Piotr Szmurło 303785 **WSI21Z**

Ćw 1. (7 pkt), data oddania: do 18.10.2021 - Zagadnienie przeszukiwania i podstawowe podejścia do niego 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, czyli np.: $x^2 +3x+8$, $x^4 - 5x^2 - 3x$

Nie trzeba implementować liczenia pochodnej z funkcji wejściowej - podajemy jako już znaną funkcję, hint: f(x) i ∇f(x) najlepiej przekazać jako argument funkcji np.: # lambda x: x ** 2 # lambda gx: 2 * gx

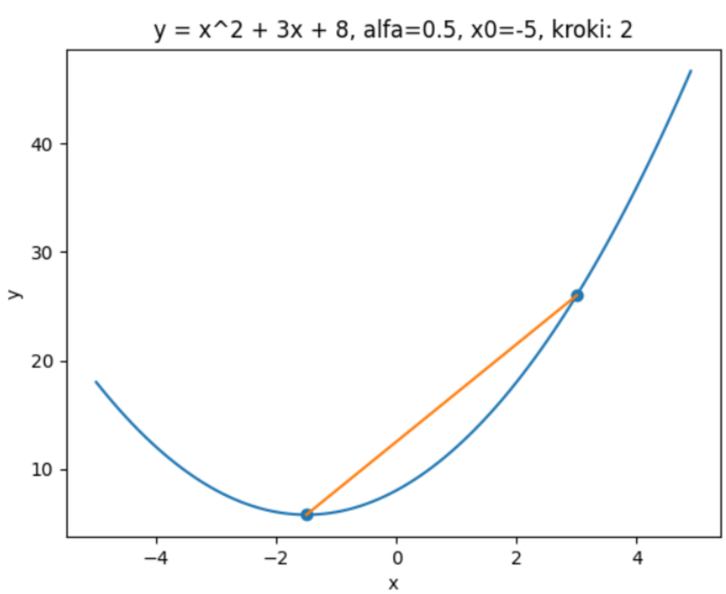
Testy:

Dla funkcji $x^2 +3x+8$:

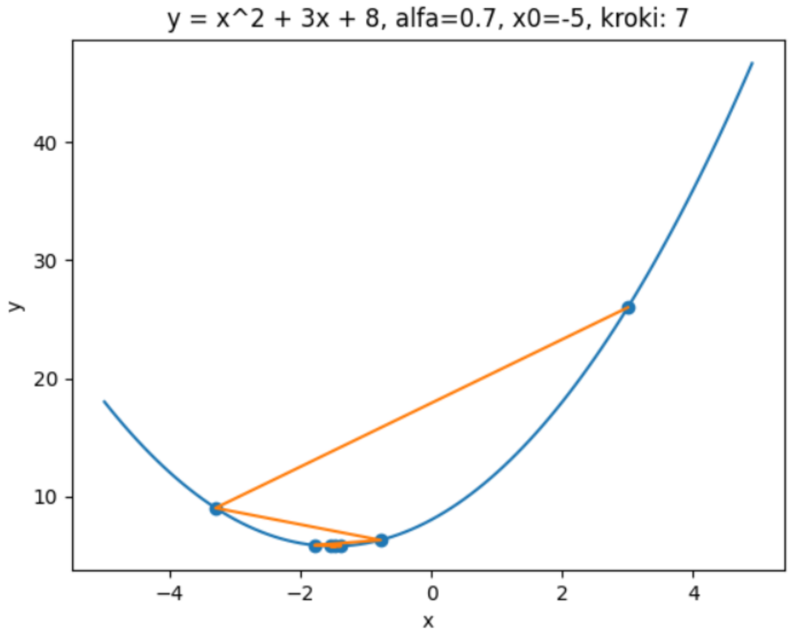
Na początku znalazłem optymalny współczynnik uczenia alfa dla konkretnego x_0:

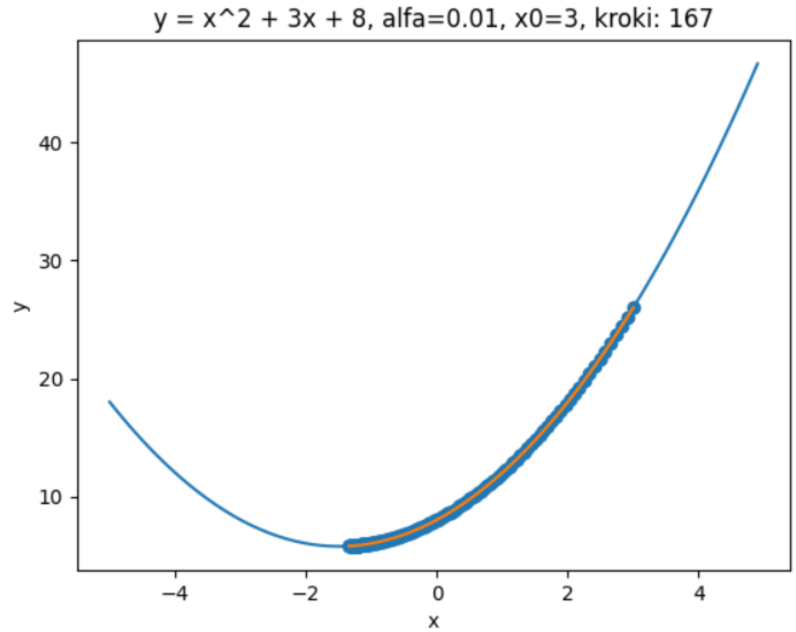
```
wynik: -1.5709644585665583; alfa: 0.05; kroki: 38
wynik: -1.5504403158265496; alfa: 0.1; kroki: 20
wynik: -1.53391115364245; alfa: 0.15000000000000002; kroki: 14
wynik: -1.535271936; alfa: 0.2; kroki: 10
wynik: -1.52734375; alfa: 0.25; kroki: 8
wynik: -1.514336; alfa: 0.30000000000000004; kroki: 7
wynik: -1.52835; alfa: 0.35000000000000003; kroki: 5
wynik: -1.528; alfa: 0.4; kroki: 4
wynik: -1.5035; alfa: 0.45; kroki: 4
wynik: -1.5; alfa: 0.5; kroki: 2
wynik: -1.4965; alfa: 0.55; kroki: 4
wynik: -1.472; alfa: 0.6000000000000001; kroki: 4
wynik: -1.52835; alfa: 0.65; kroki: 5
wynik: -1.5143360000000001; alfa: 0.7000000000000001; kroki: 7
wynik: -1.47265625; alfa: 0.75; kroki: 8
wynik: -1.464728064; alfa: 0.8; kroki: 10
wynik: -1.4660888463575499; alfa: 0.8500000000000001; kroki: 14
wynik: -1.4495596841734504; alfa: 0.9; kroki: 20
wynik: -1.429035541433441; alfa: 0.9500000000000001; kroki: 38
```

Najlepsze rezultaty otrzymałem przy użyciu alfa=0.5, niezależnie od wybranego punktu startowego algorytm znajduje minimum tej funkcji w dwóch krokach (licząc punkt startowy x_0)



