\lambda_0, \lambda_1, \lambda_2] > 0.$$

Algorytm interpolacji Powella składa się z następujących kroków:

1. Wybierz λ_0 i oblicz $F[\lambda_0 + h] < F[\lambda_0]$, $F[\lambda_0 + 2h] < F[\lambda_0 + h]$ (ewentualnie zmień znak: -h, jeśli nierówności nie są spełnione).
2. Wyznacz λ_m i sprawdź, czy jest minimum.
3. Jeśli $|\lambda_m - \lambda_n| > h$, odrzuć najdalej położony od λ_m punkt i ponownie wykonaj obliczenia z pkt. 2.
   (λ_n – najbliżej położony punkt względem λ_m).

Punkt λ_m akceptujemy jako minimum, jeśli

$$|\lambda_m - \lambda_n| < \epsilon.$$

Algorytm Powella znajduje minimum szybciej niż metoda złotego podziału, jeśli $f(x)$ nie jest skrzywiona. Jednak dla mocno niesymetrycznych funkcji, metoda złotego podziału pozostaje lepsza.

2 Zadanie do wykonania

2.1 Opis problemu

Metoda interpolacji Powella służy do lokalnego przybliżenia funkcji za pomocą wielomianu **stopnia drugiego**. W celu znalezienia lokalnego minimum tej funkcji, wykorzystujemy trzy punkty: x_1, x_2, x_3, oraz ich odpowiadające wartości funkcji: $f(x_1)$, $f(x_2)$, $f(x_3)$.
Zakładamy, że wartości funkcji maleją wraz z rosnącymi indeksami punktów. Na podstawie tych punktów możemy obliczyć lokalne minimum zgodnie z poniższym wzorem:

$$x_m = \frac{x_1 + x_2}{2} - \frac{F[x_1, x_2]}{2F[x_1, x_2, x_3]}, \tag{8}$$

gdzie:

- x_m to wartość, w której szukamy lokalnego minimum,
- $F[x_1, x_2]$ to iloraz różnicowy pierwszego rzędu między punktami x_1 i x_2,
- $F[x_1, x_2, x_3]$ to iloraz różnicowy drugiego rzędu, obliczony na podstawie trzech punktów x_1, x_2 i x_3.

Warto zauważyć, że iloraz różnicowy pierwszego rzędu, $F[x_1, x_2]$, oblicza się jako stosunek zmiany wartości funkcji do zmiany argumentu między punktami x_1 i x_2. Natomiast iloraz różnicowy drugiego rzędu, $F[x_1, x_2, x_3]$, jest wyznaczany poprzez różnicowanie ilorazu różnicowego pierwszego rzędu.

$$F[x_1, x_2] = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \tag{9}$$

$$F[x_1, x_2, x_3] = \frac{\frac{f(x_3)-f(x_2)}{x_3-x_2} - \frac{f(x_2)-f(x_1)}{x_2-x_1}}{x_3 - x_1}. \tag{10}$$

Ostatecznie, wyznaczenie x_m pozwala nam znaleźć lokalne minimum funkcji.

2.2 Wykonanie zadania

Zadaniem było zaimplementowanie metody Powella w celu znalezienia minimum dwóch funkcji:

$$f_1(x) = \ln(x^5 + 3x^2 + x + 9) \tag{11}$$

oraz

$$f_2(x) = x^6. \tag{12}$$

Dla $f(x_1)$ ustalono punkty startowe:

$$x_1 = -0.5,$$
$$x_2 = x_1 + h,$$
$$x_3 = x_2 + h,$$

gdzie h przyjął wartość 0.01.
Następnie przeprowadzono obliczenia dla kolejnych 10 przybliżeń położenia minimum, zapisując wyniki do pliku.
Kolejne obliczenia wykonano dla punktów startowych:

$$x_1 = -0.9,$$
$$x_2 = x_1 + h,$$
$$x_3 = x_1 + 2h,$$

przy $h = 0.01$.

