WSI ćwiczenie 1

Wojciech Pobocha

Październik 2023

1 Wykresy funkcji

Funkcje:

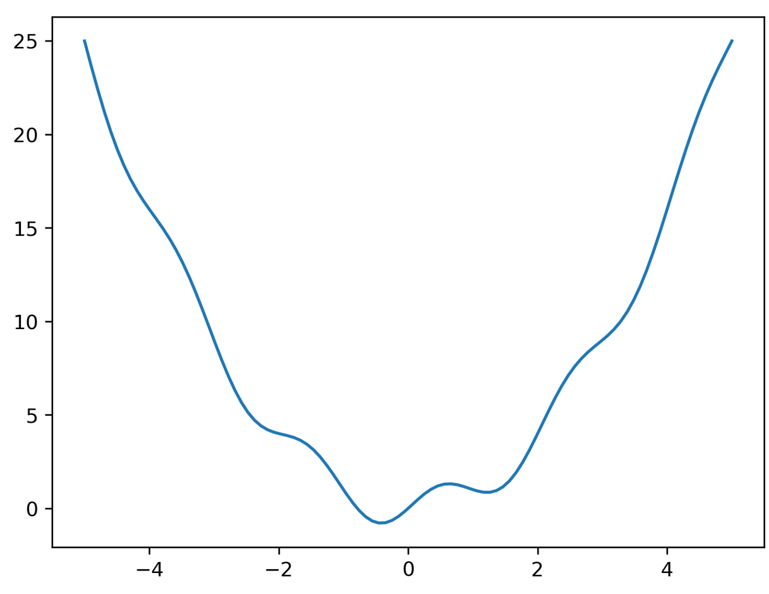
$$f(x) = \sin(\pi x) + x^2$$

$$g(x) = 5e^2 - 4ex_1 + x_1^2 + 2ex_2 + x_2^2$$

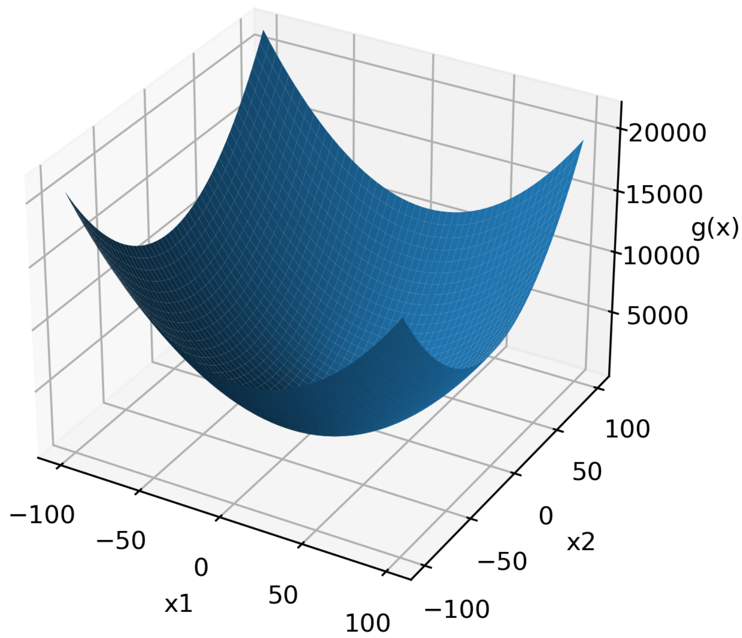
Gradienty funkcji:

$$\nabla f(x) = \pi \cos(\pi x) + 2x$$

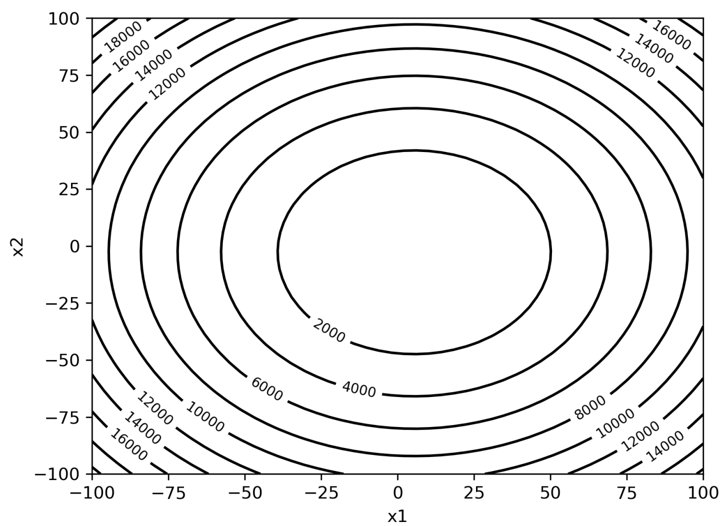
$$\nabla g(x) = \begin{bmatrix} 2(x_1 - 2e) \\ 2(x_2 + e) \end{bmatrix}$$



