

Metody gradientowe			
Adrian Jałoszewski	19 IV 2017	Środa 14:00	

1 Cel ćwiczenia

Celem ćwiczenia jest zapoznanie się z algorytmami gradientowymi oraz algorytmami zmiennej metryki w rozwiązywaniu problemów minimalizacji.

2 Przebieg ćwiczenia

Ćwiczenie polegało na minimalizacji dwóch rodzajów funkcji celu:

$$Q(x, y) = x_1^2 + ax_2^2$$

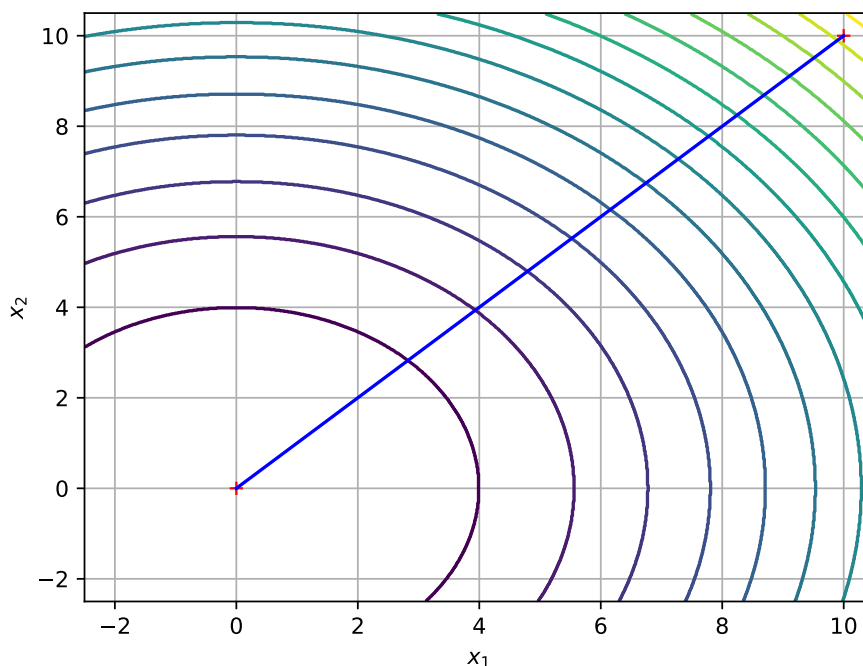
gdzie $a \in \{1, 0.5, 0.3\}$ oraz:

$$Q(x, y) = 6x_1^2 + 6x_1x_2 + x_2^2 + 4,5(e^{x_1} - x_1 - 1) + 1,5(e^{x_2} - x_2 - 1)$$

Zadania różniły się między sobą używanymi do tego metodami oraz punktami startowymi i warunkami stopu

2.1 Ćwiczenie pierwsze

Ćwiczenie pierwsze polegało na obserwacji efektu zygzakowania w przypadku gdy kierunek najszybszego spadku nie pokrywał się z kierunkiem w którym znajduje się minimum. W tym przypadku była minimalizowana forma kwadratowa. Punktem startowym dla wszystkich przypadków był punkt $(10, 10)$.



Rysunek 1: Współczynnik $a = 1$

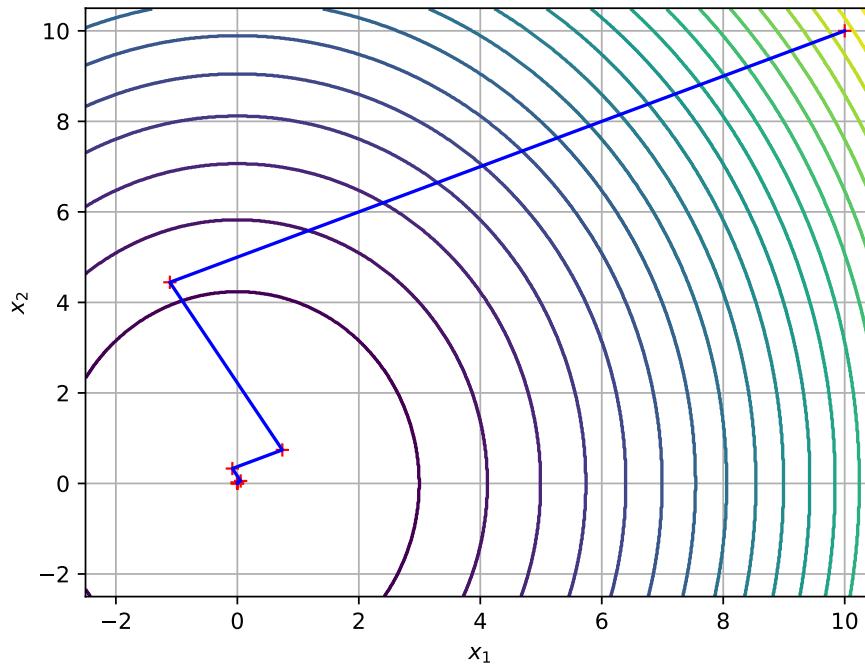
Ponieważ dla współczynnika $a = 1$ macierz minimalizowanej formy kwadratowej jest macierzą jednostkową, kierunek najszybszego spadku pokrywa się z kierunkiem minimum – minimum jest osiągnięte w jednym kroku.

$$\nabla x^T A x = \nabla x^T I x = \nabla x^T x = 2x$$

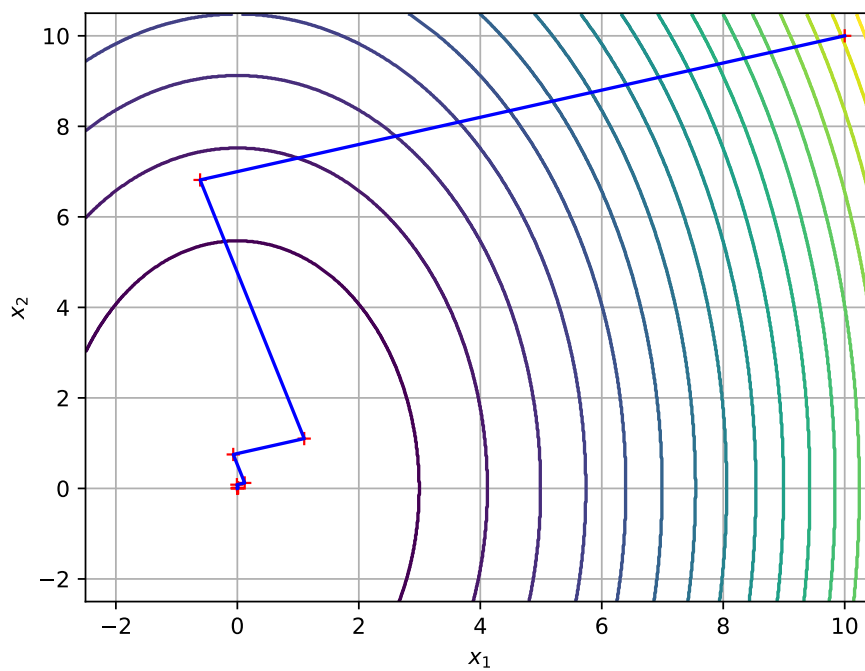
Gdzie wektor W ogólnym przypadku gradient ma następującą postać:

$$\nabla(x_1^2 + ax_2^2) = 2 \cdot [x_1 \ ax_2]^T$$

Wektor ten nie jest współliniowy z wektorem $[x_1 \ x_2]^T$, a przez to nie wskazuje w kierunku punktu (0, 0) gdzie się znajduje minimum funkcji. Im większa różnica w kierunku tych wektorów tym większy efekt zygzakowania.



Rysunek 2: Współczynnik $a = 0.5$

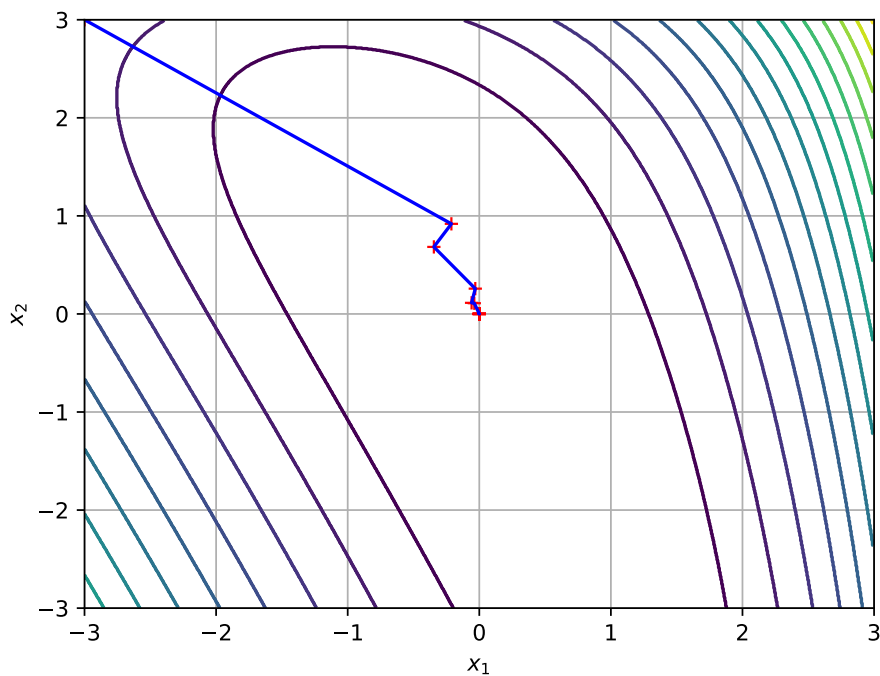


Rysunek 3: Współczynnik $a = 0.3$

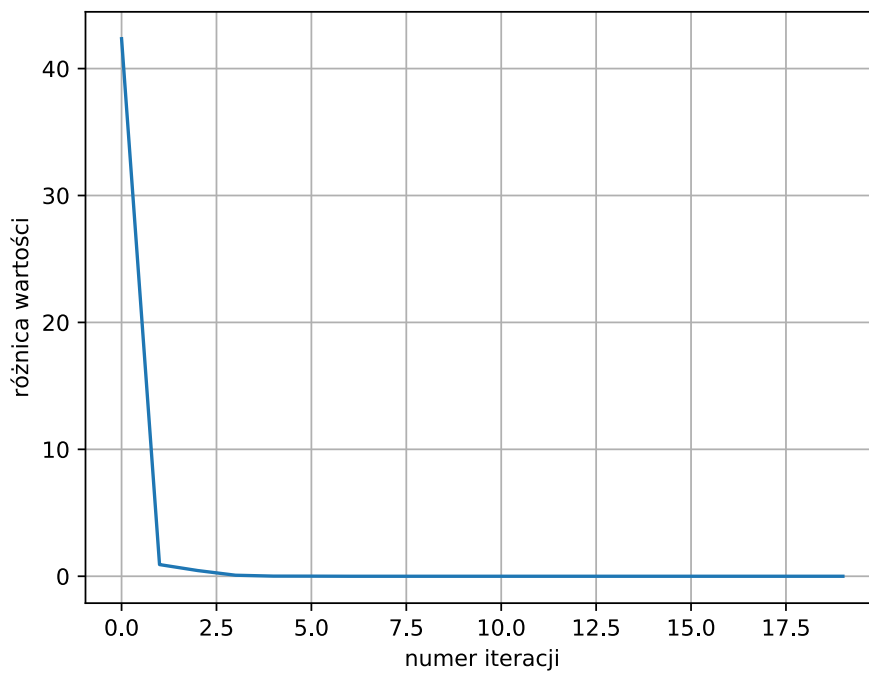
2.2 Ćwiczenie drugie

Ćwiczenie drugie polegało na zbadaniu zachowania poszczególnych metod gradientowych dla drugiej funkcji. Osiąga ona swoje minimum w punkcie $(0, 0)$.

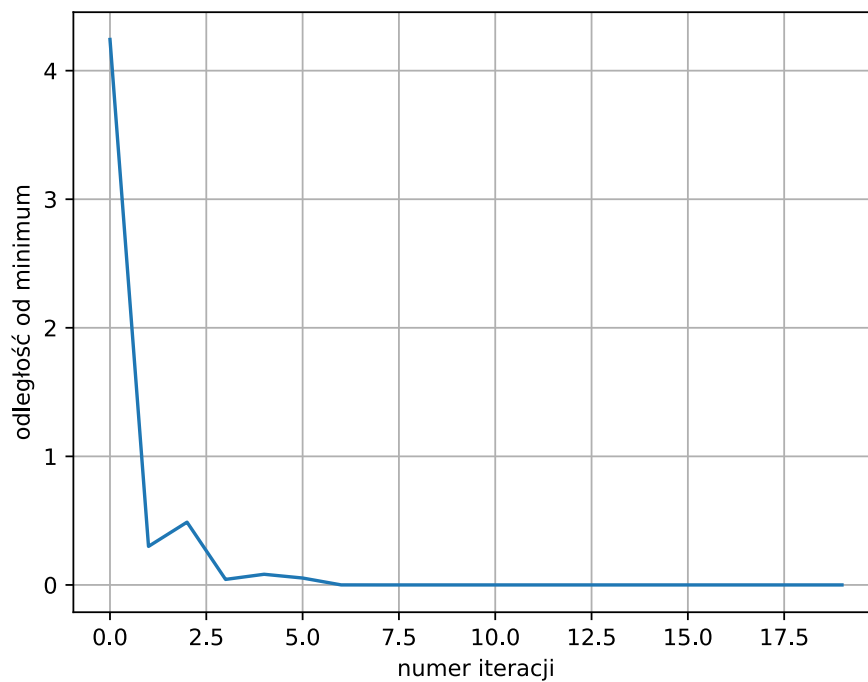
2.2.1 Metoda Fletchera-Reevesa – startująca z punktu $(-3, 3)$