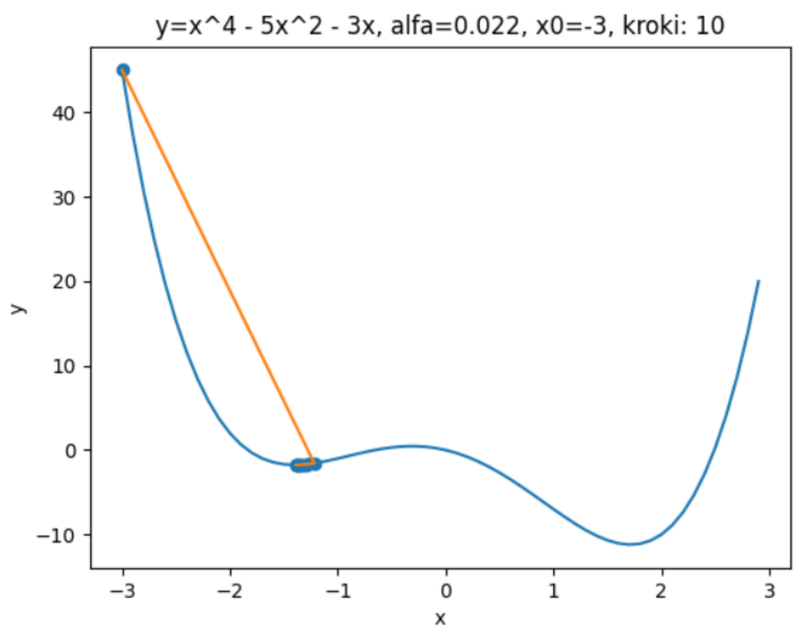
Dla większych wartości współczynnika uczenia (np. 0.7) krok jest zbyt duży, przez co pojawiają się oscylacje.