Rysunek 1: Wykres funkcji f(x)



Rysunek 2: Wykres funkcji 3D g(x_1, x_2)



Rysunek 3: Wykres konturowy g(x_1, x_2)

Analizując obie funkcje wydaje się, że prostszą w optymalizacji pomimo większej ilości parametrów będzie g. Posiada ona tylko jedno minimum lokalne, przez co łatwo będzie je zlokalizować.

2 Algorytm i implementacja

Wzór:

$$x(k+1) = x(k) - \alpha \nabla f(x(k))$$

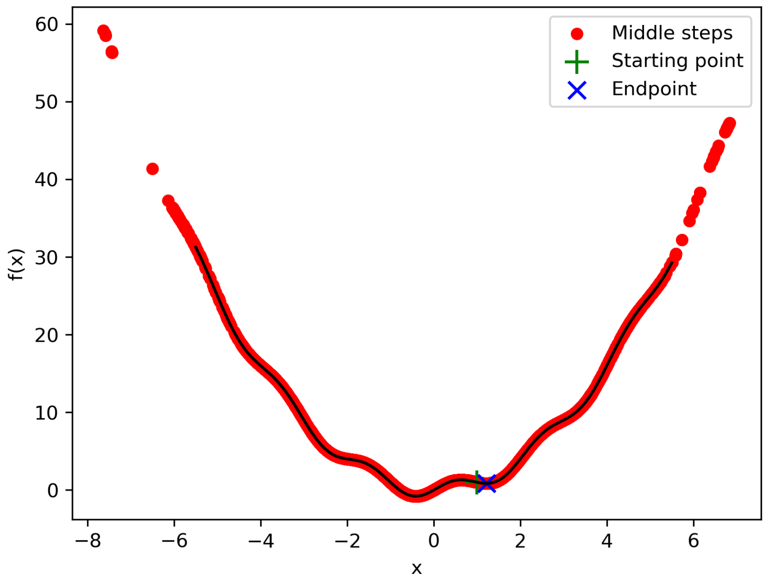
Warunek stopu:

$$||\nabla f(x(k))|| < \epsilon$$

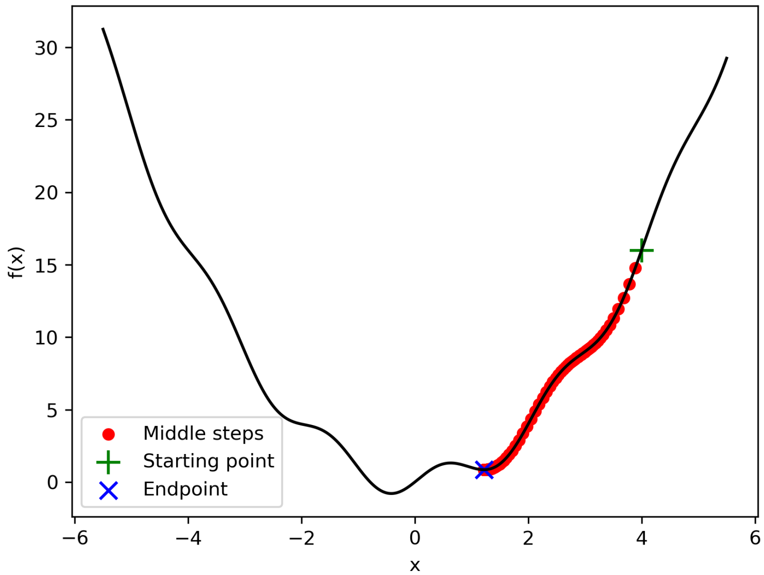
α - stały parametr regulowany
k - numer kroku algorytmu
ϵ - dopuszczalna wartość gradientu
x - wektor argumentów
$\nabla f(x(k))$ - gradient funkcji