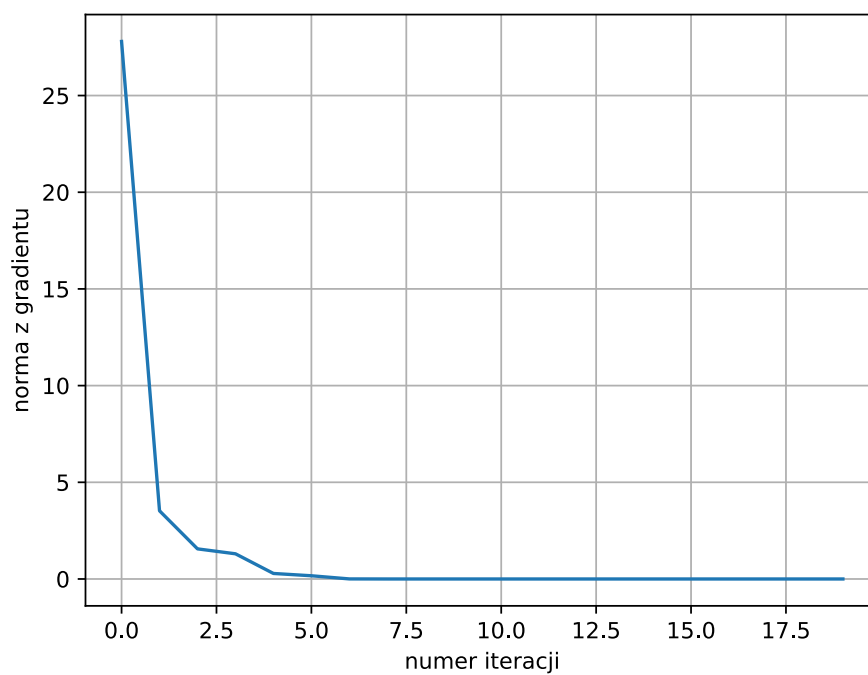
Rysunek 4: wyznaczone punkty



Rysunek 5: Różnica między minimum, a aktualną wartością

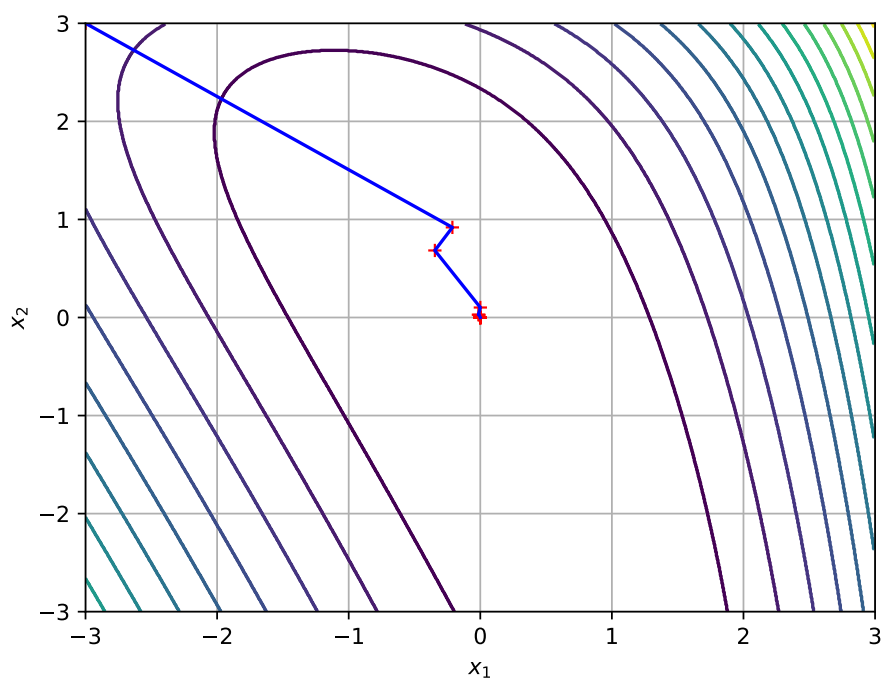


Rysunek 6: Odległość od minimum

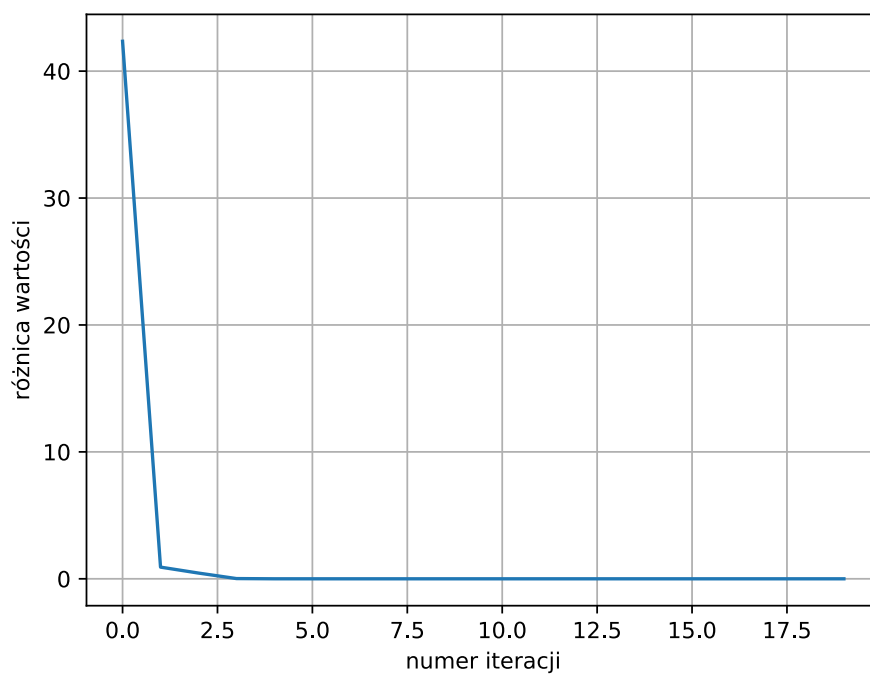


Rysunek 7: norma z gradientu

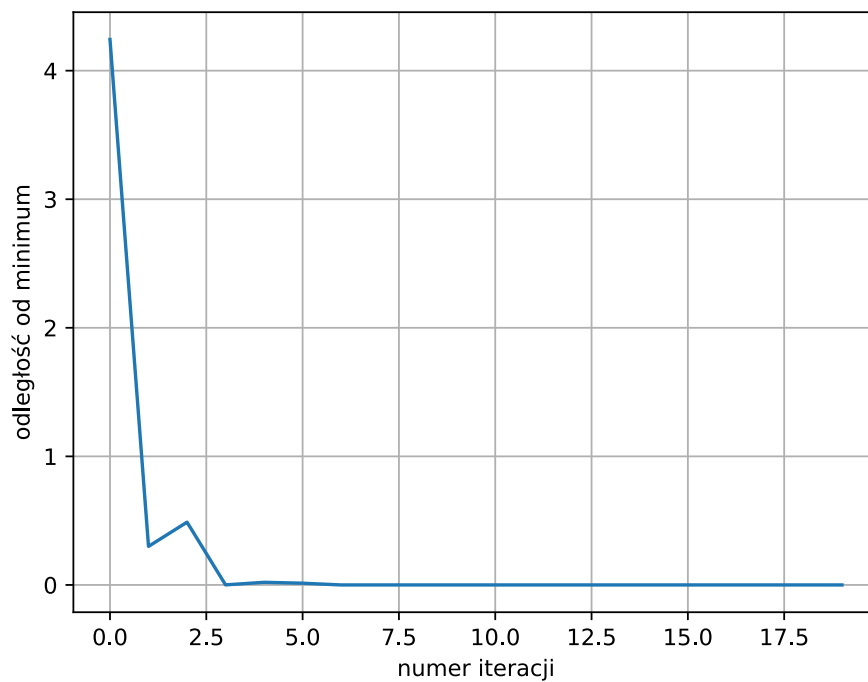
2.2.2 Metoda Polaka-Ribiera – startująca z punktu $(-3, 3)$



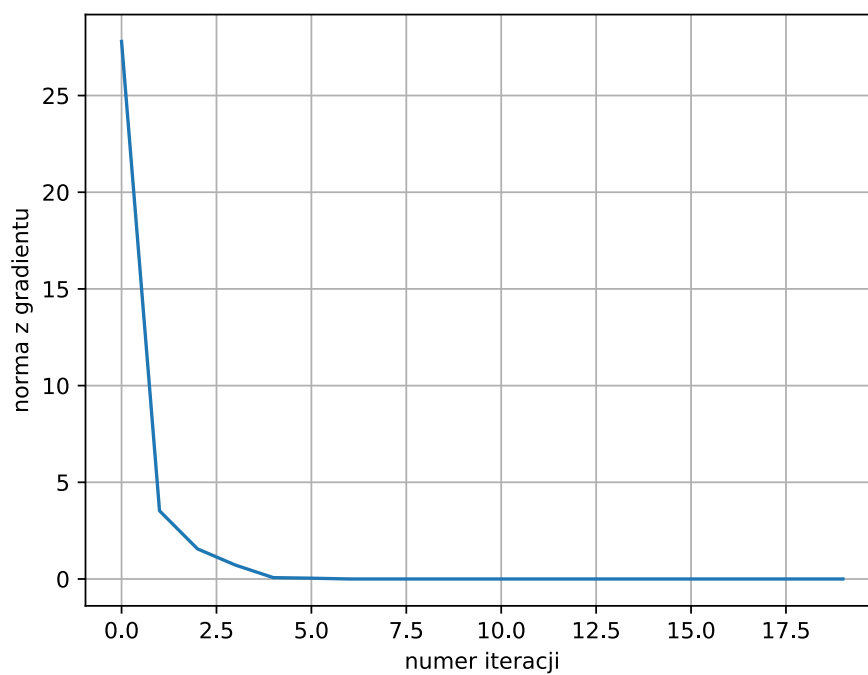
Rysunek 8: wyznaczone punkty



Rysunek 9: różnica między minimum, a aktualną wartością

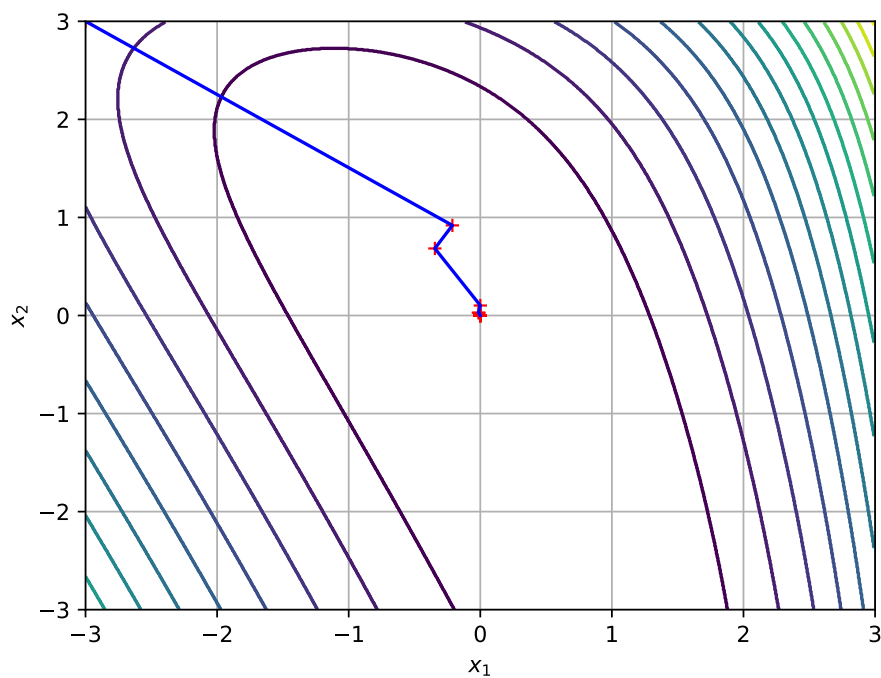


Rysunek 10: odległość od minimum

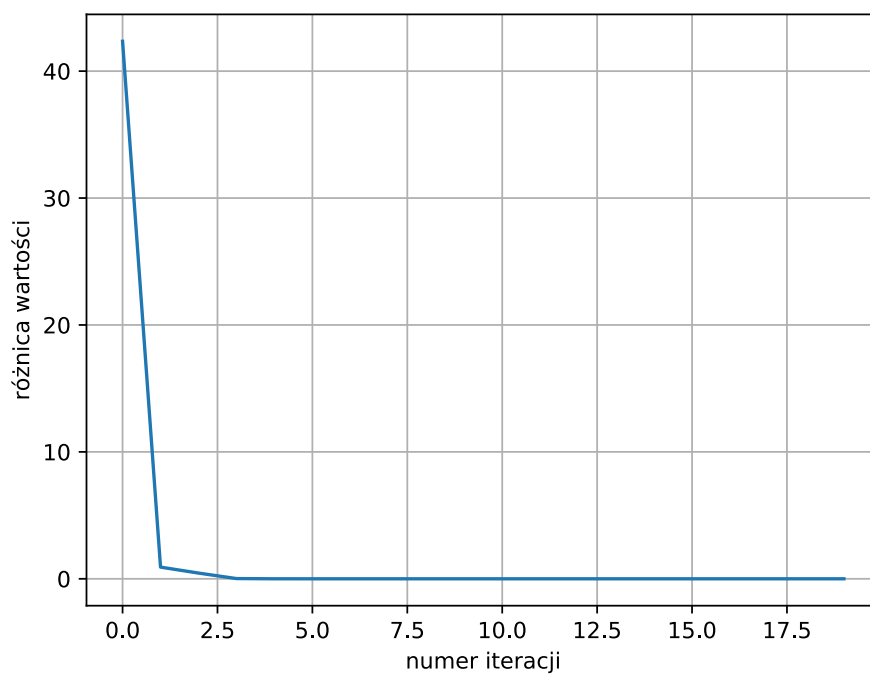


Rysunek 11: norma z gradientu

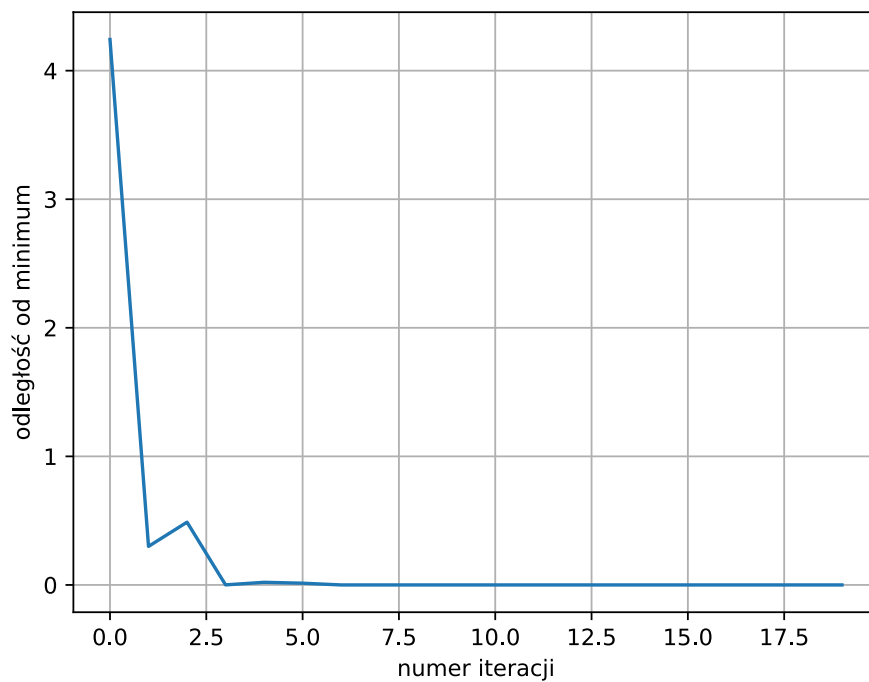
2.2.3 Pełen wzór na współczynnik β – startujący z punktu $(-3, 3)$



