Raport laboratorium 1

Ćwiczenie nr 1

Celem ćwiczenia jest minimalizacja funkcji celu i porównanie wyników dla metody najszybszego spadku gradientu i metody Newtona.

Funkcją celu jest funkcja Himmelblau postaci:

f(x,y) = (x*x+y-11)^2 + (x+y*y-7)^2 , -5 <= x <= 5, -5 <= y <= 5

Biblioteki użyte podczas implementacji rozwiązania:

- autograd
- numpy
- matplotlib
- time
- random
- pandas

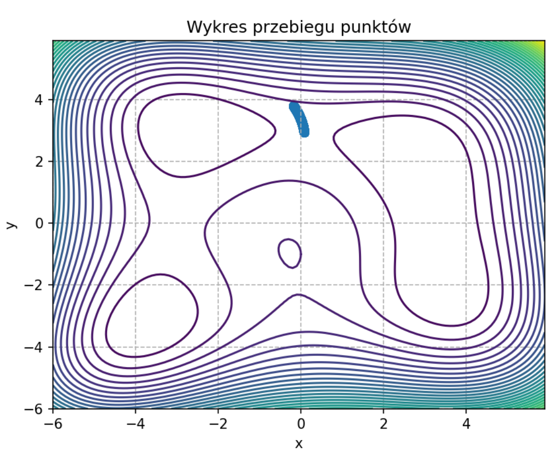
W podanym przeze mnie rozwiązaniu do obliczenia gradientu i hessianu użyłem biblioteki autograd.

W kolejnych iteracjach został losowany punkt x z zakresu -5 <= x <= 5, -5 <= y <= 5 .

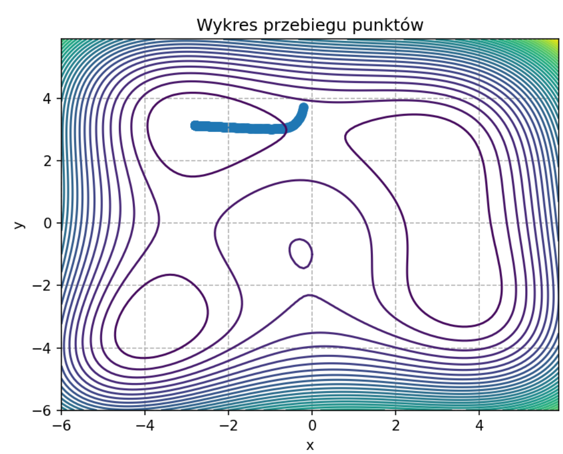
1.Ustawienia pozwalające na skuteczną prace obu metod

Punkt siodłowy

Wykres dla Metody Newtona



Wykres dla Metody najszybszego spadku gradientu

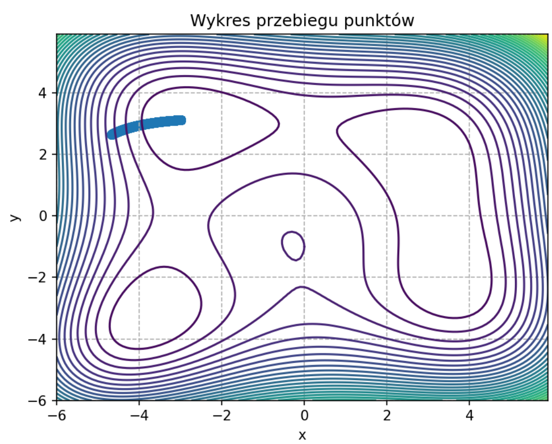


Metoda	minimum	współrzędne	wylosowany punkt
NewtonMethod	67.87	[0.08, 2.95]	[-0.19, 3.81]
GradientMethod	0.0	[-2.81, 3.13]	[-0.19, 3.81]

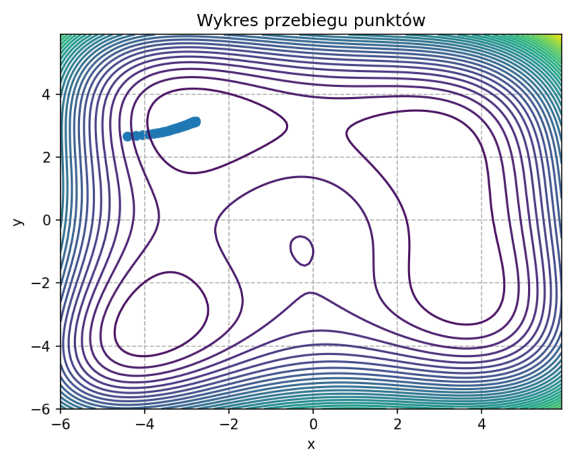
Powyżej znajdują się dwa wykresy i dane z pomiarów przez dwie metody dla jednakowego punktu [-0.19, 3.81]. Metoda Newtona nie pokazała minimum, wynik dość mocno odbiega od minimum globalnego, które dla funkcji Himmelblau wynosi 0. Wykres jednak pokazuje że punkt w którym zatrzymała się funkcja szukając minimum jest punktem siodłowym, znajdującym się pomiędzy dwoma minimami lokalnymi. Na wykresie zostały przedstawione poziomice pokazujące jak zmieniają się wartości funkcji dla danych x i y.