3 Wpływ parametrów i punktów początkowych na działanie algorytmu

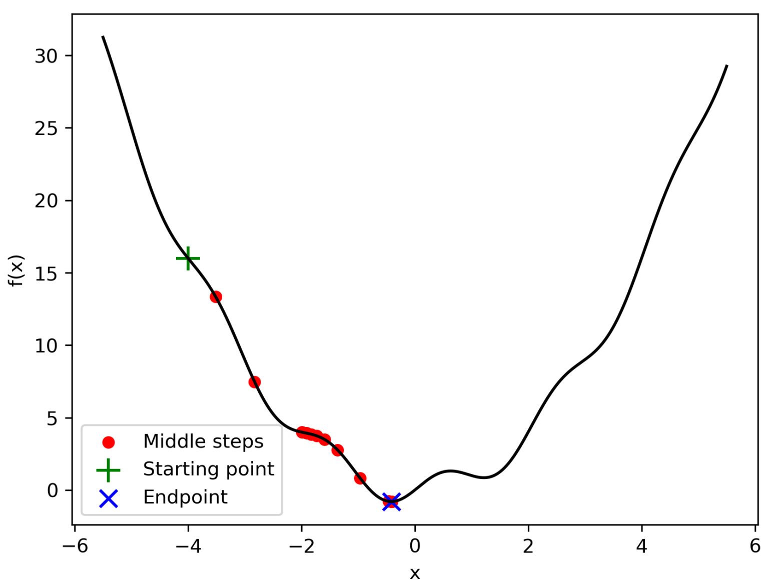
3.1 Funkcja $f(x)$



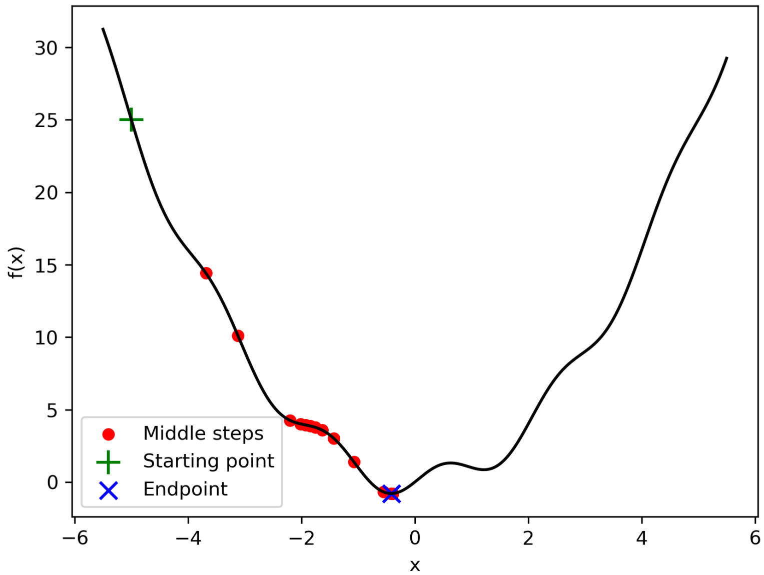
Rysunek 4: Wykres funkcji f(x), $\alpha = 0.9$



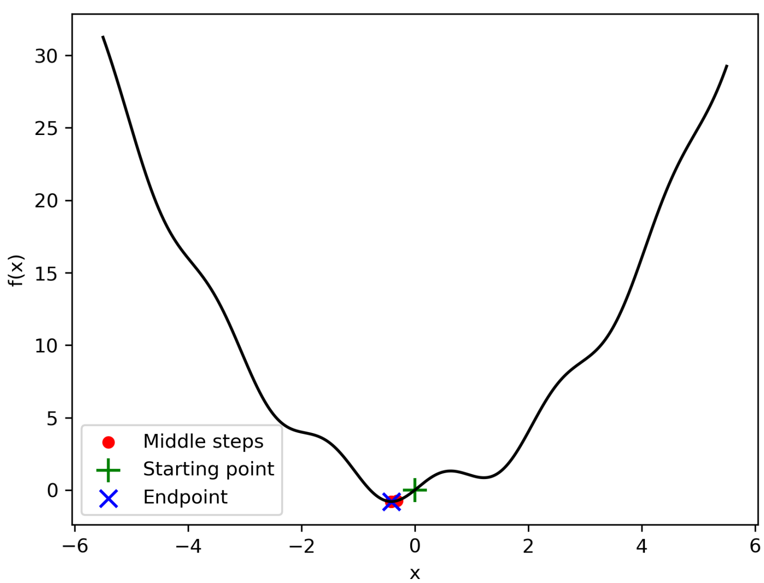
Rysunek 5: Wykres funkcji f(x), $\alpha = 0.01$