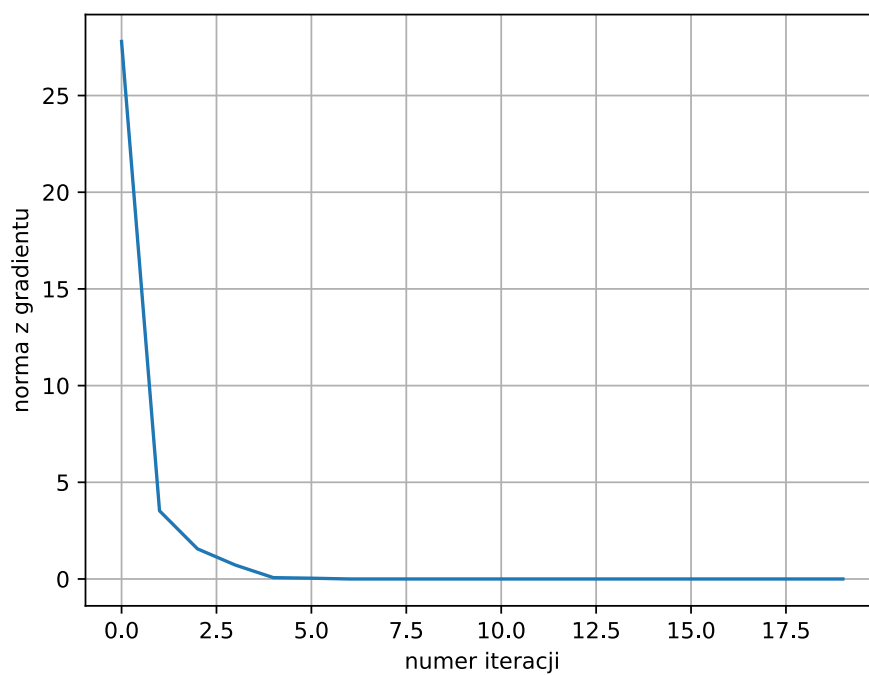
Rysunek 12: wyznaczone punkty



Rysunek 13: różnica między minimum, a aktualną wartością

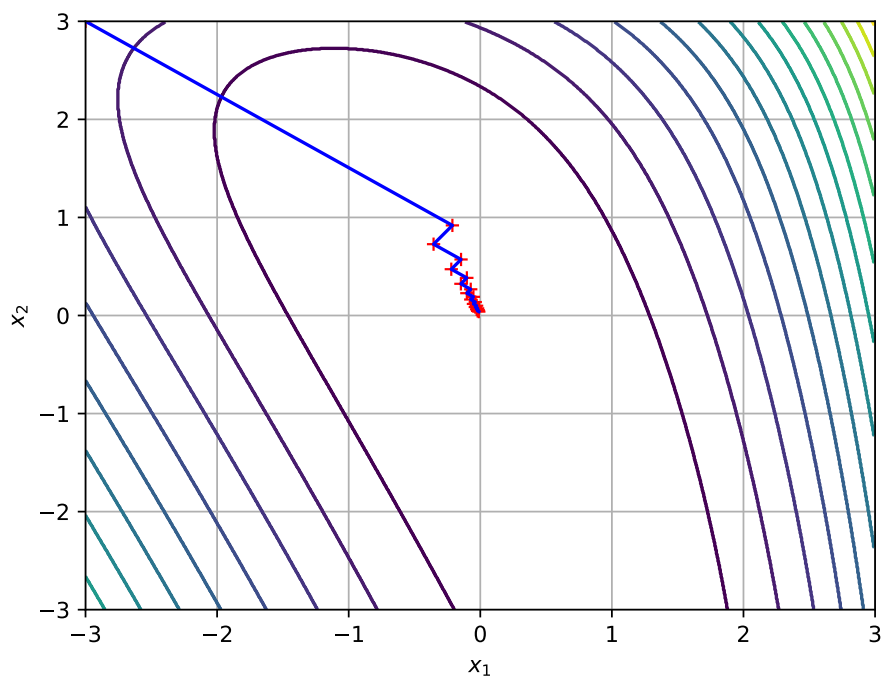


Rysunek 14: odległość od minimum

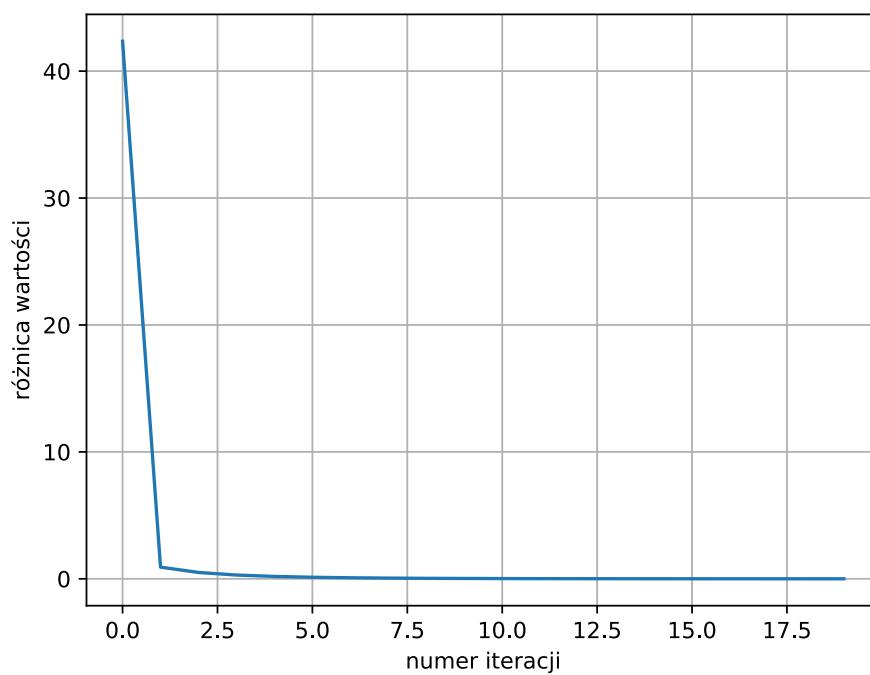


Rysunek 15: norma z gradientu

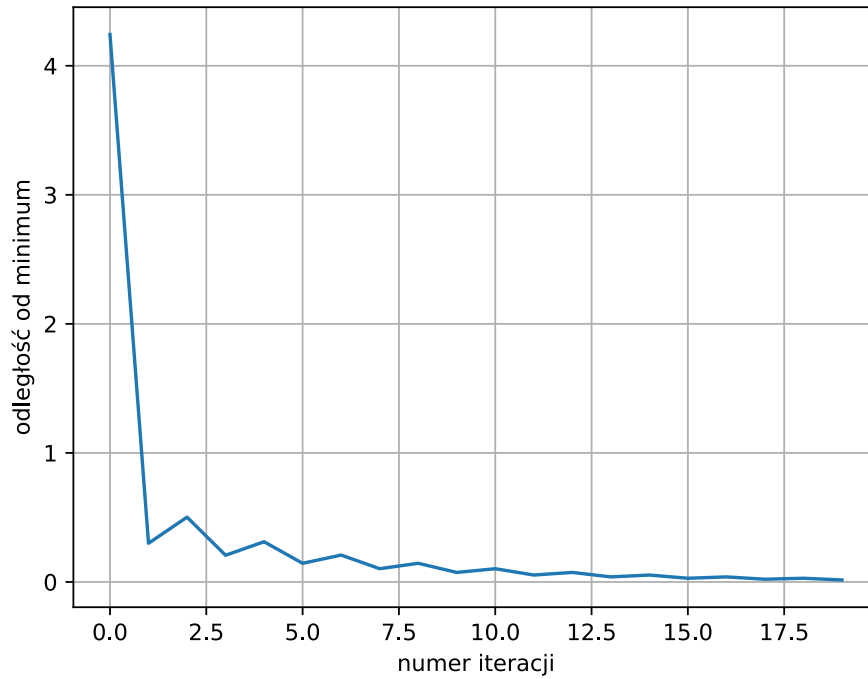
2.2.4 Metoda najszybszego spadku zaczynająca w punkcie $(-3, 3)$



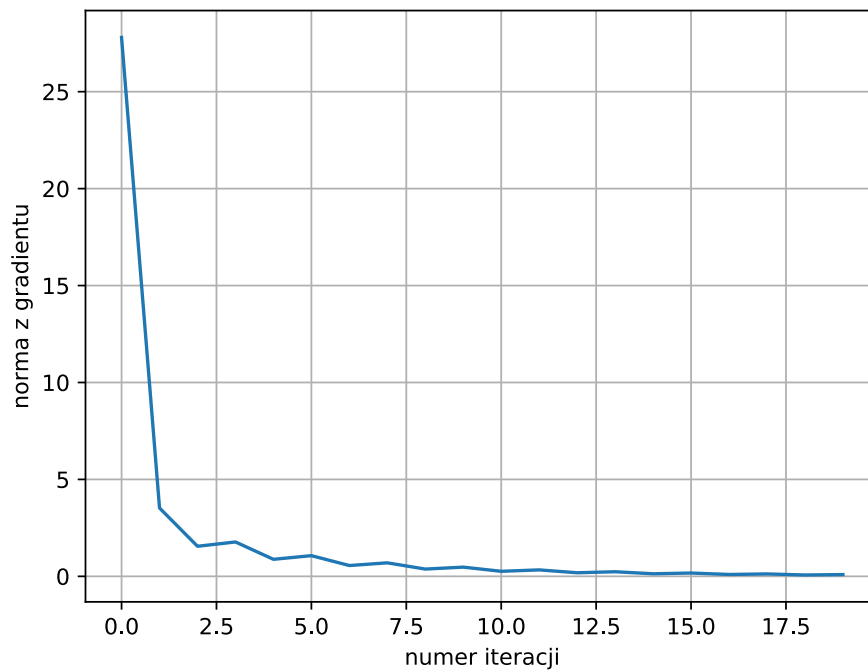
Rysunek 16: wyznaczone punkty



Rysunek 17: różnica między minimum, a aktualną wartością



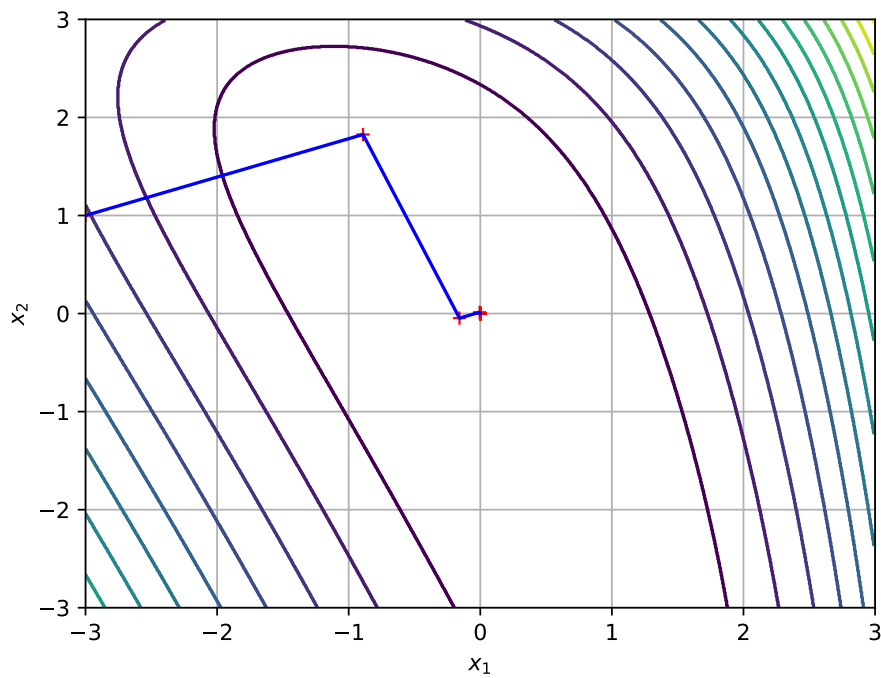
Rysunek 18: odległość od minimum



Rysunek 19: norma z gradientu

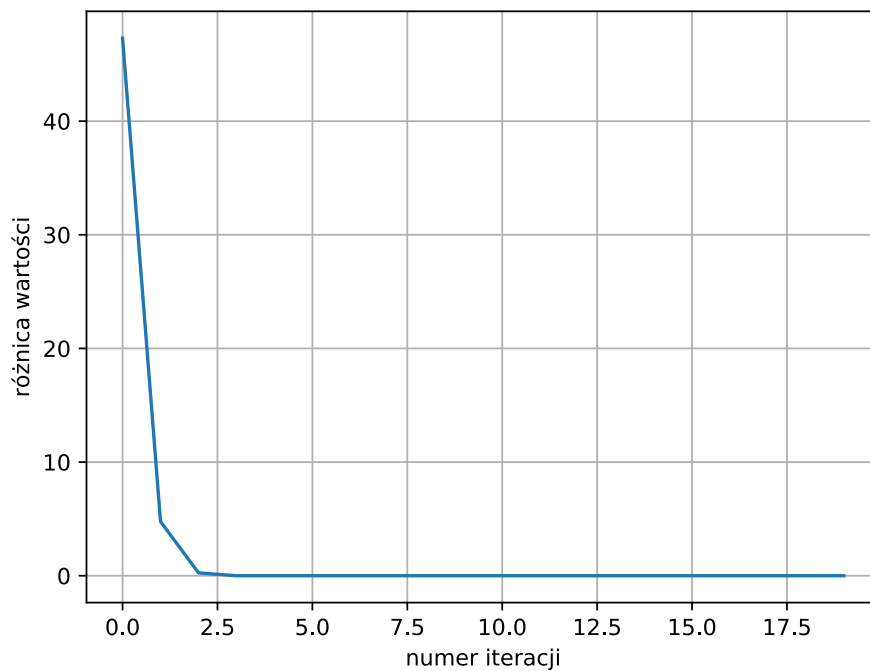
2.2.5 Metoda najszybszego spadku – startująca z punktu $(-3, 1)$

Jest to dobry przykład tego jak dobranie punktu początkowego może mieć wpływ na szybkość zbieżności metody. Poziomicie są tu przecinane przez proste na których było poszukiwane minimum.

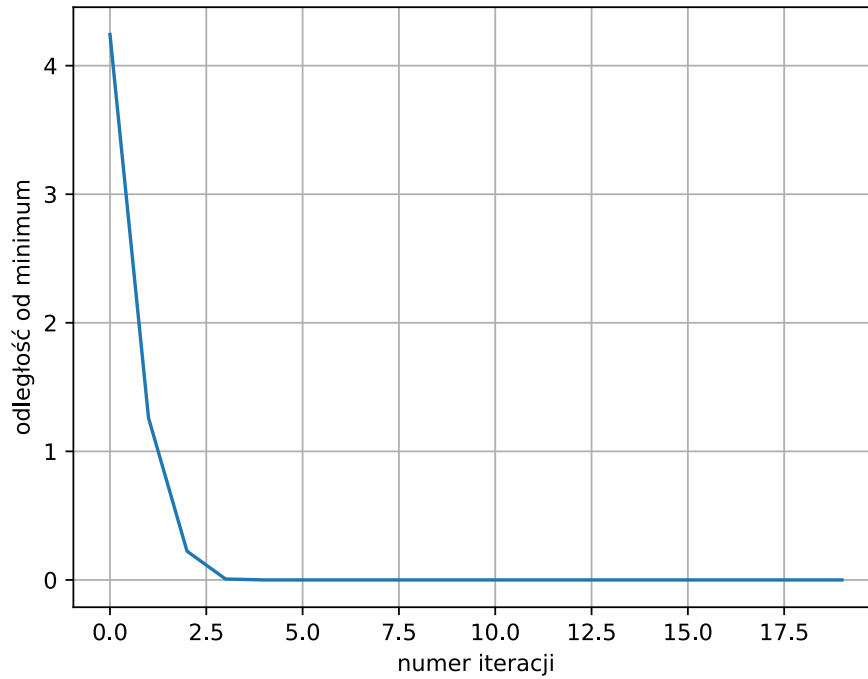


Rysunek 20: wyznaczone punkty

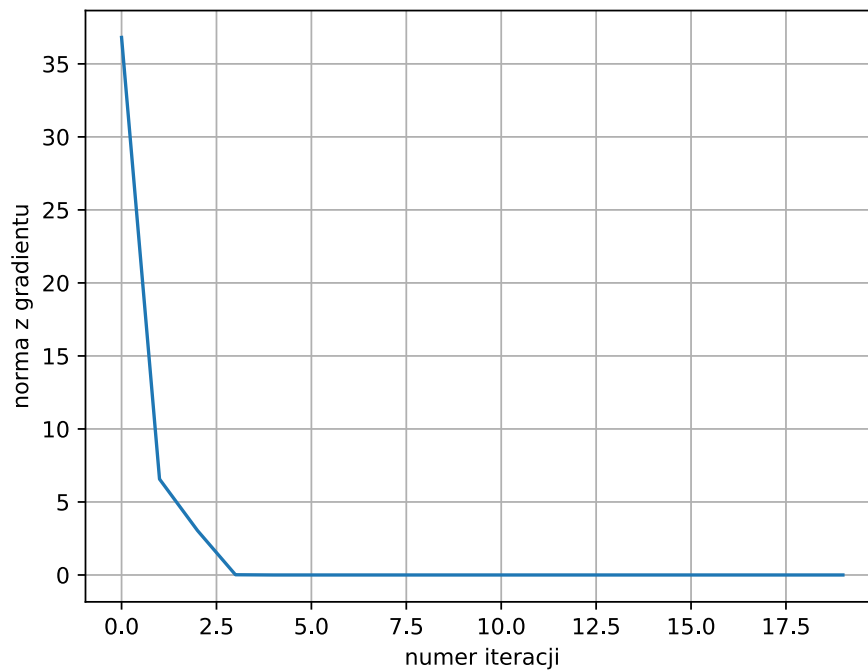
Odpowiednie dobranie punktów startowych sprawiło, że metoda ta znalazła minimum znacznie szybciej niż w przypadku gdy startowała z punktu $(-3, 3)$. Punkt startowy jest dobrany tak aby wyeliminować zyg-zakowanie w poszukiwaniu minimum.