Poprawny wynik działania obu metod

Wykres dla Metody Newtona



Wykres dla Metody najszybszego spadku gradientu

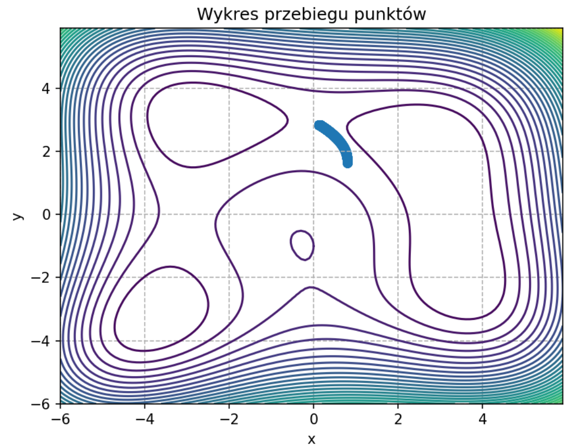


Metoda	minimum	współrzędne	wylosowany punkt
NewtonMethod	1.17	[-2.99, 3.12]	[-4.68, 2.64]
GradientMethod	0.0	[-2.81, 3.13]	[-4.68, 2.64]

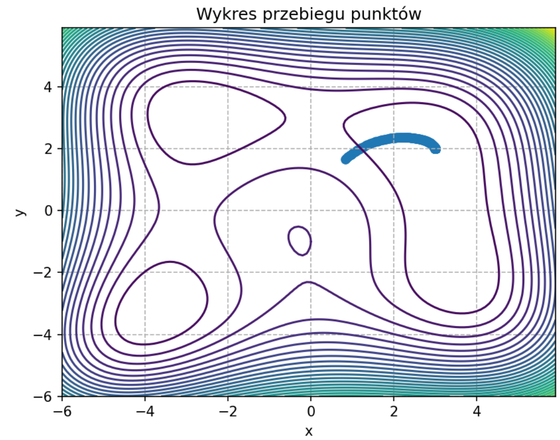
Powyżej znajdują się dane i wykresy z pomiarów przez dwie metody dla jednakowego punktu [-4.68, 2.64]. Metoda Newtona pokazała tym razem wynik zbliżony do minimum globalnego, które wynosi 0, jednak nie doszła do niego prawdopodobnie ze względu na ilość iteracji. Dla pomiarów z tego zestawienia liczba iteracji wynosiła 3000 a współczynnik Beta wynosił 0.001. Metoda Najszybszego spadku gradientu odnalazła minimum globalne równe 0.

Punkt siodłowy

Wykres dla Metody Newtona



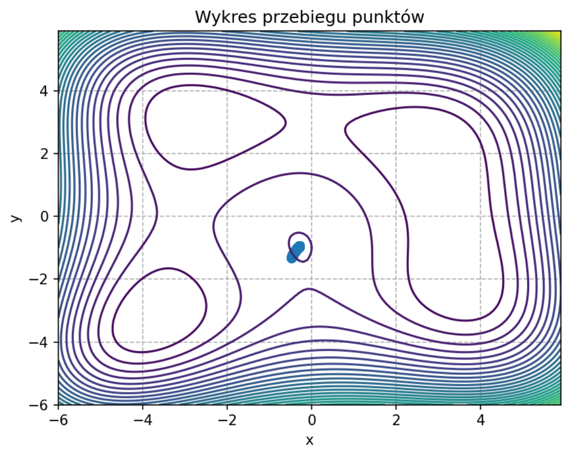
Wykres dla Metody najszybszego spadku gradientu



Metoda	minimum	współrzędne	wylosowany punkt
NewtonMethod	67.72	[0.13, 2.85]	[0.79, 1.62]
GradientMethod	0.0	[3.0, 2.0]	[0.79, 1.62]

Powyżej znajdują się dane i wykresy z pomiarów przez dwie metody dla jednakowego punktu [0.79, 1.62]. Po raz kolejny podczas pomiarów trafiliśmy na punkt siodłowy, funkcja realizująca Metodę Newtona zatrzymała się na wartości 67.72, która jak widzimy na wykresie znajduję się pomiędzy dwoma minimami lokalnymi. Metoda najszybszego spadku gradientu poprawnie znalazła minimum.