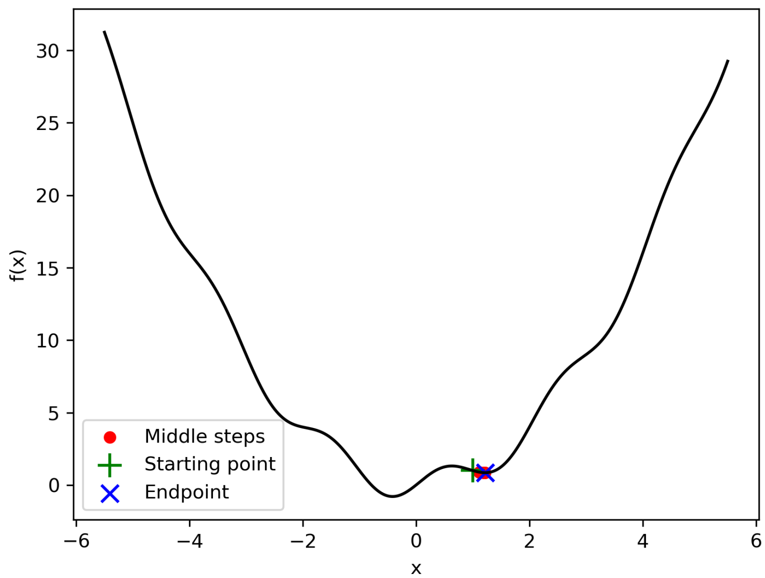
Rysunek 6: Wykres funkcji f(x), $\alpha = 0.01$



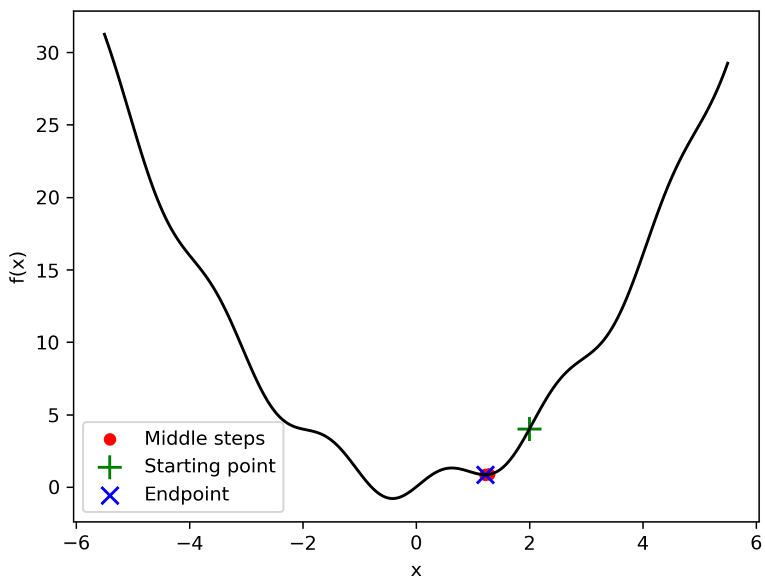
Rysunek 7: Wykres funkcji f(x), $\alpha = 0.1$



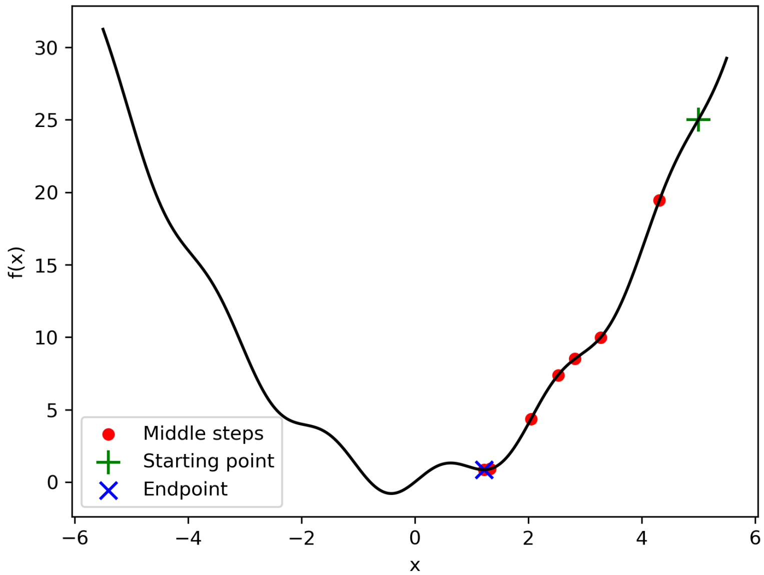
Rysunek 8: Wykres funkcji f(x), $\alpha = 0.1$