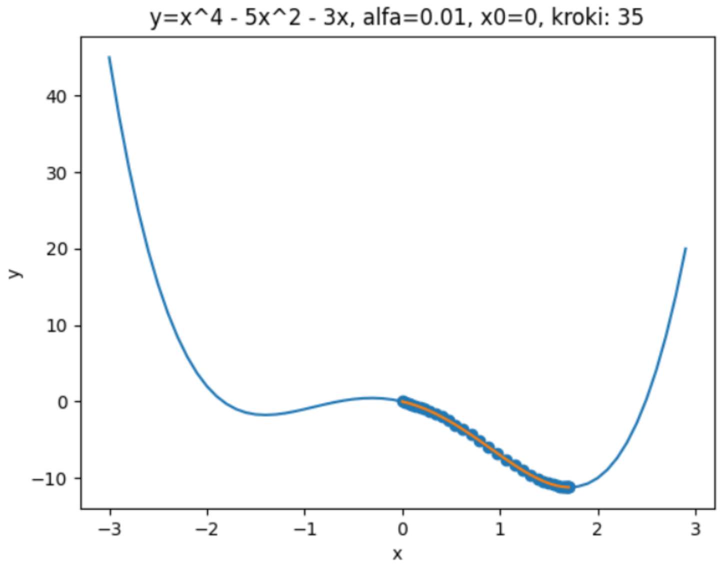


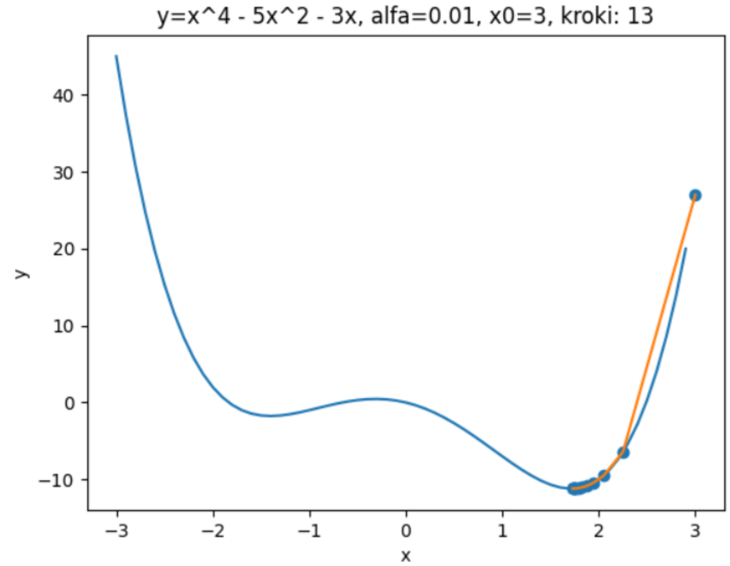


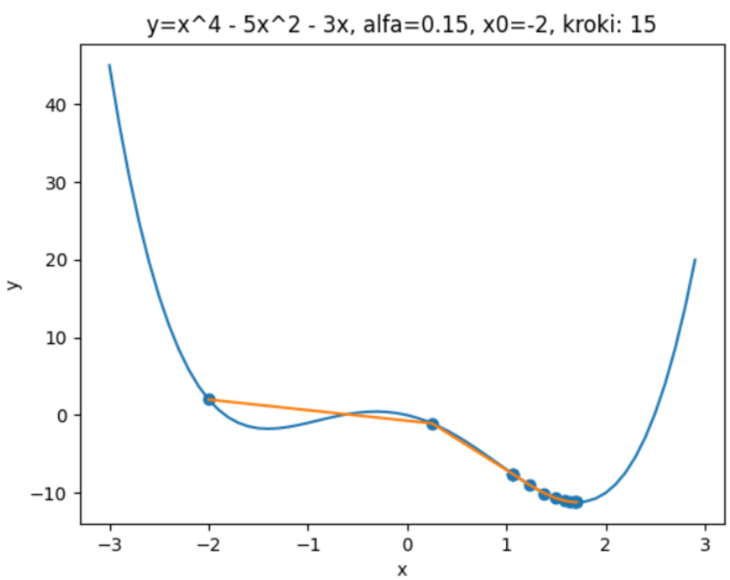
Dla mniejszych wartości alfa kroki są mniejsze, więc algorytm wolniej zbiega do minimum.

