

## Ćw. 1

Zaimplementować metodę gradientu prostego dla funkcji jednej zmiennej. Zbadać działanie metody w zależności od parametrów wejściowych: - punkt startowy - współczynnika uczenia Eksperymenty przeprowadzić dla funkcji z jednym minimum oraz dla funkcji z minimum lokalnym.

## Założenia wstępne

Metoda gradientu prostego służy do iteracyjnego odnajdywania minimum lokalnego zadanej funkcji. Przed użyciem tej metody należy zdefiniować parametry takie jak:

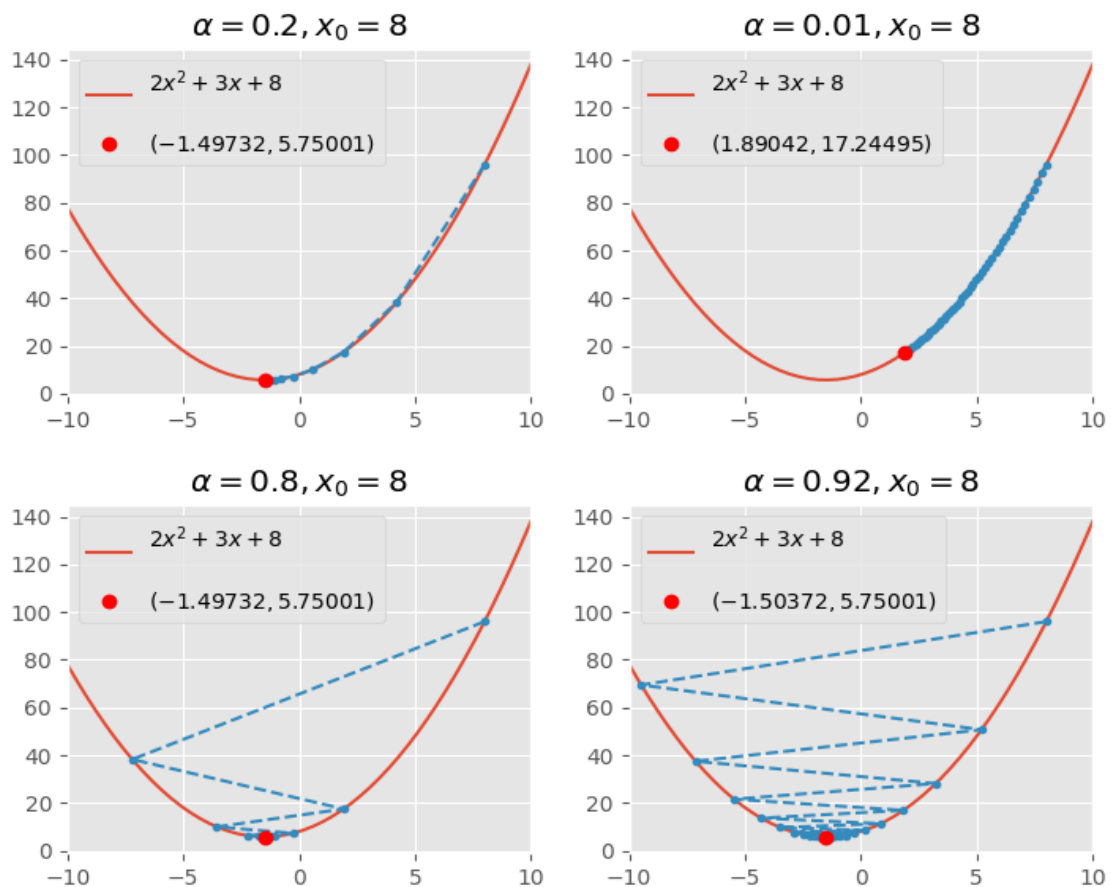
$\alpha$  – współczynnik uczenia,

$\varepsilon$  – zadana precyzja, przyjęto  $\varepsilon = 0,01$ ,

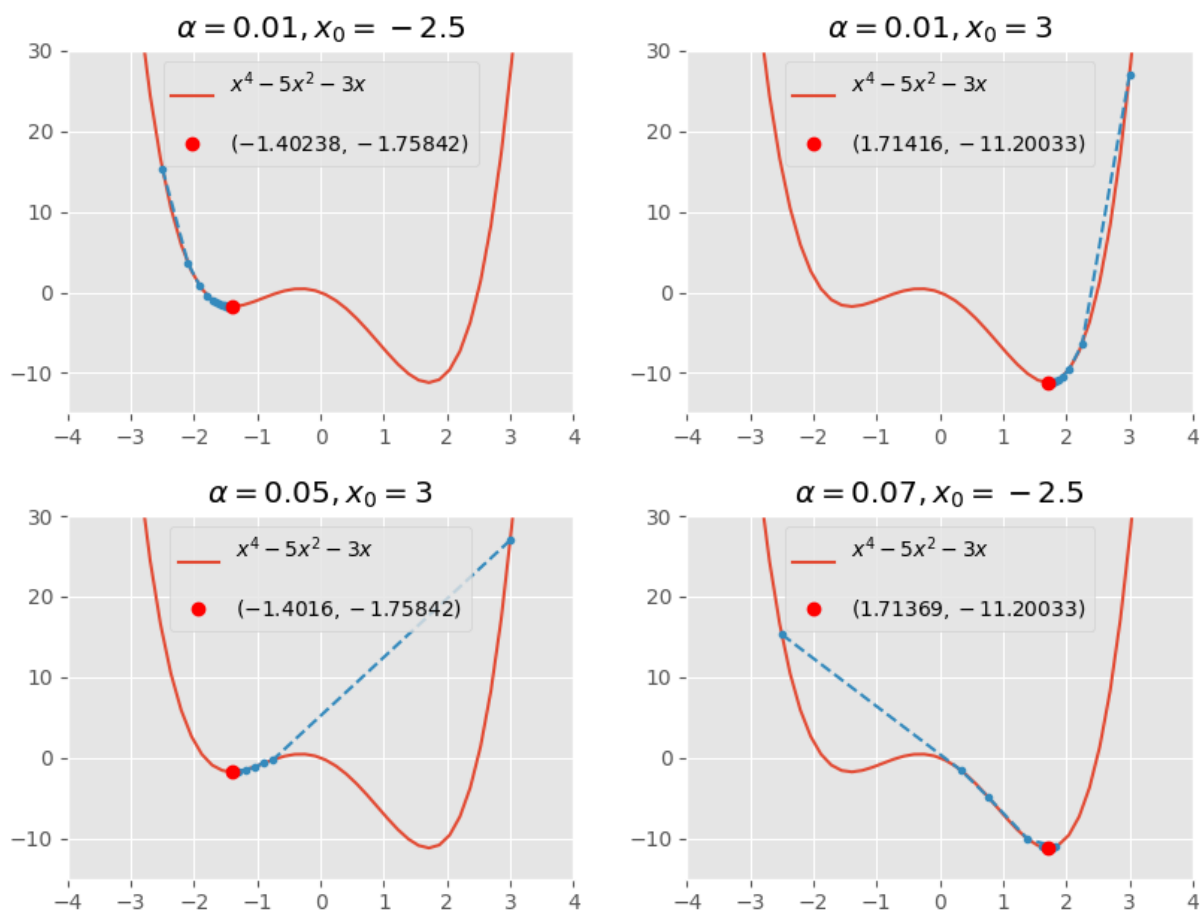
$x_0$  – punkt startowy.

W celu określenia, czy punkt wygenerowany w danej iteracji dostatecznie dobrze przybliży minimum funkcji użyto następującego kryterium stopu (test stacjonarności):  $|\nabla f(x_i)| \leq \varepsilon$  oraz ustawiono limit iteracji na 50.

## Działanie algorytmu w zależności od współczynnika uczenia i punktu startowego



Rys 1 – Kolejne iteracje algorytmu przedstawione graficznie dla różnych  $\alpha$



Rys 2 – Kolejne iteracje algorytmu przedstawione graficznie dla różnych  $\alpha$  i  $x_0$

Wykonano eksperymenty dla różnych wartości współczynnika uczenia, punktu początkowego i różnych funkcji. Wyniki przedstawiono na rysunkach 1 i 2. Trasę algorytmu i punkty wyznaczone w kolejnych iteracjach zaznaczono kolorem niebieskim. Kolorem czerwonym zaznaczono znalezione minimum funkcji.