Rysunek 9: Wykres funkcji f(x), $\alpha = 0.1$



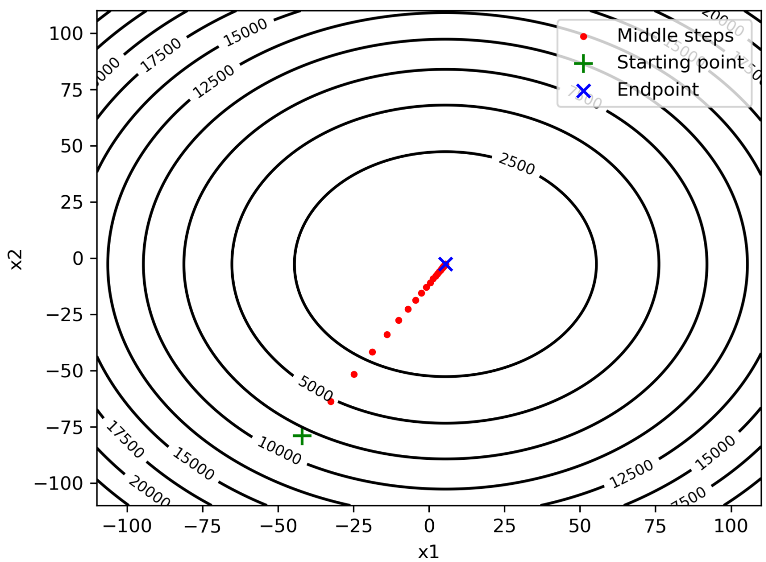
Rysunek 10: Wykres funkcji f(x), $\alpha = 0.1$



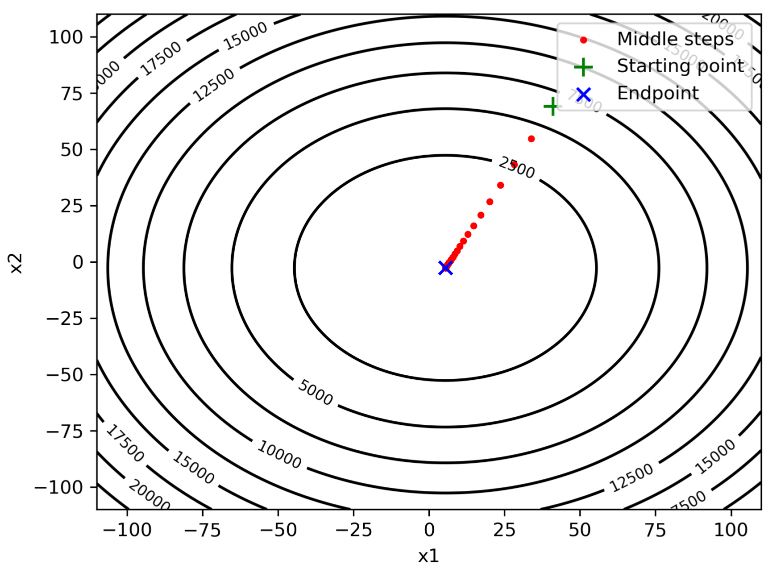
Rysunek 11: Wykres funkcji f(x), $\alpha = 0.1$

Najlepszym parametrem dla $f(x)$ z powyższych testów wydaje się być $\alpha = 0.1$. Przy zbyt dużej α algorytm robi zbyt duże kroki, przez co czasami potrafi "przeskoczyć" minimum funkcji, które ma zlokalizować. Zbyt niska wartość tego parametru powoduje spowolnienie algorytmu, ponieważ musi on wykonać więcej kroków, aby znaleźć minimum. Przy ustawieniu $\alpha = 1$ algorytm w ogóle nie znajduje minimum, a raczej nie wiadomo kiedy mógłby je znaleźć, ponieważ po zbyt długim czasie działania wyłączałem go. Algorytm dla tej funkcji nie zawsze jest w stanie znaleźć minimum globalne, ze względu na to, że nic nie zabezpiecza go przed zatrzymywaniem się w minimum lokalnym, ponieważ w tym miejscu również $\nabla f(x) < \epsilon$.