Znalezienie maksimum zamiast minimum



minimum	współrzędne	wylosowany punkt
181.61	[-0.28, -0.94]	[-0.48, -1.32]

Metoda Newtona znalazła zamiast minimum maksimum globalne, zadziało się to podczas paru pomiarów, ponieważ w tych pomiarach punkt startowy został wylosowany bardzo blisko maksimum funkcji. Pochodna w punkcie, w którym jest maksimum podobnie jak w przypadku minimum wynosi zero, dlatego metoda Newtona uznała punkt maksimum jako iteracyjnie znalezione minimum.

Poniżej znajdują się wyniki z 25 pomiarów dla ustawienia:

liczba iteracji: 3000

współczynnik Beta: 0.001

Metoda Newtona

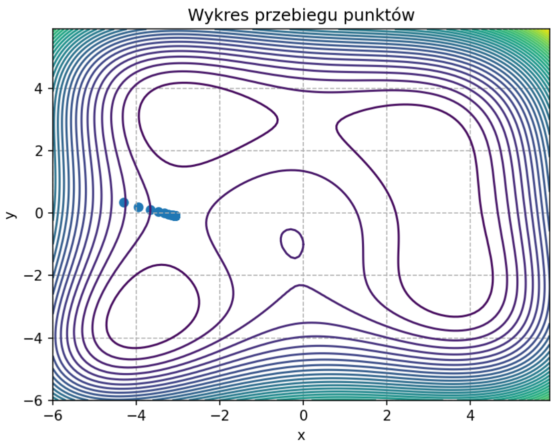
Minimum	Czas - Newton	iteracje	beta
0.64	8.42	3000	0.001
0.01	10.33	3000	0.001
13.32	10.41	3000	0.001
67.87	10.88	3000	0.001
0.0	10.04	3000	0.001
67.81	9.6	3000	0.001
1.17	9.25	3000	0.001
0.43	9.55	3000	0.001
178.33	8.93	3000	0.001
0.0	10.24	3000	0.001
1.82	9.84	3000	0.001
5.4	9.61	3000	0.001
13.34	9.85	3000	0.001
0.4	9.78	3000	0.001
67.72	9.06	3000	0.001
3.65	10.62	3000	0.001
181.62	9.03	3000	0.001
0.03	10.61	3000	0.001
181.49	9.32	3000	0.001
0.41	9.86	3000	0.001
69.05	9.92	3000	0.001
0.01	9.9	3000	0.001
104.9	9.26	3000	0.001
0.28	9.03	3000	0.001
0.0	8.88	3000	0.001

Metoda najszybszego spadku gradientu

Minimum	Czas - gradient	iteracje	bet
0.0	1.33	3000	0.0
0.0	1.57	3000	0.0
0.0	1.27	3000	0.0
0.0	1.43	3000	0.0
0.0	1.27	3000	0.0
0.0	1.18	3000	0.0
0.0	1.18	3000	0.0
0.0	1.24	3000	0.0
0.0	1.21	3000	0.0
0.0	1.39	3000	0.0
0.0	1.22	3000	0.0
0.0	1.21	3000	0.0
0.0	1.27	3000	0.0
0.0	1.22	3000	0.0
0.0	1.17	3000	0.0
0.0	1.38	3000	0.0
0.0	1.2	3000	0.0
0.0	1.35	3000	0.0
0.0	1.21	3000	0.0
0.0	1.25	3000	0.0
0.0	1.18	3000	0.0
0.0	1.2	3000	0.0
0.0	1.23	3000	0.0
0.0	1.19	3000	0.0
0.0	1.21	3000	0.0