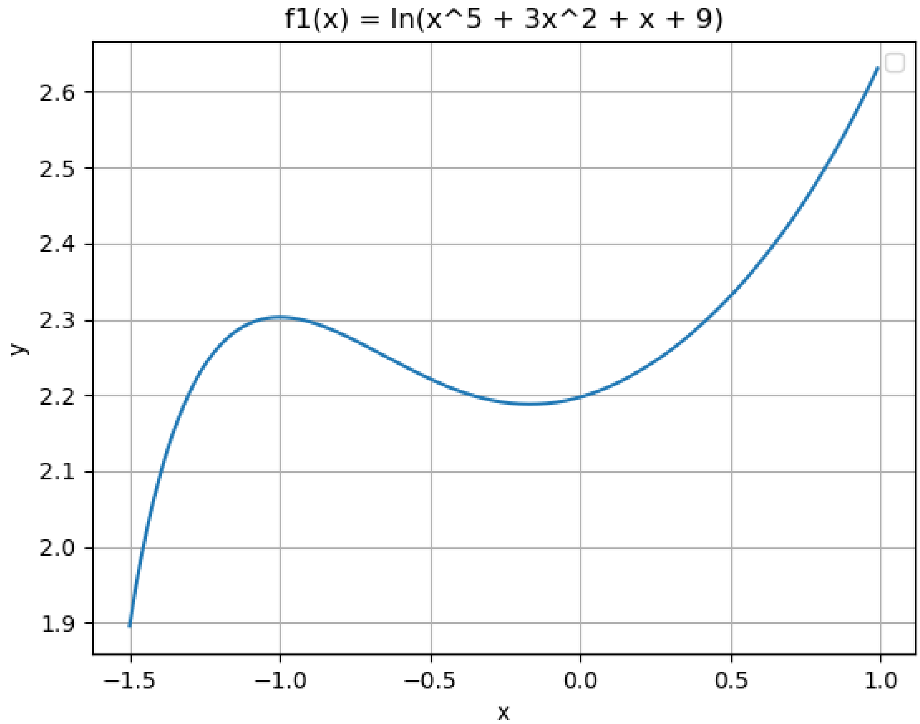
Dla funkcji $f(x_2)$ jako punkty startowe przyjęto

$$x_1 = 1.5,$$
$$x_2 = x_1 + h,$$
$$x_3 = x_2 + h,$$

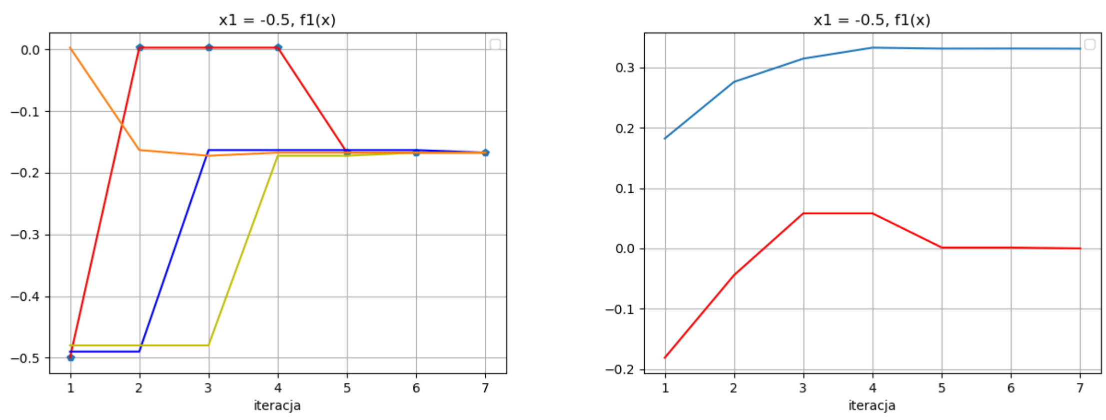
przy $h = 0.01$.
W pliku zapisano 100 przybliżeń położenia funkcji.
Za dostateczne przybliżenie przyjęto $\epsilon = 10^{-7}$.

2.3 Wyniki

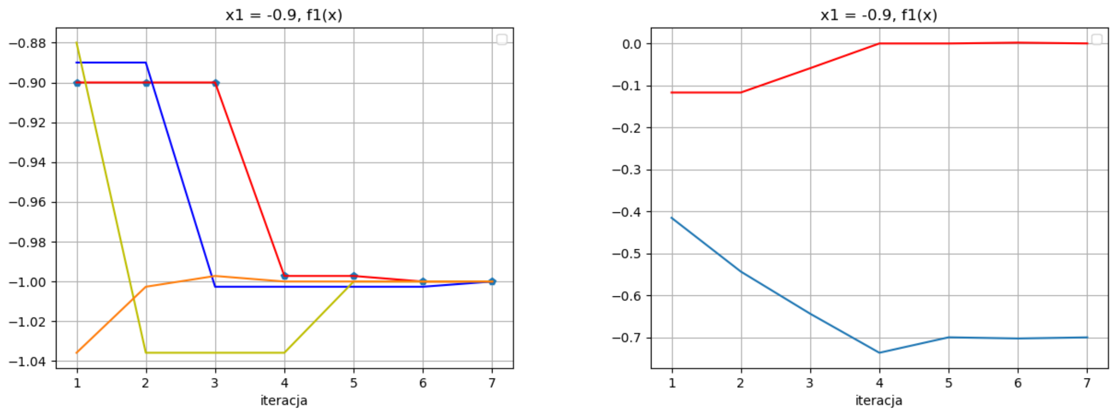
Na podstawie wyników zawartych w plikach tekstowych zostały wygenerowane następujące wykresy:



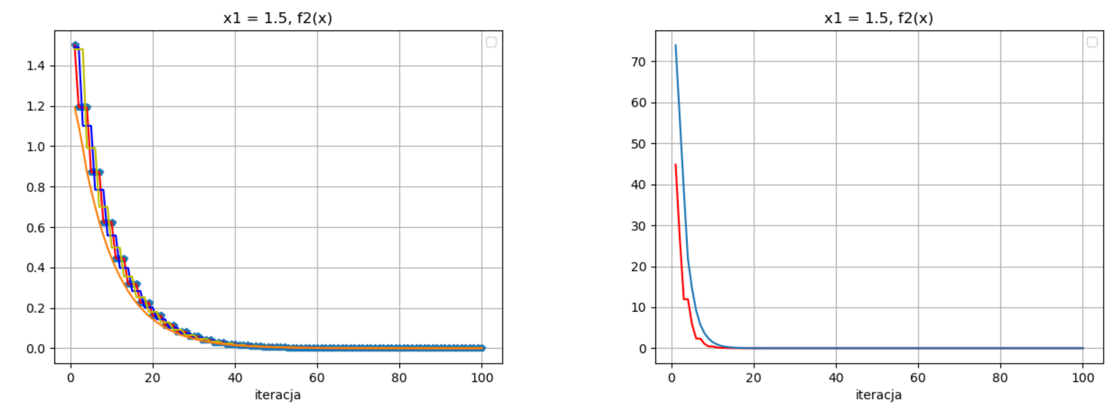
Rysunek 2: Wykres funkcji $f_1(x) = \ln(x^5 + 3x^2 + x + 9)$



Rysunek 3: Kolejne przybliżenia położenia minimum funkcji $f_1(x)$ oraz ilorazów różnicowych dla $x_1 = -0.5$ i $\epsilon = 10^{(-7)}$.



Rysunek 4: Kolejne przybliżenia położenia minimum funkcji $f_1(x)$ oraz ilorazów różnicowych dla $x_1 = -0.9$ i $\epsilon = 10^{(-7)}$.



Rysunek 5: Kolejne przybliżenia położenia minimum funkcji $f_2(x)$ oraz ilorazów różnicowych dla $x_1 = 1.5$ i $\epsilon = 10^{(-7)}$.

3 Wnioski

- **Analiza wartości ilorazów:** Dla funkcji $f_1(x)$, wybierając drugi zestaw punktów startowych, iloraz różnicowy $F[x_1, x_2, x_3]$ jest ujemny, co oznacza, że wartości funkcji maleją wraz ze zmianą argumentów. W przypadku, gdy iloraz ten jest ujemny, według algorytmu Powella, szukamy maksimum funkcji, a nie minimum. Zjawisko to wynika z kształtu funkcji oraz zachowania się wartości ilorazów różnicowych w okolicy punktu startowego.

- **Wolnozbieżność metody dla funkcji $f_2(x)$:** Metoda jest wolnozbieżna dla funkcji $f_2(x)$ ze względu na dużą różnicę w wartościach między kolejnymi iteracjami. Funkcja $f_2(x)$, będąca funkcją wielomianową szóstego stopnia, charakteryzuje się nagłymi zmianami wartości w okolicy punktów ekstremalnych, co utrudnia zbieżność metody. Najbardziej odpowiednim warunkiem stopu dla takich funkcji może być osiągnięcie maksymalnej liczby iteracji, gdyż metoda może nie zbiegać w rozsądnym czasie.

- **Rozstrzyganie istnienia minimum/maksimum za pomocą drugiej pochodnej:** Dla funkcji $f_2(x)$, istnienie minimum/maksimum w danym punkcie decyduje druga pochodna. Gdy druga pochodna jest dodatnia, mamy do czynienia z minimum lokalnym, natomiast gdy jest ujemna, jest to maksimum lokalne. W przypadku $f_2(x)$, druga pochodna w punkcie $x = 0$ wynosi 0, co oznacza, że nie możemy jednoznacznie określić, czy mamy do czynienia z minimum czy maksimum.

Funkcja $f(x) = x^6$ ma minimum w punkcie $x = 0$, ponieważ jest to punkt przegięcia dla tej funkcji. Można to zobaczyć poprzez analizę wykresu funkcji $f(x) = x^6$. W okolicy punktu $x = 0$, funkcja ta jest rosnąca po lewej stronie oraz malejąca po prawej stronie, co oznacza, że w punkcie $x = 0$ występuje minimum lokalne. Jednakże, w rzeczywistości, to minimum jest również globalne, ponieważ funkcja x^6 zawsze jest dodatnia, więc nie może być mniejsza niż 0.

Funkcja $f(x) = x^6$ nie ma maksimum. Można to również zauważyć, analizując jej wykres. Funkcja x^6 jest zawsze nieujemna, więc nie może osiągnąć maksimum, ponieważ nie ma górnego ograniczenia. Może ona przyjmować coraz większe wartości, gdy x zbliża się do nieskończoności dodatniej. Dlatego nie istnieje punkt, w którym funkcja x^6 osiąga maksimum globalne.