Rysunek 21: różnica między minimum, a aktualną wartością



Rysunek 22: odległość od minimum

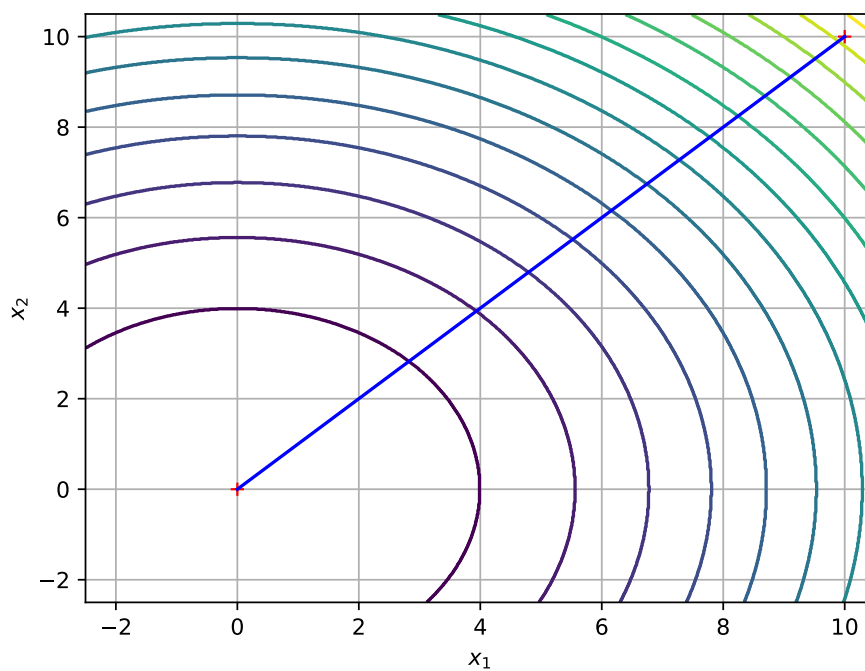


Rysunek 23: norma z gradientu

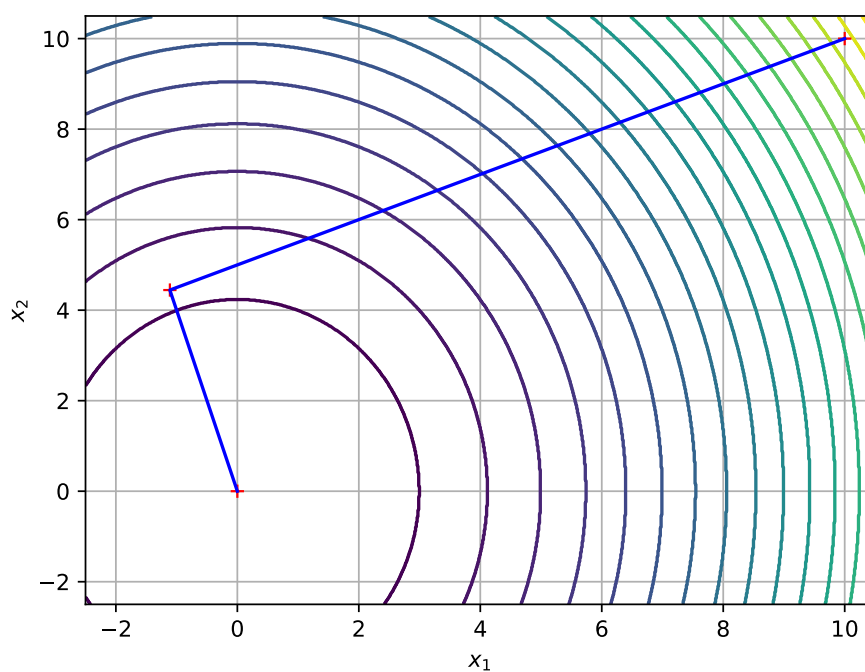
2.3 Ćwiczenie trzecie

Celem zadania trzeciego jest zbadania metod zmiennej metryki w kontekście optymalizacji funkcji celu w postaci formy kwadratowej dla różnych metod dla różnych współczynników a . Jest to ta sama funkcja, co była rozważana w ćwiczeniu pierwszym z tym samym punktem startowym $(10, 10)$.

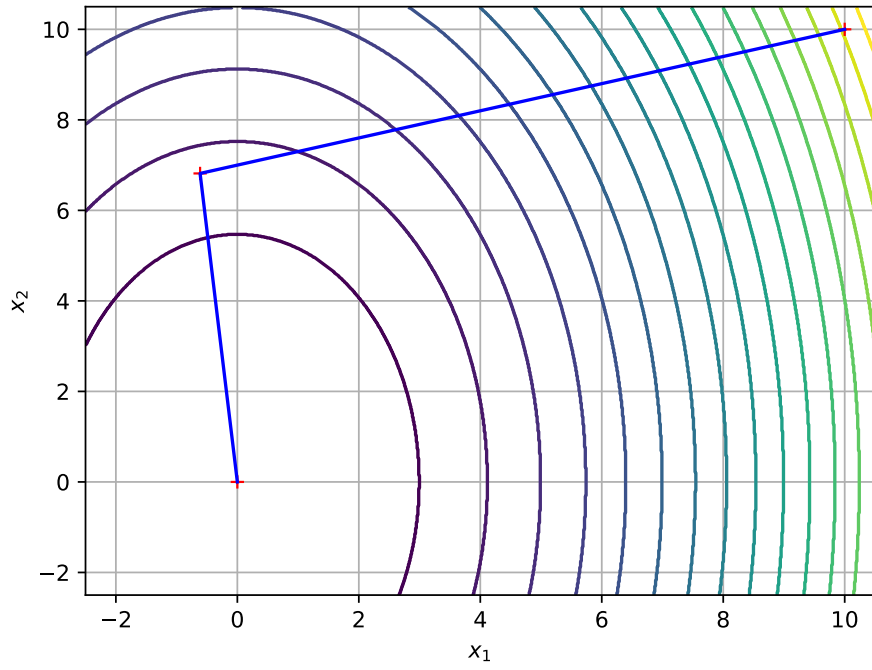
2.3.1 Metoda Davidona-Fletcher-Powella



Rysunek 24: Współczynnik $a = 1$

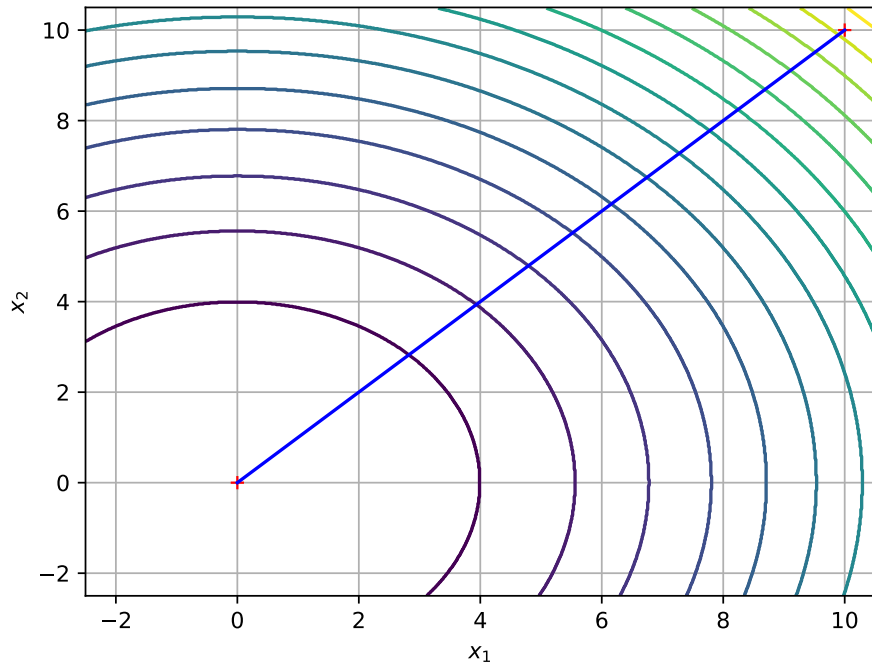


Rysunek 25: Współczynnik $a = 0.5$

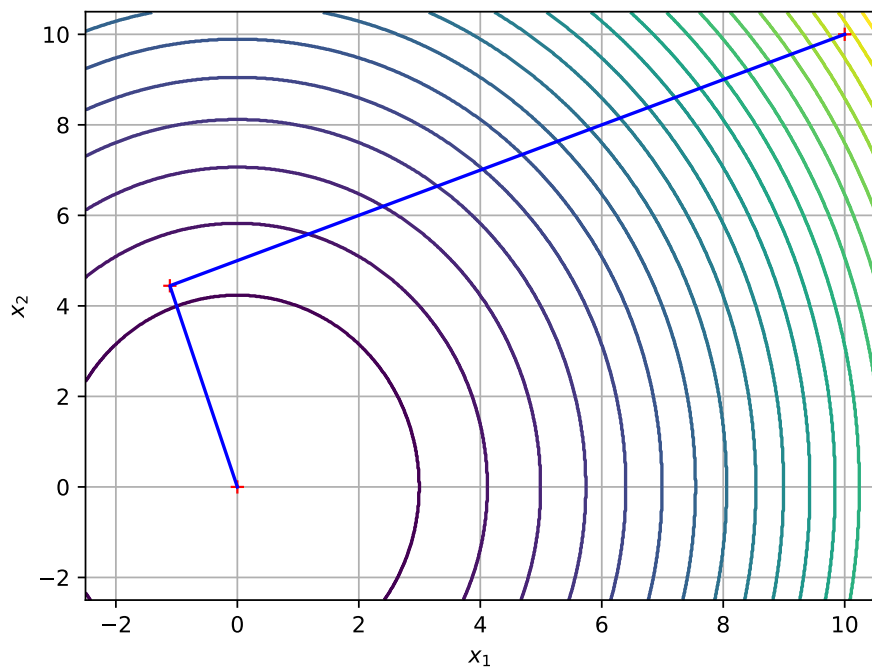


Rysunek 26: Współczynnik $a = 0.3$

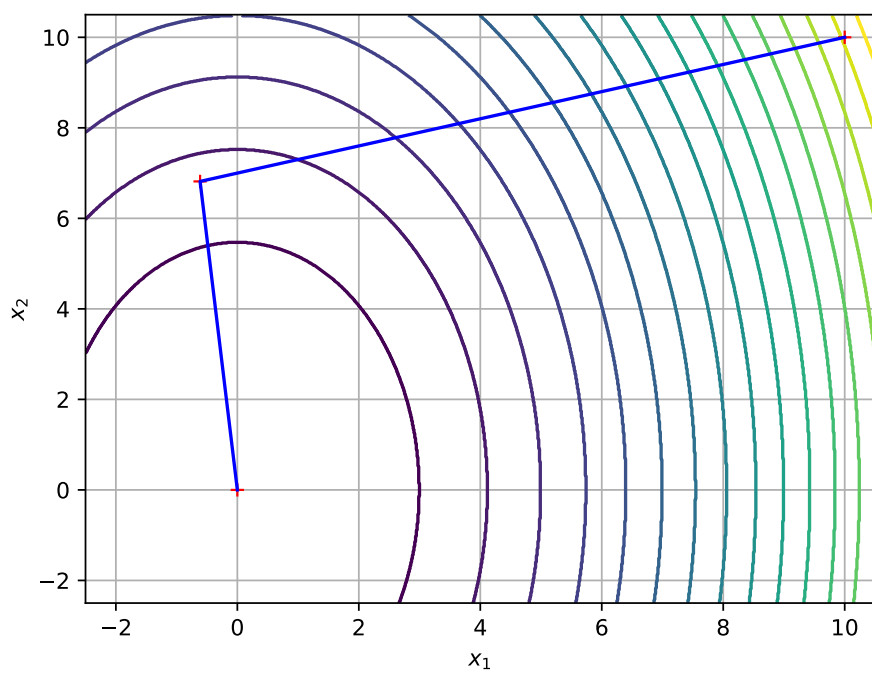
2.4 Metoda Wolfe'a-Broydena-Davidona



Rysunek 27: Współczynnik $a = 1$

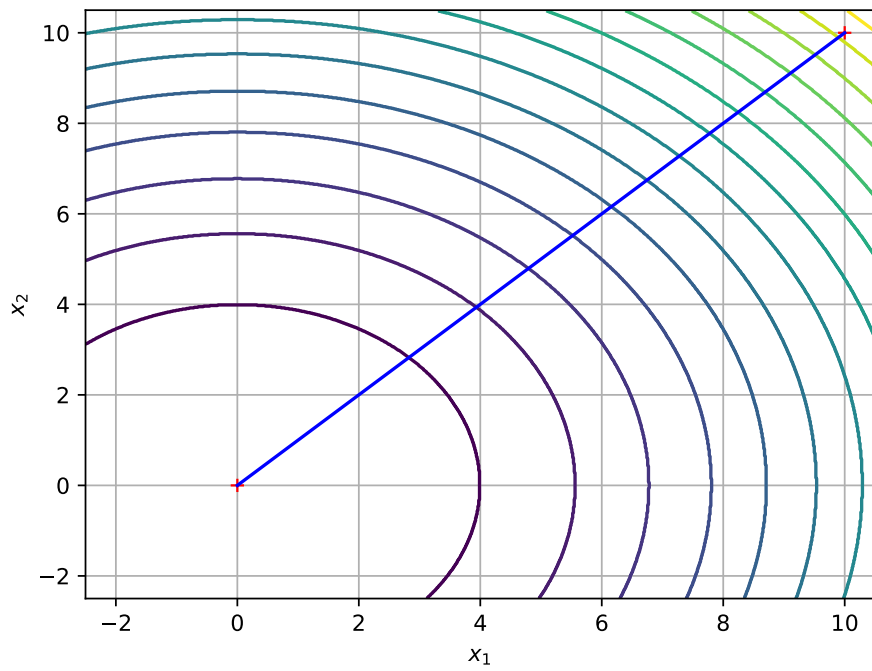


Rysunek 28: Współczynnik $a = 0.5$

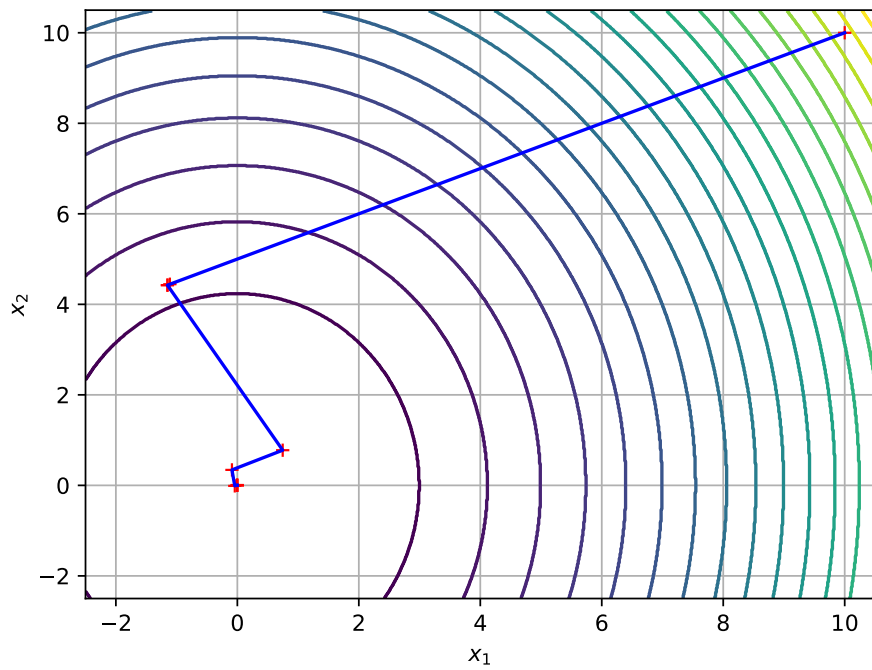


Rysunek 29: Współczynnik $a = 0.3$

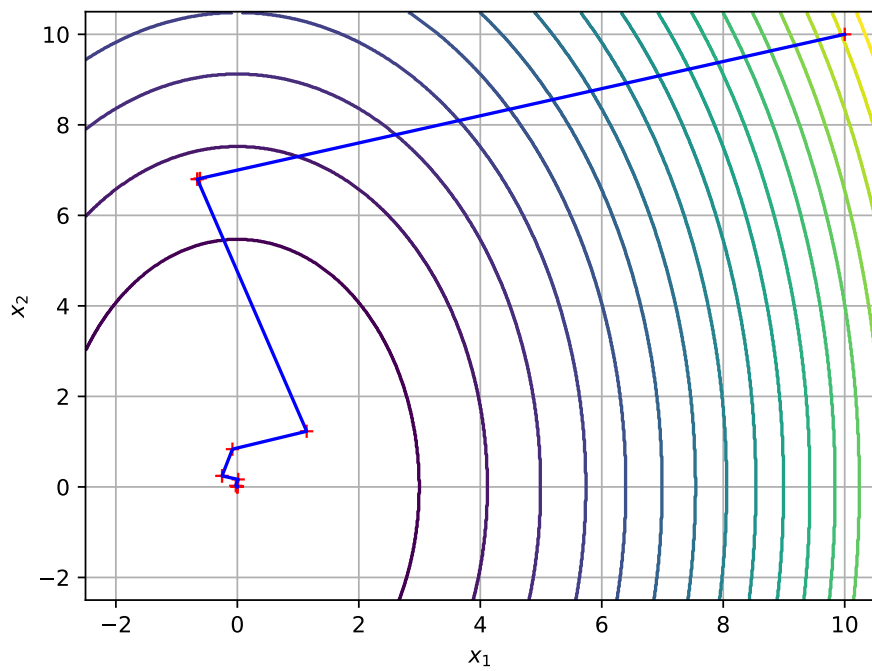
2.4.1 Metoda Broydena-Fletcher-Goldfarba-Shanno