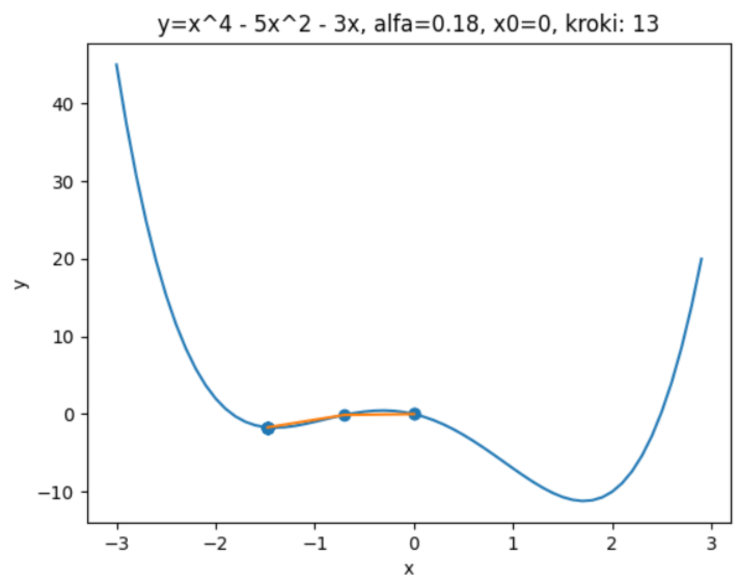
Dla funkcji $x^4 - 5x^2 - 3x$



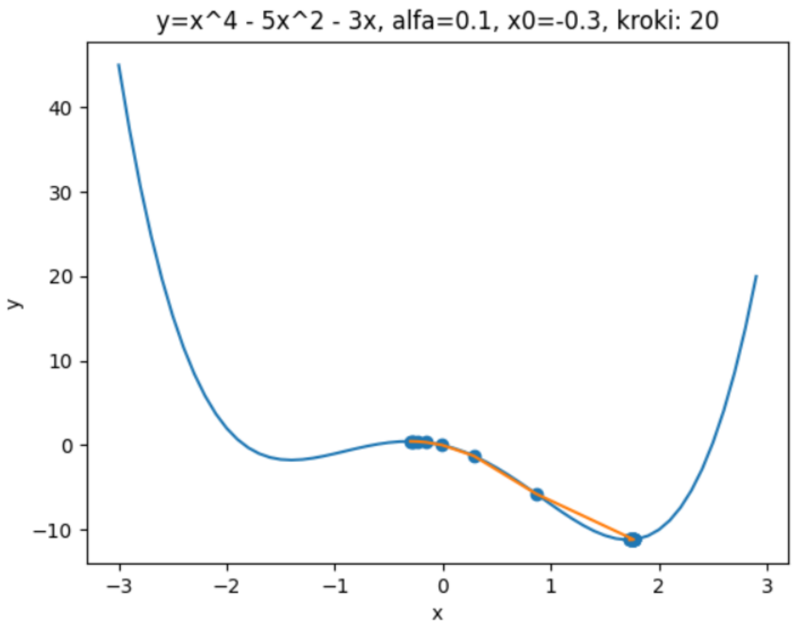

y=x^4 - 5x^2 - 3x, alfa=0.01, x0=0, kroki: 35
y
x
−3 −2 −1 0 1 2 3
−10 0 10 20 30 40

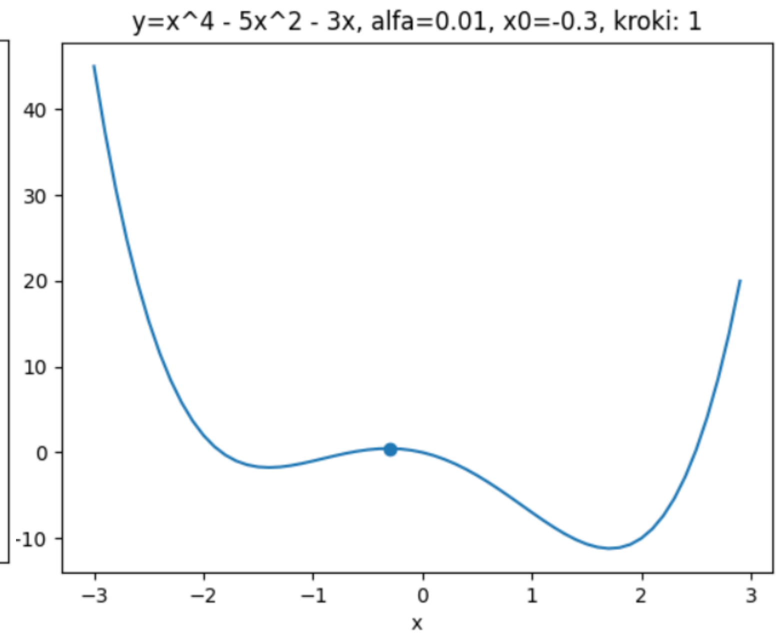

y=x^4 - 5x^2 - 3x, alfa=0.01, x0=3, kroki: 13
y
x
−3 −2 −1 0 1 2 3
−10 0 10 20 30 40


y=x^4 - 5x^2 - 3x, alfa=0.15, x0=-2, kroki: 15
y
x
−3 −2 −1 0 1 2 3
−10 0 10 20 30 40


y=x^4 - 5x^2 - 3x, alfa=0.18, x0=0, kroki: 13
y
x
−3 −2 −1 0 1 2 3
−10 0 10 20 30 40

Algorytm znacznie trudniej sobie radzi z trudniejszymi funkcjami, gdzie jest wiele ekstremów. Gdy punkt startowy jest w okolicy maksimum lokalnego oraz alfa jest zbyt małe, algorytm nie znajduje minimum I "wpada" w bliskie temu punktowi maksimum lokalne, gdyż gradient jest zbyt mały; należy wtedy zwiększyć współczynnik uczenia alfa. Zbyt duży współczynnik powoduje rozbieżność algorytmu.





Wnioski:

Metoda gradientu prostego nie zawsze znajdzie ekstremum funkcji bardziej złożonych (w wielu kombinacjach punktu startowego i współczynnika uczenia minimum nie zostaje znalezione lub ze sporym błędem) oraz jest w stanie znaleźć tylko jedno ekstremum – zwykle te najbliższe punktu startowego, chyba że krok będzie na tyle duży, że je "przeskoczy". Algorytm nie jest dobry do szukania minimów funkcji gdzie nachylenie (gradient) jest bardzo małe (np. w okolicy maksimum).