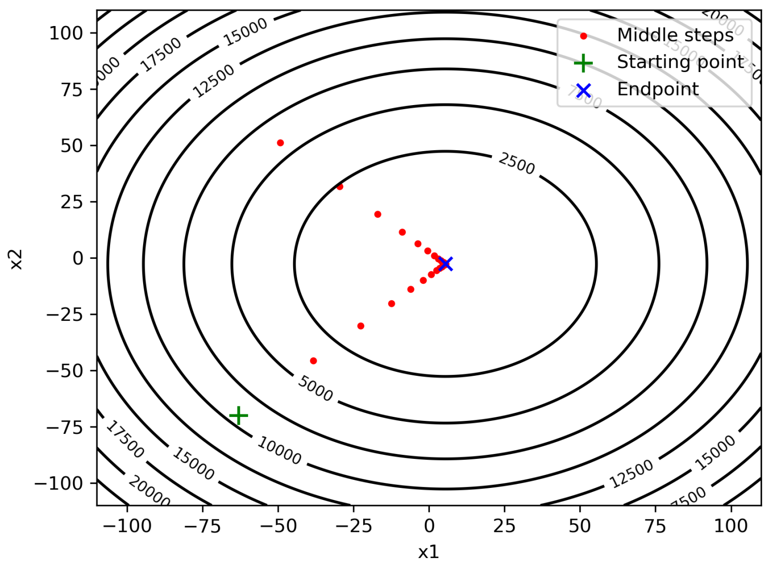
3.2 Funkcja $g(x_1, x_2)$



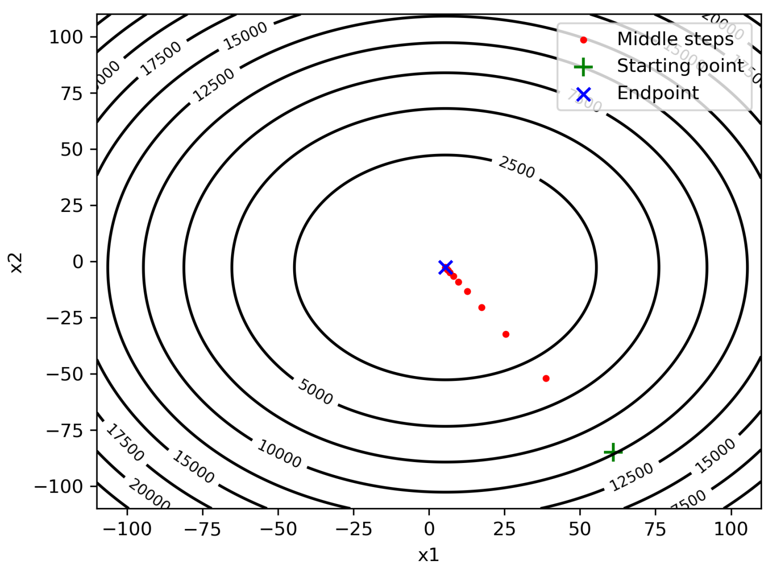
Rysunek 12: Wykres funkcji $g(x_1, x_2), \alpha = [0.1, 0.1]$



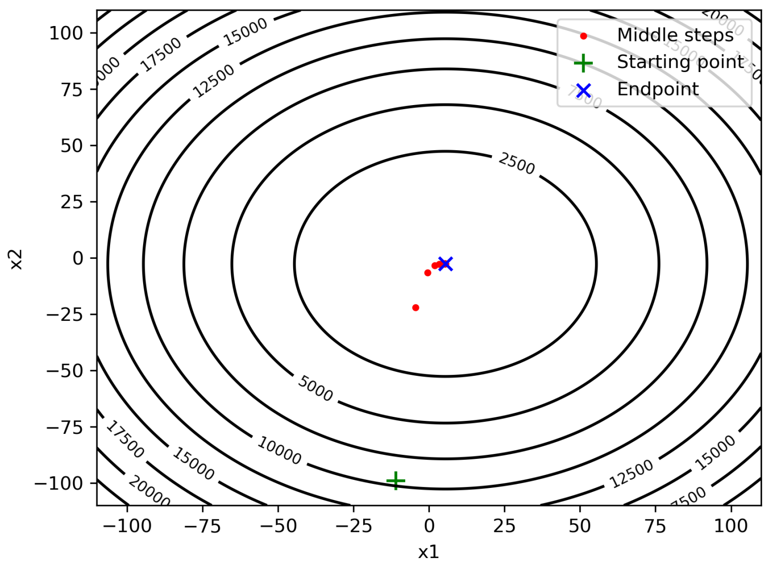
Rysunek 13: Wykres funkcji $g(x_1, x_2), \alpha = [0.1, 0.1]$