Rysunek 30: Współczynnik $a = 1$

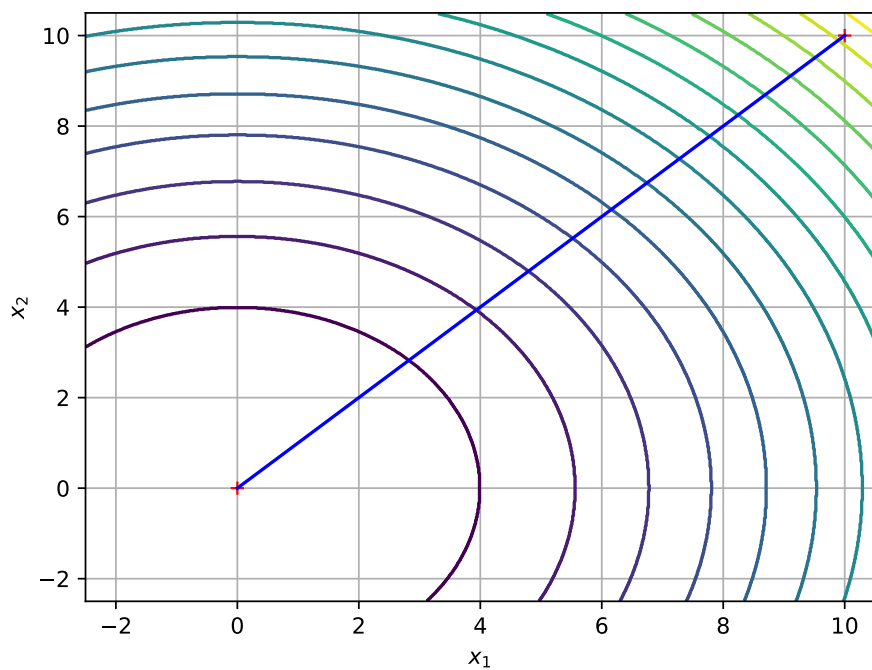


Rysunek 31: Współczynnik $a = 0.5$

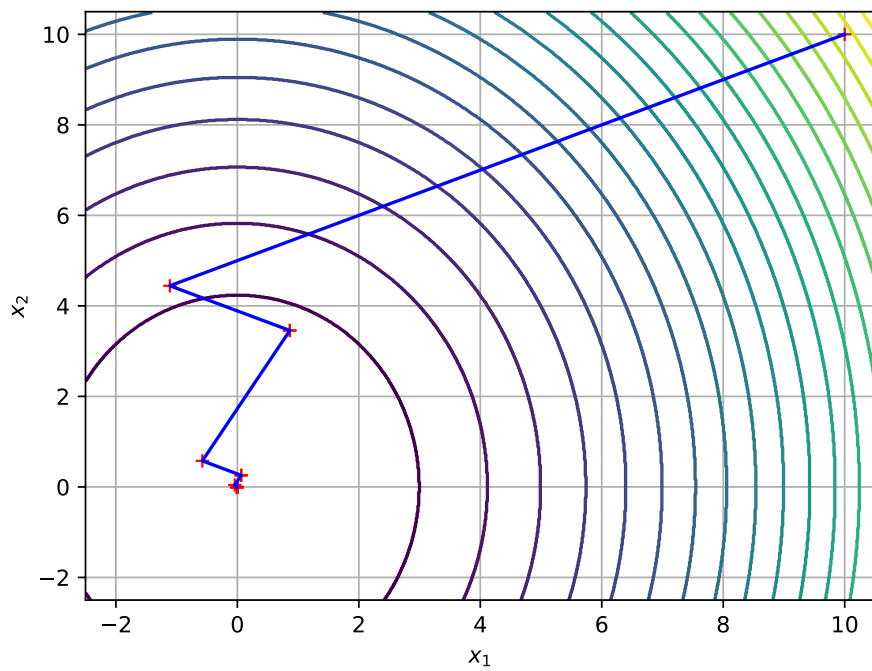


Rysunek 32: Współczynnik $a = 0.3$

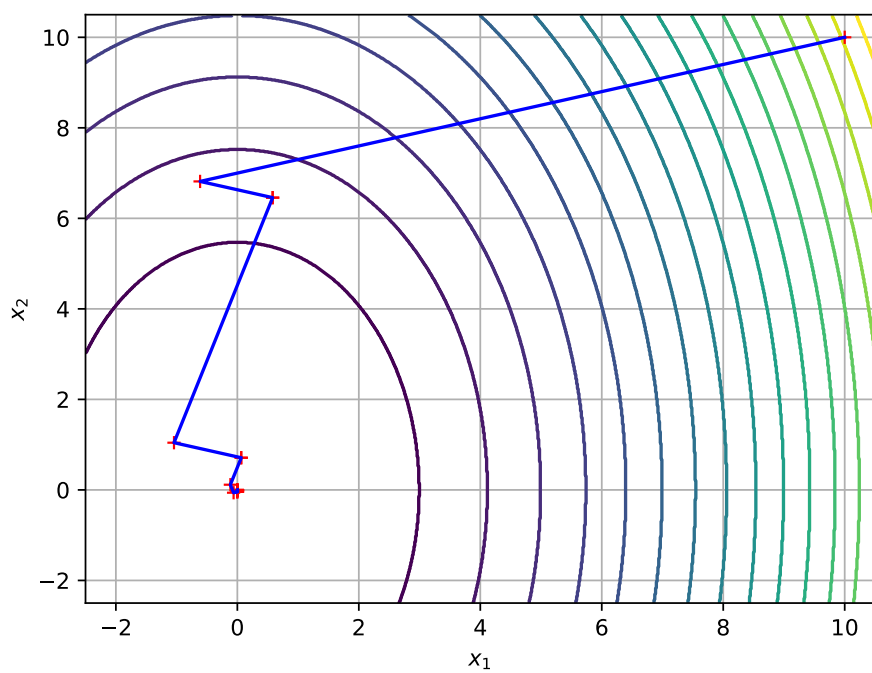
2.4.2 Pierwsza metoda Pearsona



Rysunek 33: Współczynnik $a = 1$

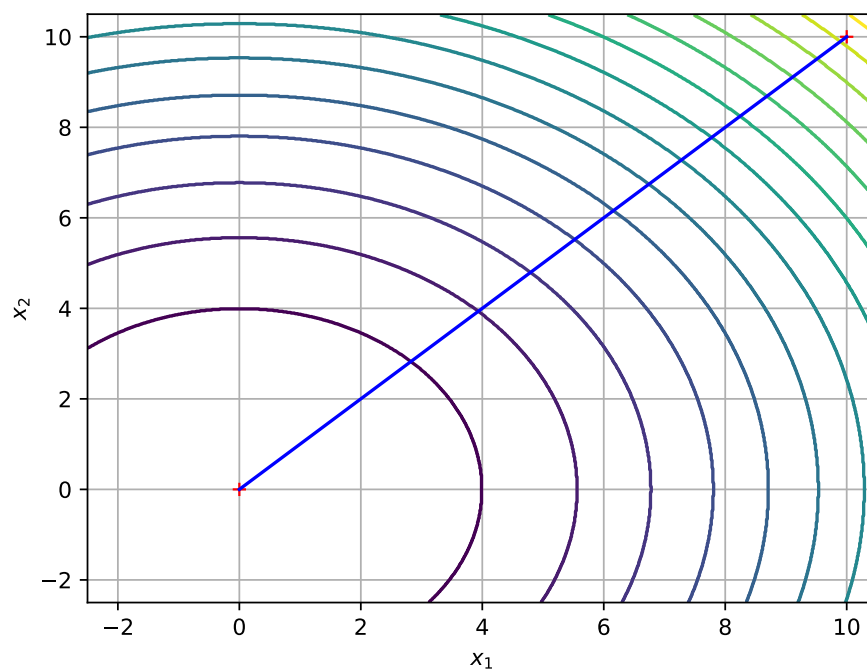


Rysunek 34: Współczynnik $a = 0.5$

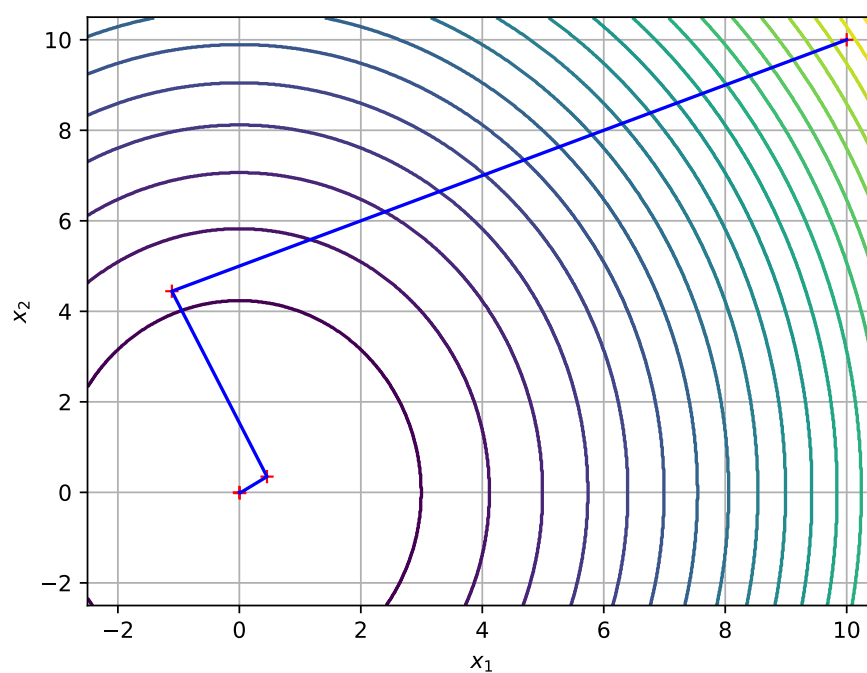


Rysunek 35: Współczynnik $a = 0.3$

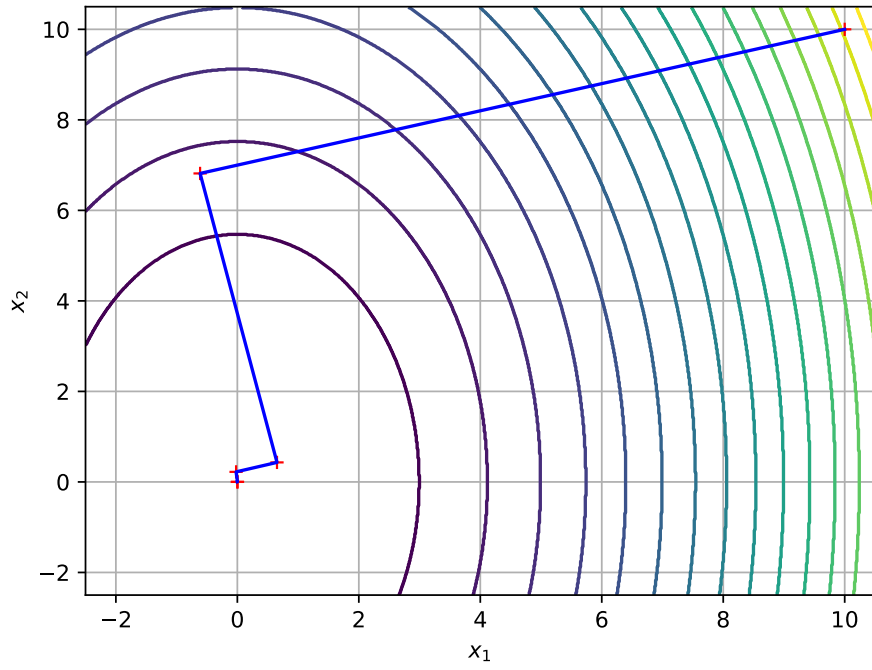
2.4.3 Druga metoda Pearsona



Rysunek 36: Współczynnik $a = 1$

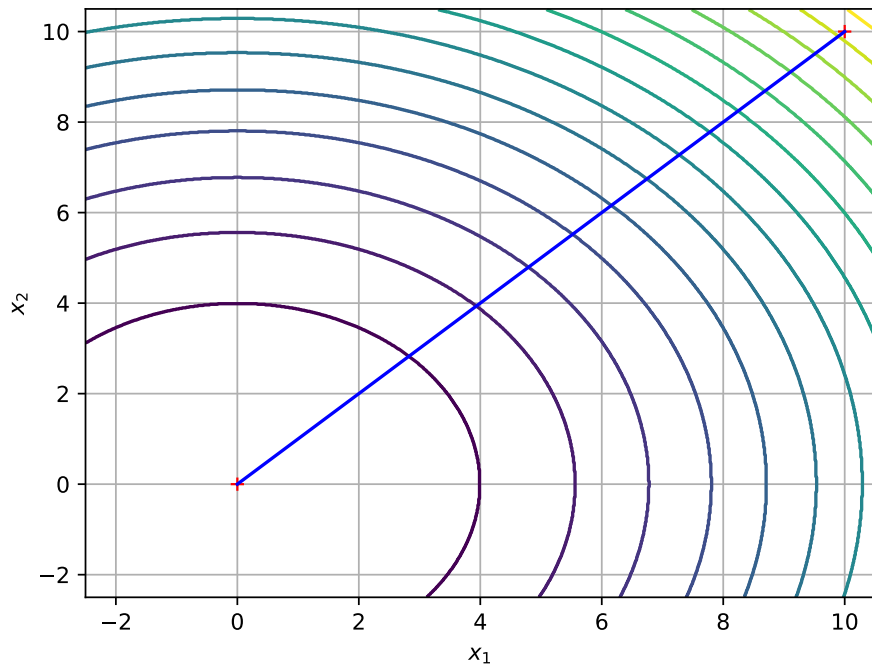


Rysunek 37: Współczynnik $a = 0.5$

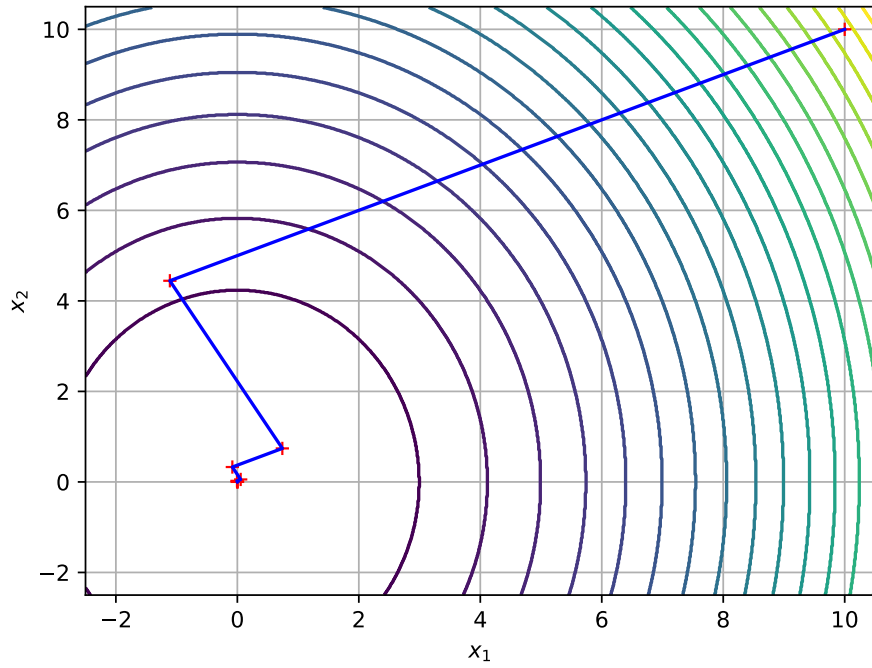


Rysunek 38: Współczynnik $a = 0.3$

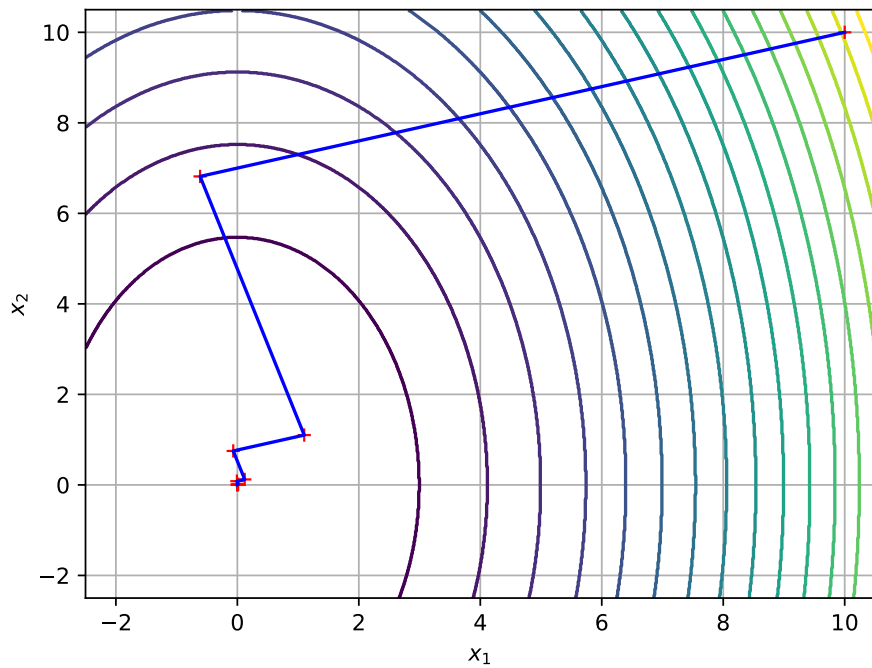
2.4.4 Metoda McCormicka



Rysunek 39: Współczynnik $a = 1$



Rysunek 40: Współczynnik $a = 0.5$



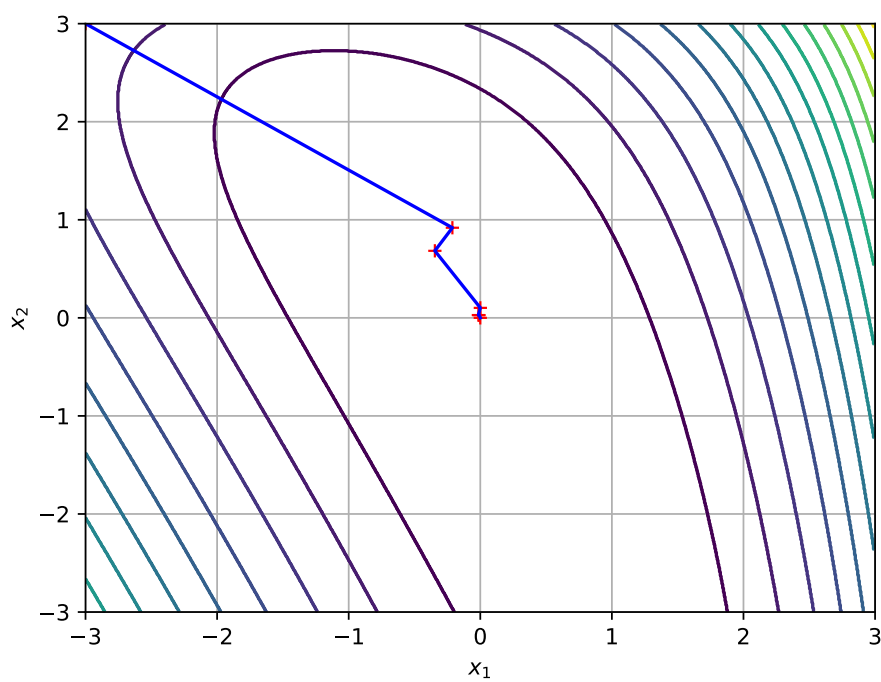
Rysunek 41: Współczynnik $a = 0.3$

Wszystkie metody znajdują w przypadku macierzy jednostkowej rozwiązanie w jednym kroku – nie wypadają w tym gorzej niż metody gradientowe. Wszystkie metody znajdują minimum szybko

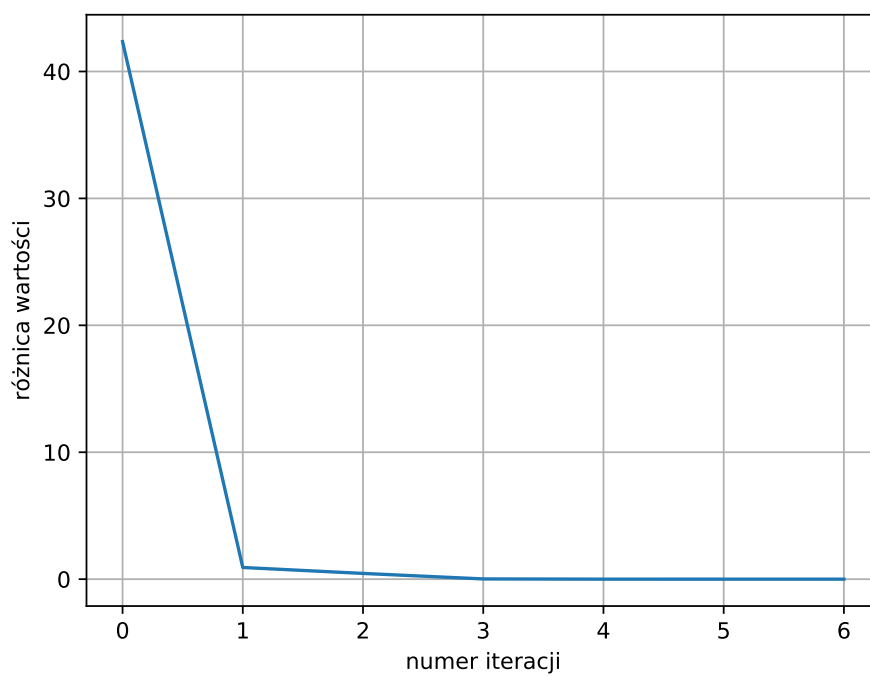
2.5 Ćwiczenie czwarte

Ćwiczenie czwarte polega na zbadaniu zachowania metod zmiennej metryki dla tej samej funkcji celu, co była rozważana w zadaniu drugim.