Wykres dla Metody Newtona

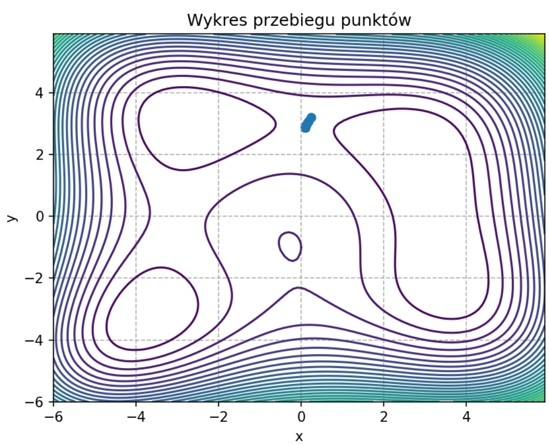
Wykres dla Metody Newtona



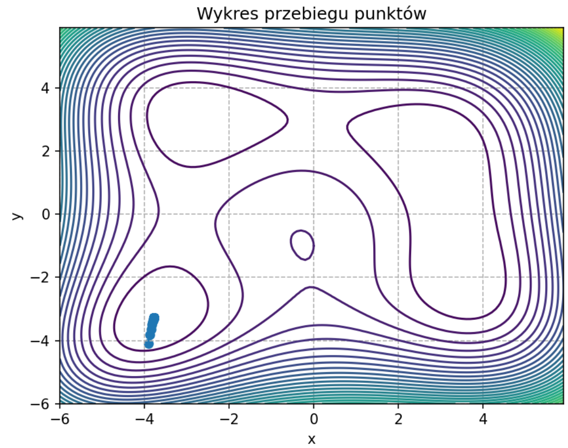
Powyżej i poniżej znajduję się wykres z pomiaru, w którym został znaleziony punkt siodłowy. Pokazują to wykresy, dzięki którym widzimy, że metoda zatrzymała się w punkcie pomiędzy dwoma minimalmi lokalnymi.

minimum	współrzędne	wylosowany punkt
104.02	[-3.07, 0.08]	[-4.33, -0.72]

Wykres dla Metody Newtona



minimum	współrzędne	wylosowany punkt
67.72	[0.09, 2.88]	[1.28, 3.22]



Powyżej znajduję się wykres z pomiaru, w którym metoda Newtona poprawnie odnalazła minimum.

minimum	współrzędne	wylosowany punkt
0.0	[-3.78, 3.28]	[-3.96, -4.44]

Poniżej wyniki z pomiarów dla 2000 iteracji i współczynnika Beta 0.43:

Metoda Newtona

Minimum	Czas - Newton	iteracje	beta

Metoda najszybszego spadku gradientu

Minimum	Czas - gradient	iteracje	bet

104.02	4.86	2000	0.43			0.88	2000	0.4
104.02	7.0	2000	0.43			1.39	2000	0.4
0.0	6.64	2000	0.43			1.29	2000	0.4
0.0	7.36	2000	0.43			1.04	2000	0.4
0.0	6.94	2000	0.43			1.31	2000	0.4
0.0	6.81	2000	0.43			1.03	2000	0.4
181.62	6.83	2000	0.43			1.11	2000	0.4
0.0	7.3	2000	0.43			1.04	2000	0.4
0.0	7.24	2000	0.43			1.0	2000	0.4
0.0	6.93	2000	0.43			1.07	2000	0.4
13.31	6.85	2000	0.43			1.08	2000	0.4
0.0	6.77	2000	0.43			1.03	2000	0.4
178.34	7.12	2000	0.43			1.17	2000	0.4
0.0	6.81	2000	0.43			1.3	2000	0.4
0.0	7.18	2000	0.43			1.4	2000	0.4
0.0	6.88	2000	0.43			1.18	2000	0.4
67.72	6.94	2000	0.43			1.1	2000	0.4
0.0	6.84	2000	0.43			1.31	2000	0.4
104.02	7.19	2000	0.43			1.2	2000	0.4
0.0	6.81	2000	0.43			1.27	2000	0.4
0.0	6.79	2000	0.43			1.17	2000	0.4
0.0	6.78	2000	0.43			1.26	2000	0.4
13.31	6.91	2000	0.43			1.28	2000	0.4

Dla dużych wartości

Poniżej wyniki z pomiarów dla 6300 iteracji i współczynnika Beta 0.01:

Metoda Newtona

Metoda najszybszego spadku gradientu

Minimum	Czas - Newton	iteracje	beta	Minimum		Czas - gradient	iteracje	bet
13.31	16.41	6300	0.01	0.0		2.54	6300	0.0
0.0	16.18	6300	0.01	0.0		2.59	6300	0.0
104.02	16.41	6300	0.01	0.0		2.58	6300	0.0
0.0	16.73	6300	0.01	0.0		2.57	6300	0.0
13.31	17.15	6300	0.01	0.0		2.72	6300	0.0
104.02	16.44	6300	0.01	0.0		2.53	6300	0.0
0.0	18.24	6300	0.01	0.0		2.64	6300	0.0
0.0	17.38	6300	0.01	0.0		2.85	6300	0.0
0.0	17.12	6300	0.01	0.0		2.63	6300	0.0