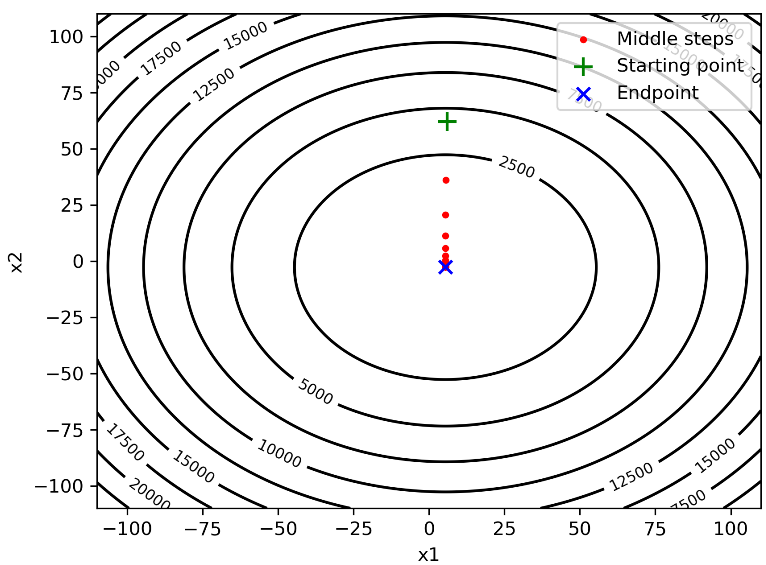
Rysunek 14: Wykres funkcji $g(x_1, x_2), \alpha = [0.1, 0.9]$



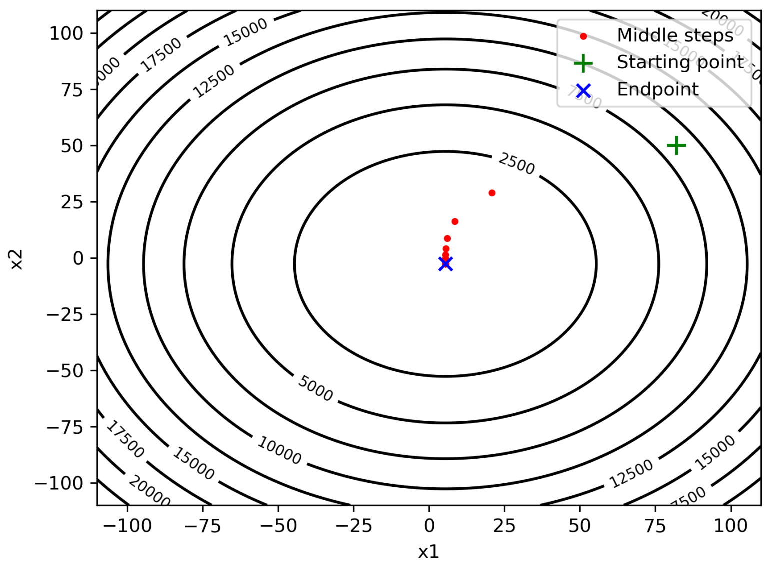
Rysunek 15: Wykres funkcji $g(x_1, x_2), \alpha = [0.2, 0.2]$