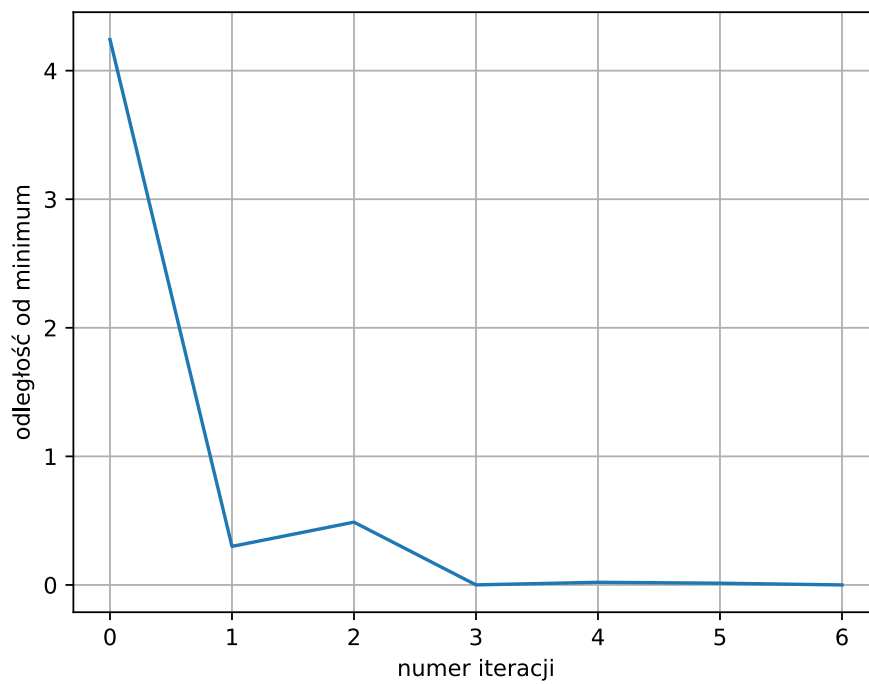
2.5.1 Metoda Davidona-Fletcher-Powella



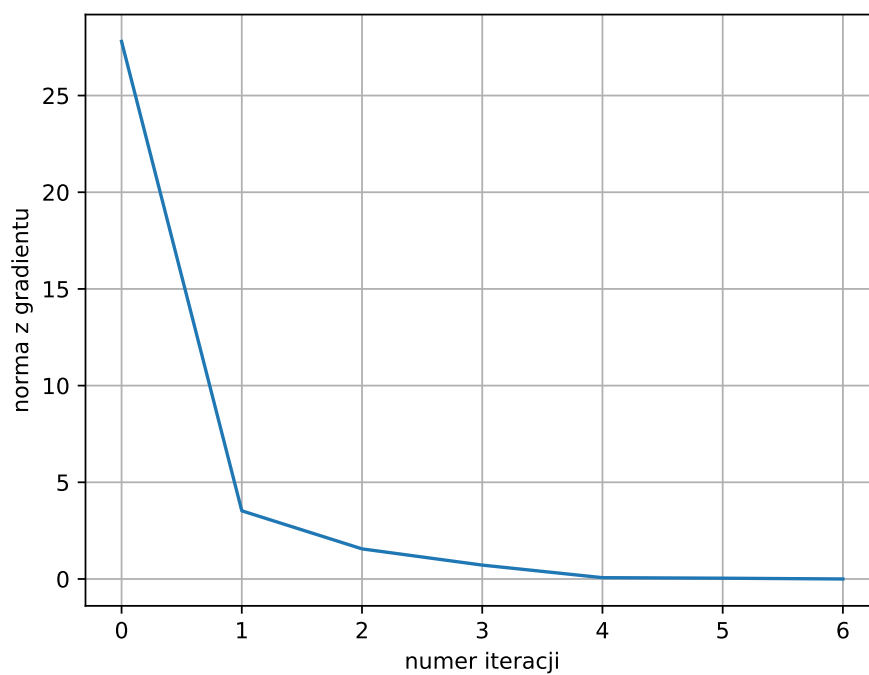
Rysunek 42: wyznaczone punkty



Rysunek 43: różnica między minimum, a aktualną wartością

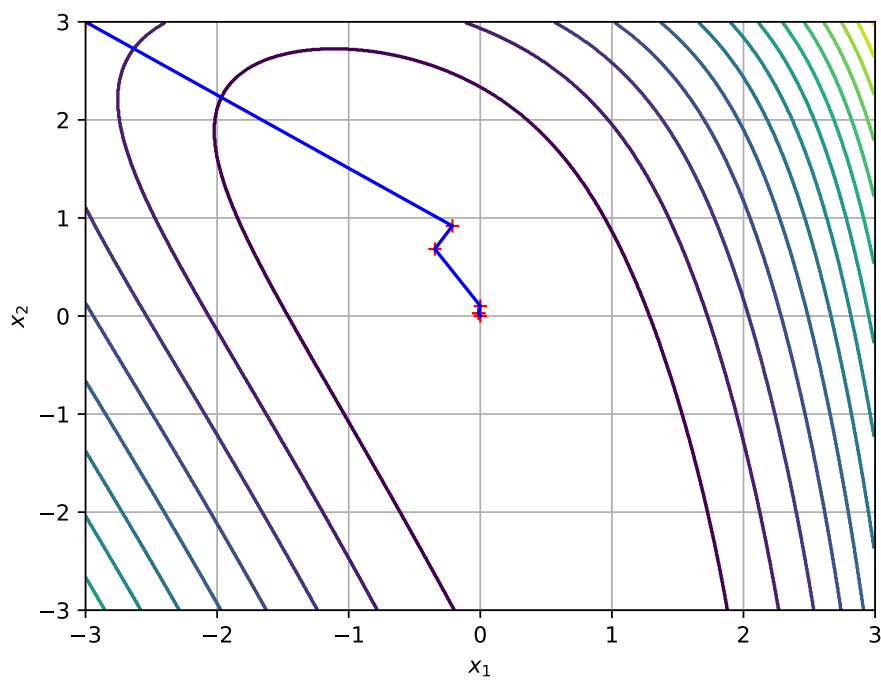


Rysunek 44: odległość od minimum

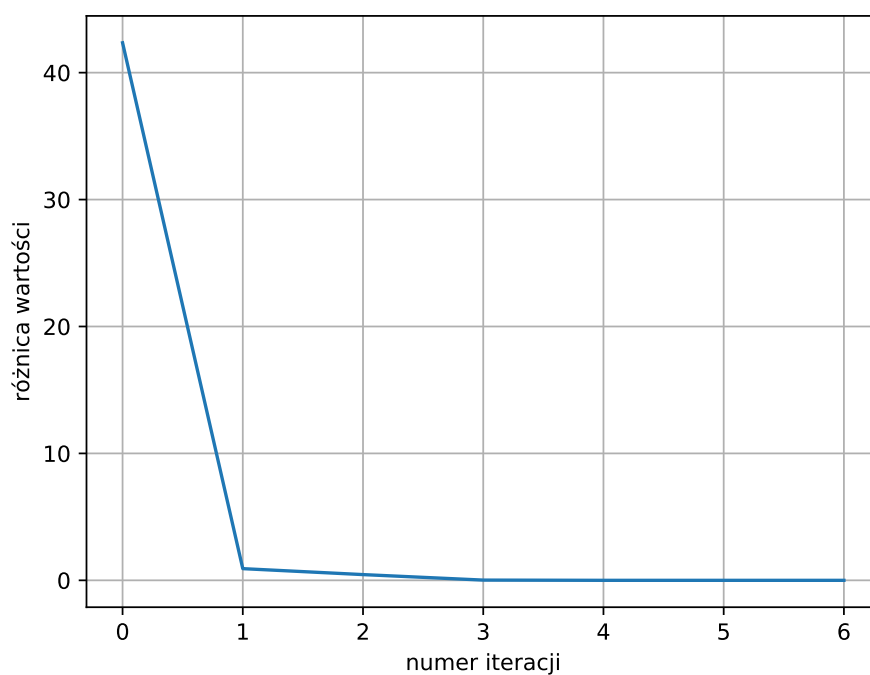


Rysunek 45: norma z gradientu

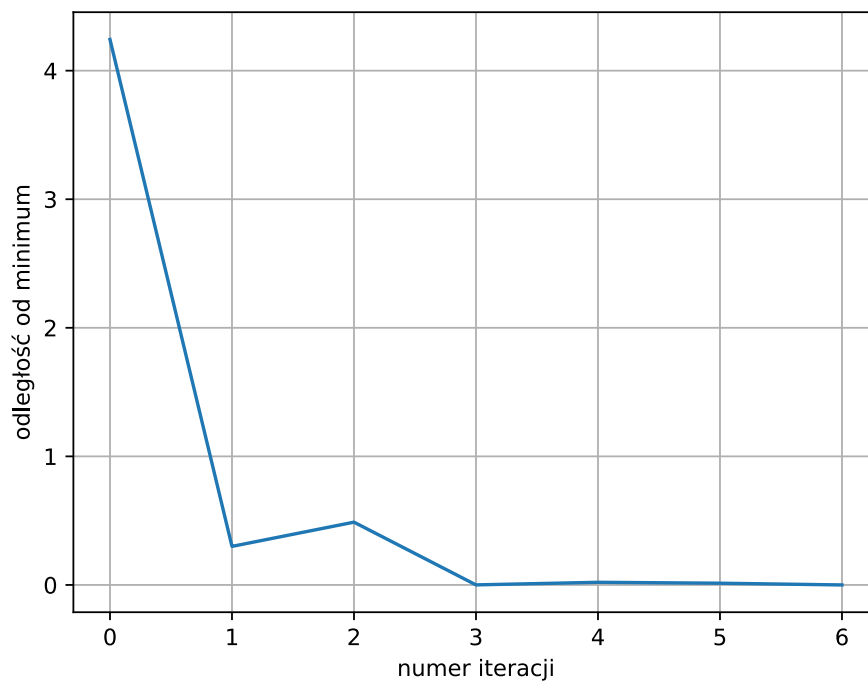
2.5.2 Metoda Wolfe'a-Broydena-Davidona



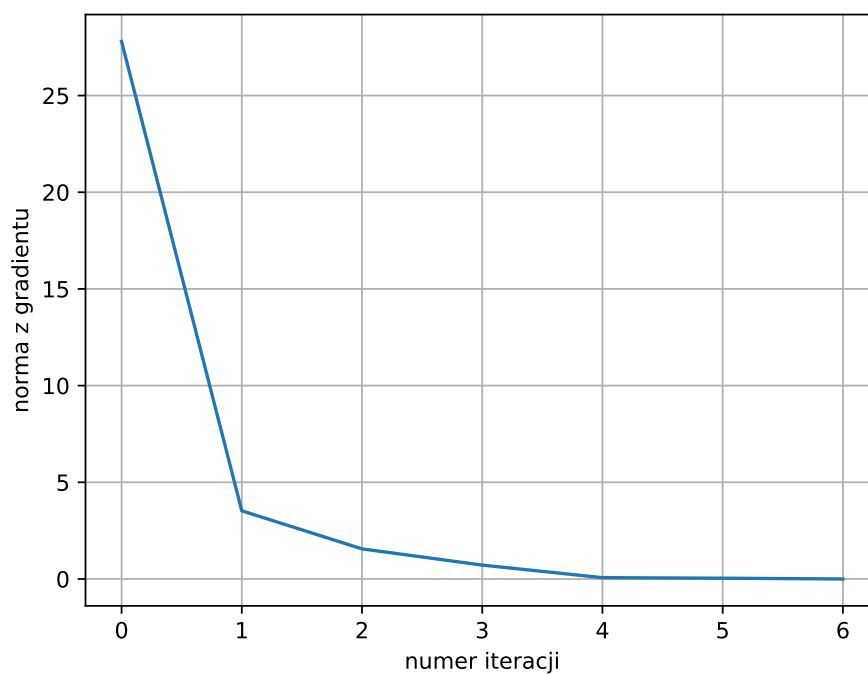
Rysunek 46: wyznaczone punkty



Rysunek 47: różnica między minimum, a aktualną wartością

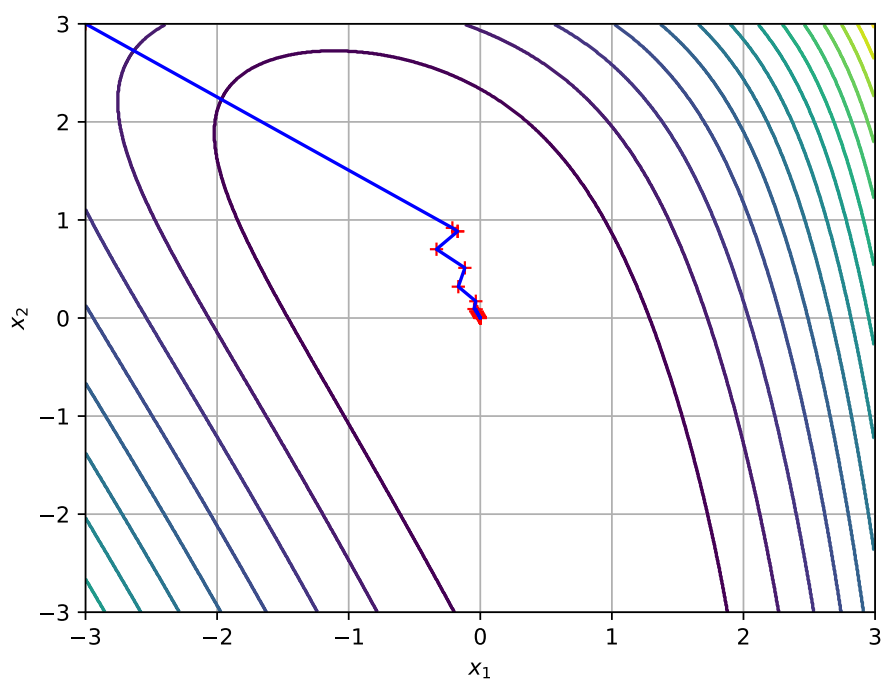


Rysunek 48: odległość od minimum

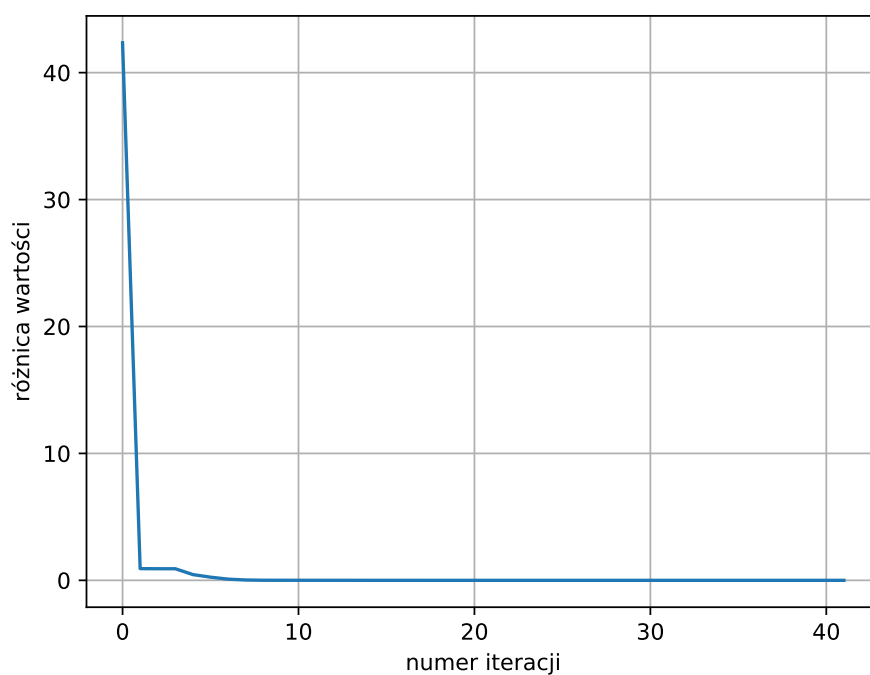


Rysunek 49: norma z gradientu

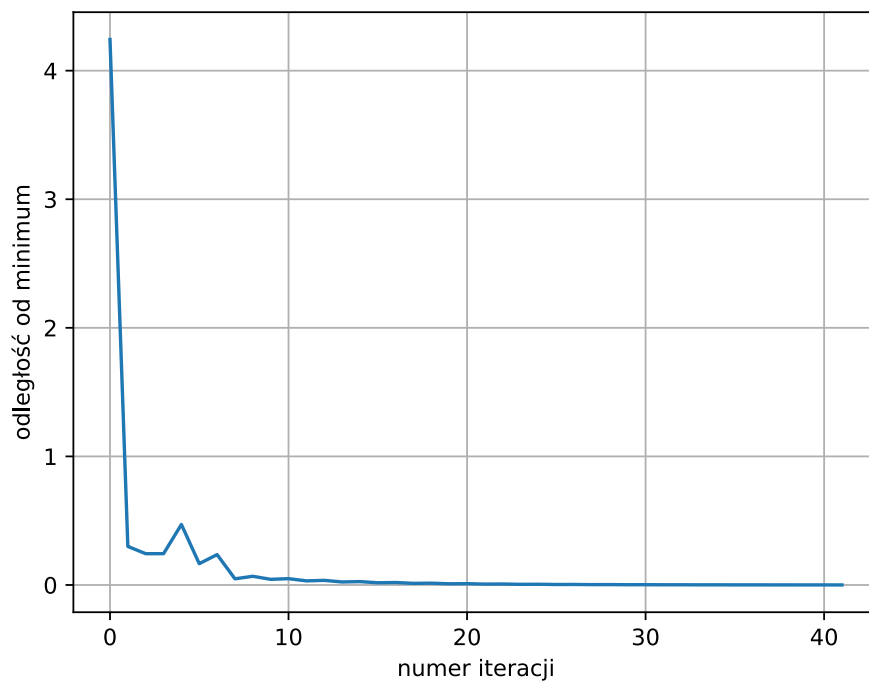
2.5.3 Metoda Broydena-Fletcher-Goldfarba-Shanno