### Komentarze i wnioski

Na rysunku 1 optymalną wartością parametru  $\alpha$  jest  $0,2$ . Minimum funkcji zostało znalezione poprawnie. Z eksperymentów wynika, że zbyt mała wartość współczynnika uczenia skutkuje tym, że kroki algorytmu są małe, a w związku z tym algorytm będzie działał dłużej. Dla  $\alpha=0,01$  znalezione minimum nie jest poprawne ponieważ algorytm napotkał limit iteracji i zakończył działanie przed znalezieniem minimum. Zwiększanie wartości  $\alpha$  może zmniejszyć ilość iteracji algorytmu, ale zbyt duża wartość powoduje, że algorytm zaczyna generować zbyt duże kroki i zachowuje się niestabilnie (tak jak na rysunku 1 dla  $\alpha=0,92$ ), chociaż znalezione minimum jest poprawne. Dalsze zwiększanie  $\alpha$  spowoduje, że algorytm nie zadziała poprawnie, ponieważ nie będzie miał szans trafić w minimum z tak dużym krokiem.

Metoda gradientu prostego służy do odnajdywania minimum lokalnego, a nie globalnego, więc przy rozsądnym doborze współczynnika uczenia znalezione minimum będzie zależało od punktu początkowego i będzie tym najbliższym punktu początkowego. Na rysunku 2 pokazano, że rzeczywiście tak jest. Minimum jest inne dla  $x_0=-2,5$  i  $x_0=3$ . Przy zbyt dużym współczynniku uczenia istnieje ryzyko przeskoczenia najbliższego minimum i znalezieniu innego (tak jak na rysunku 2 dla  $\alpha=0,05$  i  $\alpha=0,07$ ).