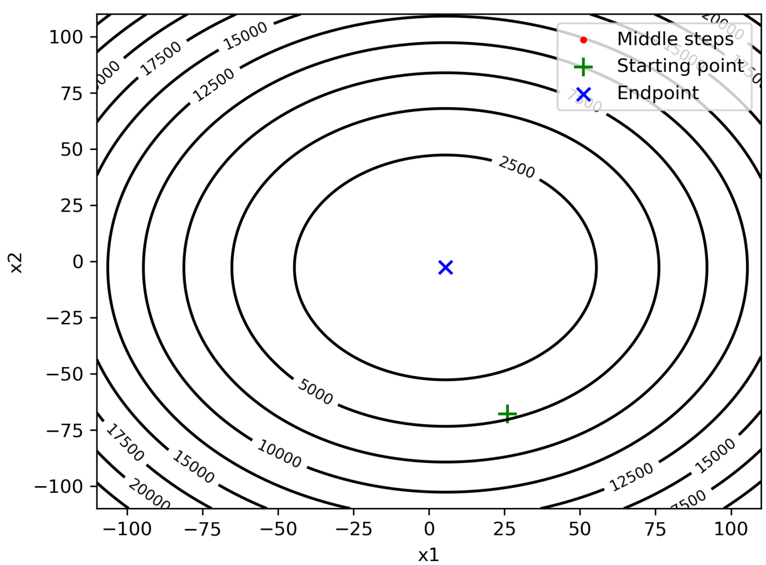
Rysunek 16: Wykres funkcji $g(x_1, x_2), \alpha = [0.2, 0.4]$



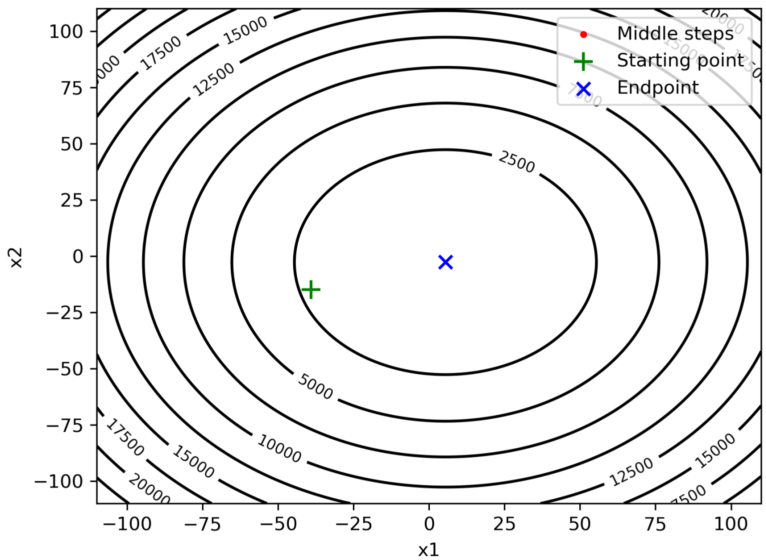
Rysunek 17: Wykres funkcji $g(x_1, x_2), \alpha = [0.4, 0.2]$



Rysunek 18: Wykres funkcji $g(x_1, x_2), \alpha = [0.4, 0.2]$



Rysunek 19: Wykres funkcji $g(x_1, x_2), \alpha = [0.5, 0.5]$



Rysunek 20: Wykres funkcji $g(x_1, x_2), \alpha = [0.5, 0.5]$

Dla funkcji $g(x_1, x_2)$ najlepszy parametr to $\alpha = [0.5, 0.5]$, w każdym drugim kroku algorytmu jest znajdowany punkt, który spełnia warunki zatrzymania. Algorytm przy zmianie parametru α dla tej funkcji zachowuje się podobnie jak dla $f(x)$ - gdy ustawiony jest za duży algorytm "przeskakuje" minima, natomiast za mały powoduje wzrost liczby iteracji do znalezienia rozwiązania. Dla $g(x_1, x_2)$ algorytm zawsze znajduje minimum globalne, ponieważ istnieje tylko jedno.

4 Podsumowanie i ocena algorytmu

Po zbadaniu obu funkcji, algorytm okazuje się efektywny w znajdowaniu minimów, jednak w zależności od funkcji czasami nie jest w stanie znaleźć minimum globalnego. Przy dobraniu odpowiednich parametrów za każdym razem znajdzie rozwiązanie (globalne lub lokalne). Posiada mało parametrów do dobrania, a czas i ilość iteracji jest niska w stosunku do tego jak dużą przestrzeń musi przeszukać. Jednym z jego minusów pozostaje, jednak to, że $\nabla funkcji()$ musi być znany.
Podsumowując, jest to idealny algorytm do znajdowania minimum globalnego w momencie gdy istnieje tylko jedno, a $\nabla funkcji()$ jest znany.