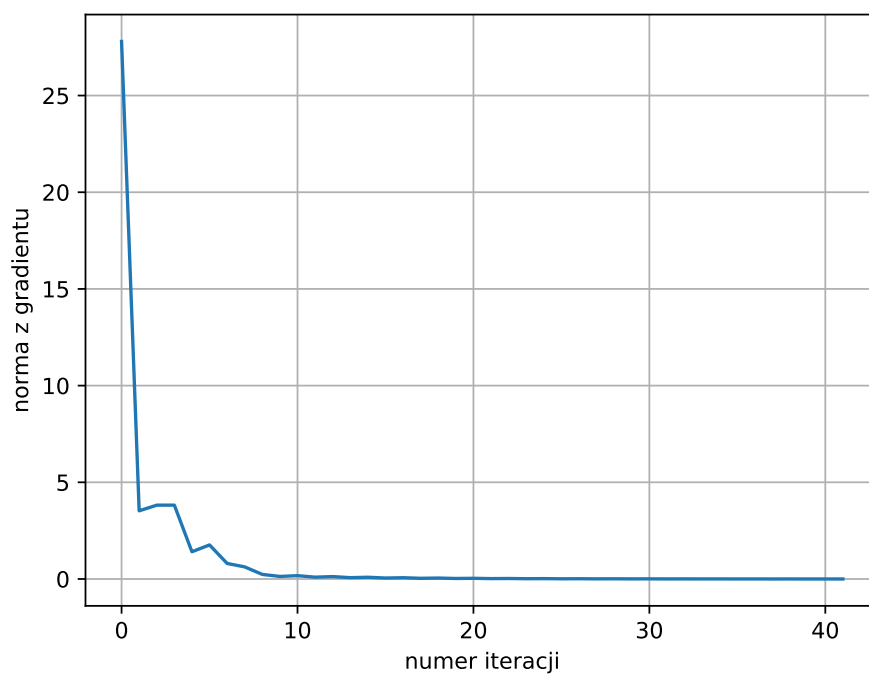
Rysunek 50: wyznaczone punkty



Rysunek 51: różnica między minimum, a aktualną wartością

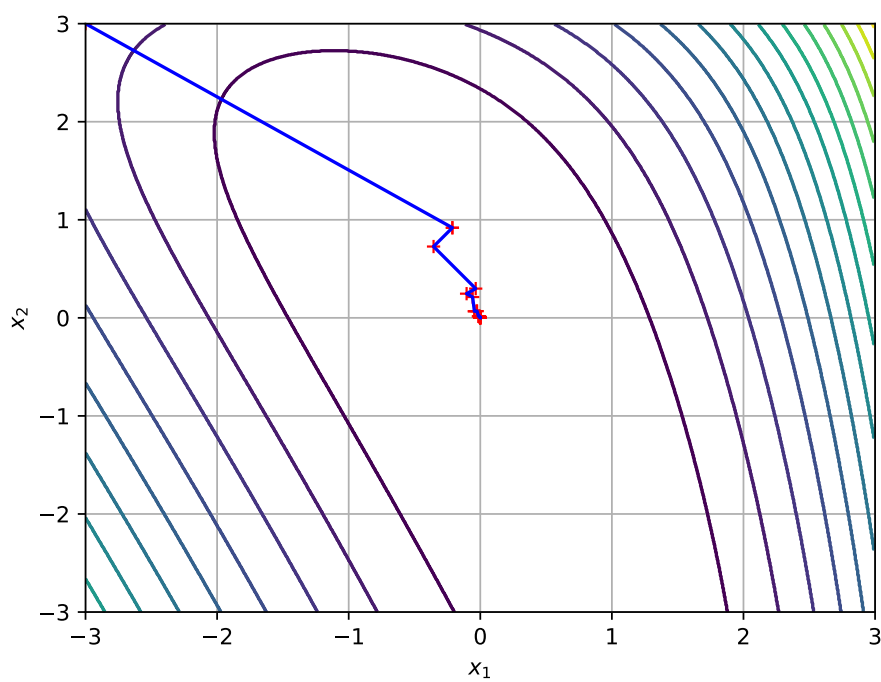


Rysunek 52: odległość od minimum

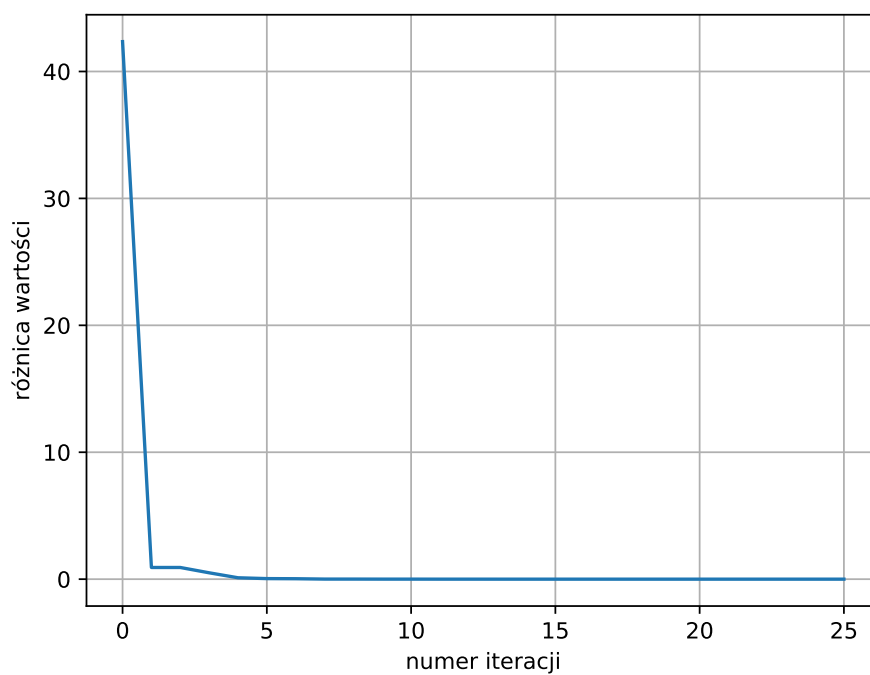


Rysunek 53: norma z gradientu

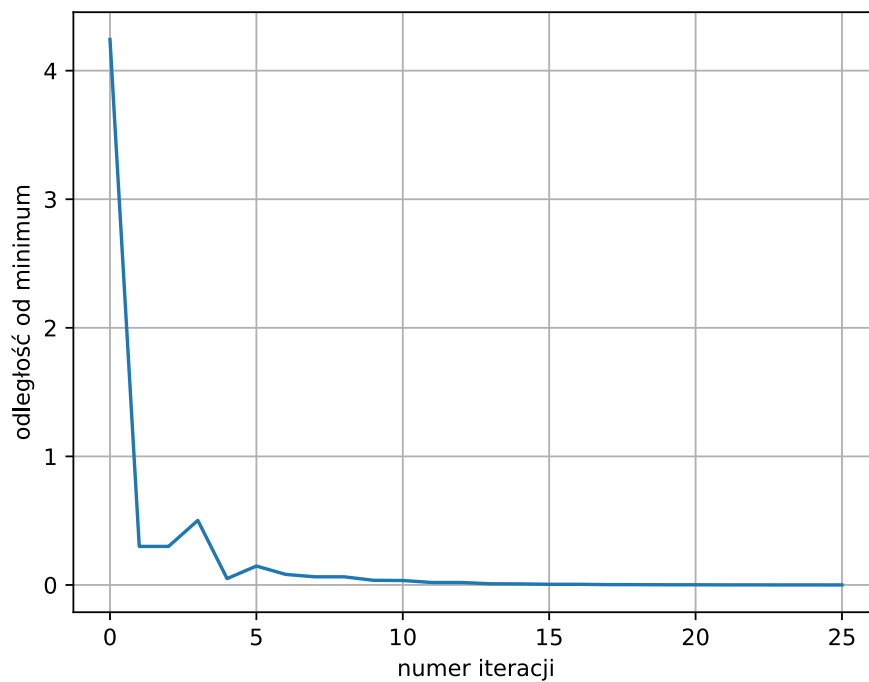
2.5.4 Pierwsza metoda Pearsona



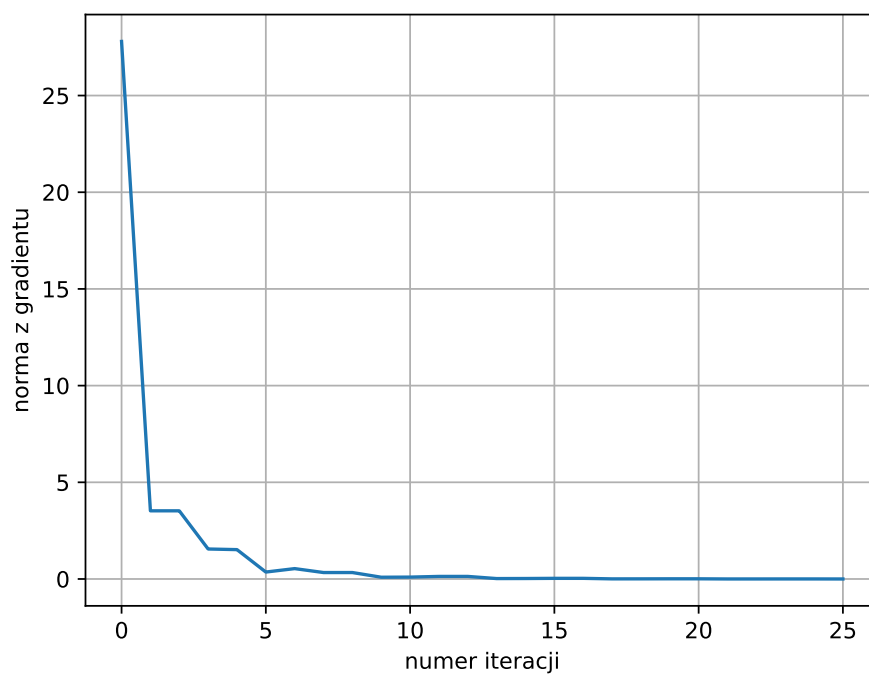
Rysunek 54: wyznaczone punkty



Rysunek 55: różnica między minimum, a aktualną wartością

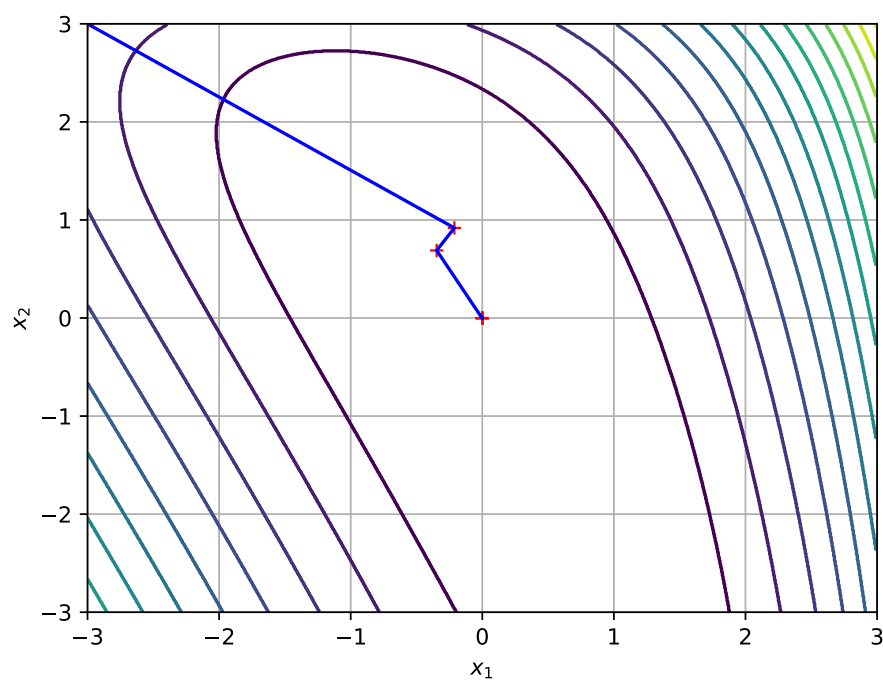