0.0	16.4	6300	0.01	0.0		2.74	6300	0.0

Poniżej wyniki z pomiarów dla 8200 iteracji i współczynnika Beta 0.0001:

Metoda Newtona

Metoda najszybszego spadku gradientu

Minimum	Czas - Newton	iteracje	beta	Minimum		Czas - gradient	iteracje	bet
120.67	18.41	8200	0.0001	0.0		3.28	8200	0.0
105.67	20.01	8200	0.0001	0.0		3.22	8200	0.0
19.51	19.86	8200	0.0001	0.0		3.39	8200	0.0
121.53	19.16	8200	0.0001	0.0		3.29	8200	0.0
56.4	20.0	8200	0.0001	0.0		3.25	8200	0.0
68.02	20.37	8200	0.0001	0.0		3.32	8200	0.0
175.08	19.66	8200	0.0001	0.0		3.34	8200	0.0
172.02	20.2	8200	0.0001	0.0		3.31	8200	0.0
7.1	20.7	8200	0.0001	0.0		3.29	8200	0.0
29.2	21.28	8200	0.0001	0.0		3.4	8200	0.0

Poniżej wyniki z pomiarów dla 20200 iteracji przy współczynniku Beta 0.00001:

Minimum	Czas - Newton	iteracje	beta	Minimum		Czas - gradient	iteracje	bet
116.27	43.8	20200	0,00001	0.0		8.09	20200	0,0
15.29	39.21	20200	0,00001	0.0		8.01	20200	0,0
33.99	39.15	20200	0,00001	10.65		8.02	20200	0,0
42.9	39.34	20200	0,00001	0.0		8.12	20200	0,0
81.92	39.82	20200	0,00001	0.0		7.93	20200	0,0
38.6	38.93	20200	0,00001	1.46		7.84	20200	0,0
22.71	38.95	20200	0,00001	0.0		7.95	20200	0,0
134.82	38.95	20200	0,00001	0.0		7.82	20200	0,0
5.84	38.81	20200	0,00001	0.0		7.85	20200	0,0
7.27	38.9	20200	0,00001	0.0		7.71	20200	0,0

Działanie obu metod

- Wyniki kolejnych pomiarów dla różnych ustawień liczby iteracji i współczynnika Beta pokazują, że metoda najszybszego wzrostu gradientu działa dobrze im mniejszy jest współczynnik beta, natomiast odwrotnie jest w przypadku metody Newtona, która działa poprawnie dla większych wartości współczynnika beta.

Występowanie punktów siodłowych

- Przypisana funkcja celu w swojej dziedzinie posiada cztery minima globalne, które wynoszą wszystkie 0. Punkt siodłowy to punkt, w którym funkcja nie posiada ekstremum(ani minimum ani maksimum) jednak pochodne cząstkowe w tym punkcie są równe 0, przy czym punkt siodłowy nie jest minimum ani maksimum funkcji.

- W metodzie Newtona korzystamy z hessianu, który jest drugą pochodną funkcji, czyli określa jak zmienia się gradient (nachylenie/wzrost/spadek funkcji).
- W metodzie najszybszego spadku gradientu korzystamy z gradientu, który jest pierwszą pochodną funkcji, czyli określa jak zmieniają się wartości funkcji.
- Z tego powodu metoda Newtona jest narażona na wystąpienie punktów siodłowych. Funkcja, która ma więcej niż jedno ekstremum(w tym przypadku minimum) posiada obszar, na którym z obu stron występuje nachylenie(w tym przypadku spadek), przez co funkcja wykorzystująca tą metodę przechodzi z kolejnymi iteracjami w środek krzywizny pomiędzy dwoma ekstremami. Druga pochodna w tym punkcie jest różna od 0, natomiast pierwsza pochodna równa 0 , dlatego punkt ten jest traktowany jako potencjalne ekstremum.

Porównanie szybkości metod

- Metoda Newtona znajduję minimum szybciej niż metoda najszybszego spadku gradientu, ponieważ krok każdej iteracji jest skierowany w stronę minimum, a nie jak w przypadku drugiej metody w stronę największego spadku wartości funkcji. Dzięki temu metoda Newtona porusza się bardziej po linii prostej a nie zakręca jak w przypadku metody spadku gradientu. Z tego powodu metoda Newtona potrzebuje mniej iteracji na znalezienie minimum.
- Czasy podane w tabelach są większe dla metody Newtona, ponieważ jest to wynik obliczeń. Metoda ta wymaga obliczenia odwrotności hessianu co zajmuję więcej czasu.