Rysunek 56: odległość od minimum

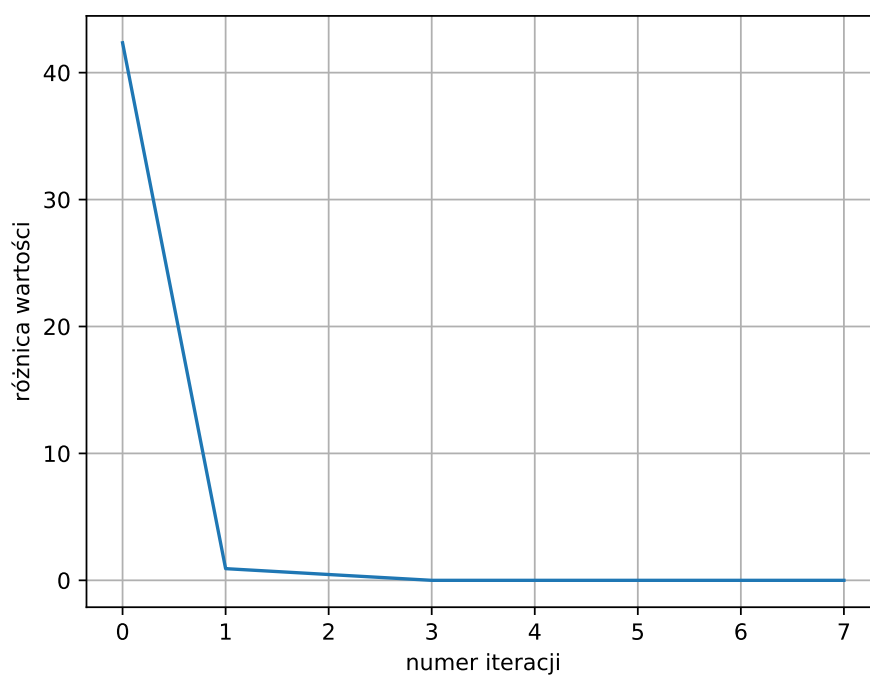


Rysunek 57: norma z gradientu

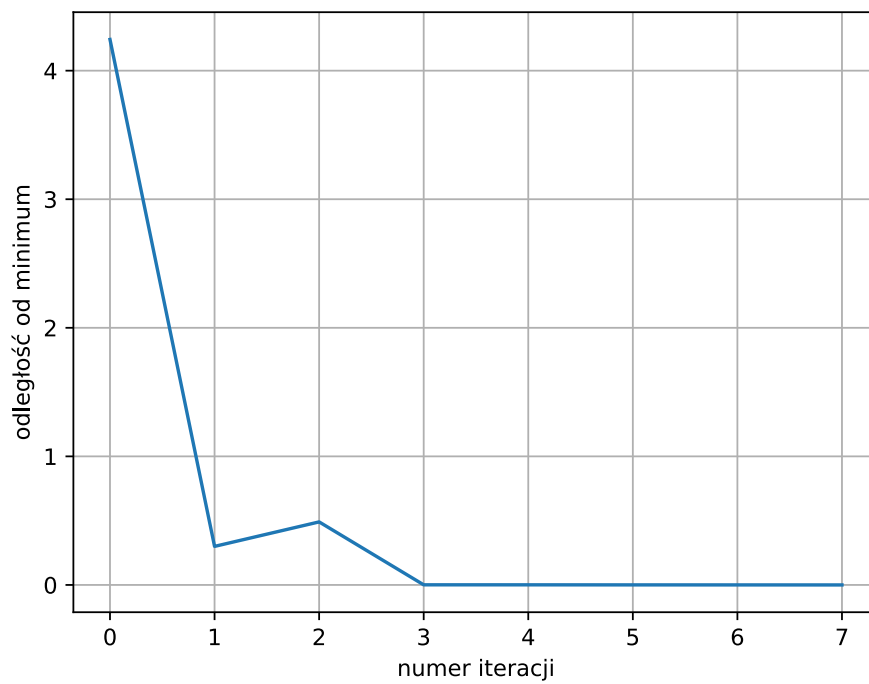
2.5.5 Druga metoda Pearsona



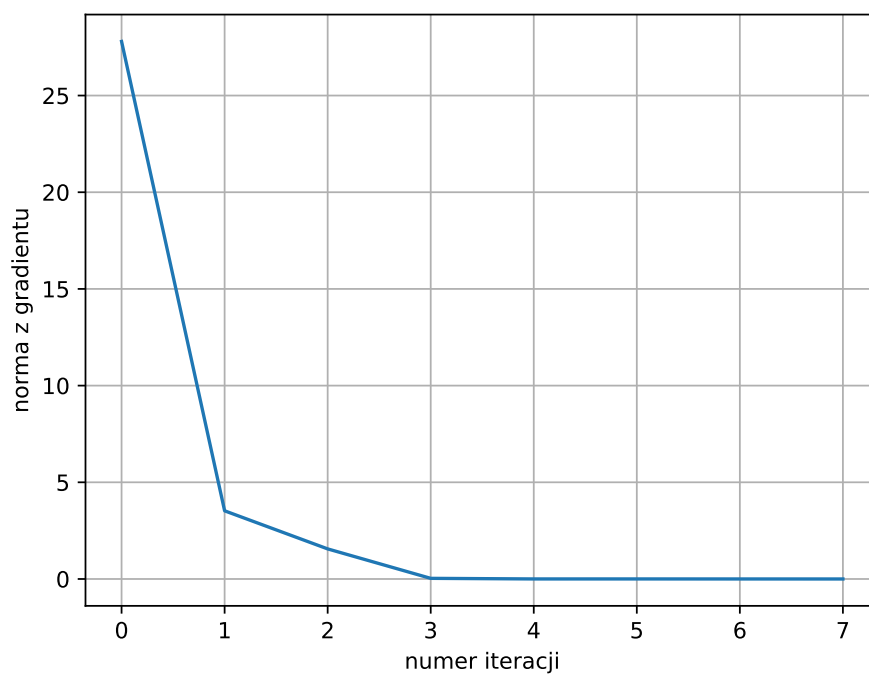
Rysunek 58: wyznaczone punkty



Rysunek 59: różnica między minimum, a aktualną wartością

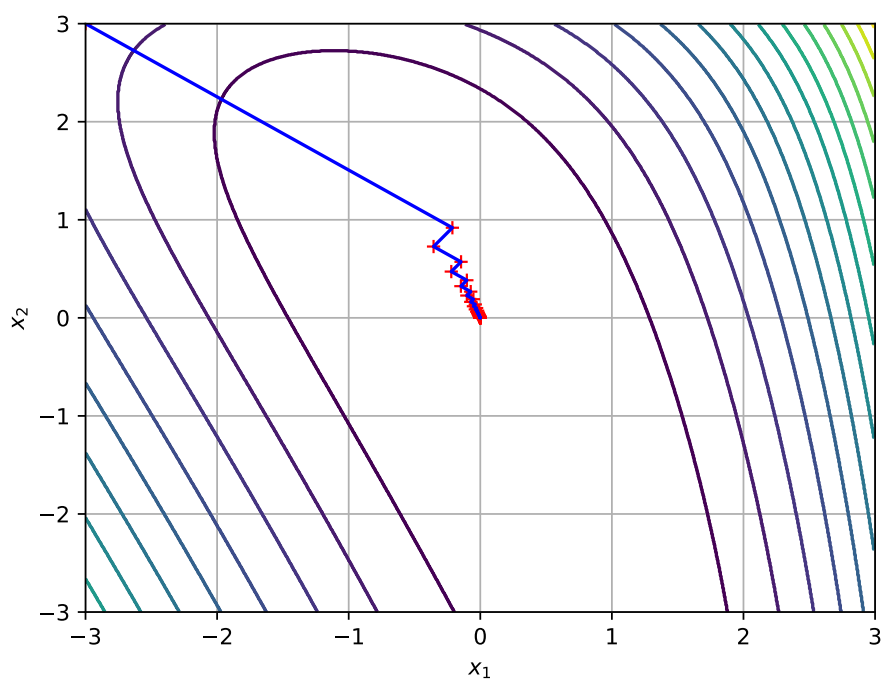


Rysunek 60: odległość od minimum

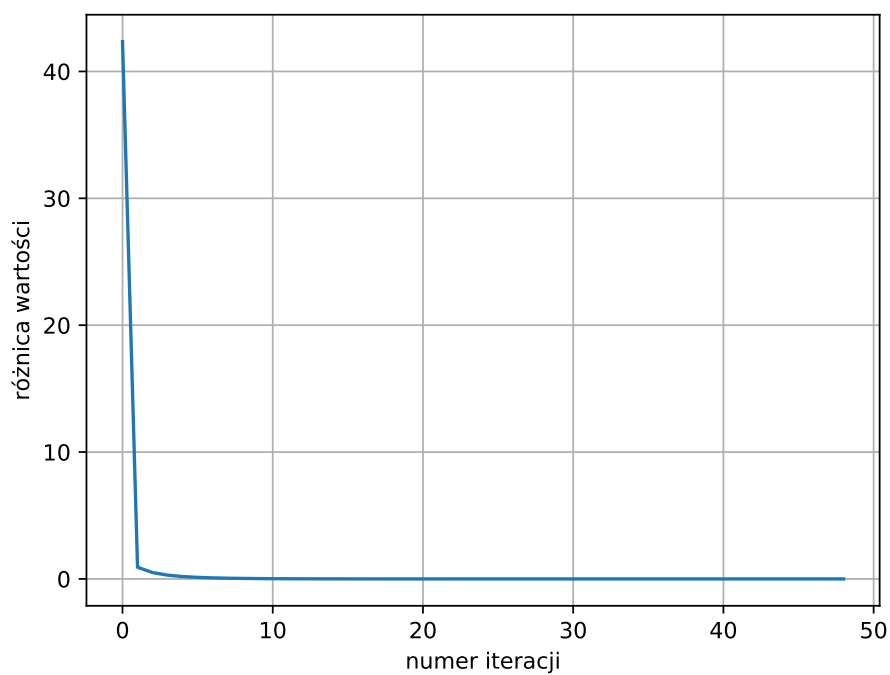


Rysunek 61: norma z gradientu

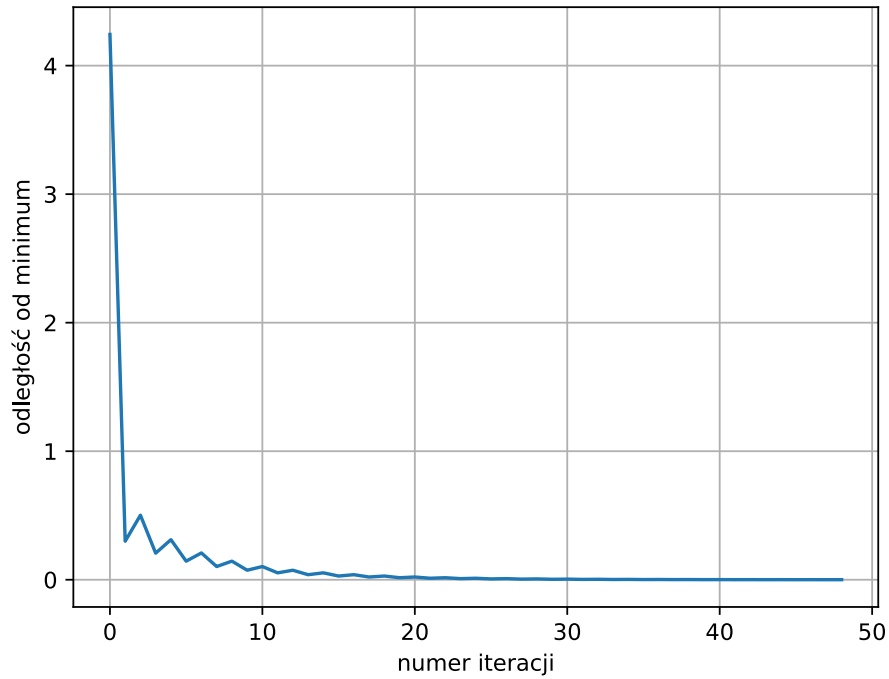
2.5.6 Metoda McCormicka



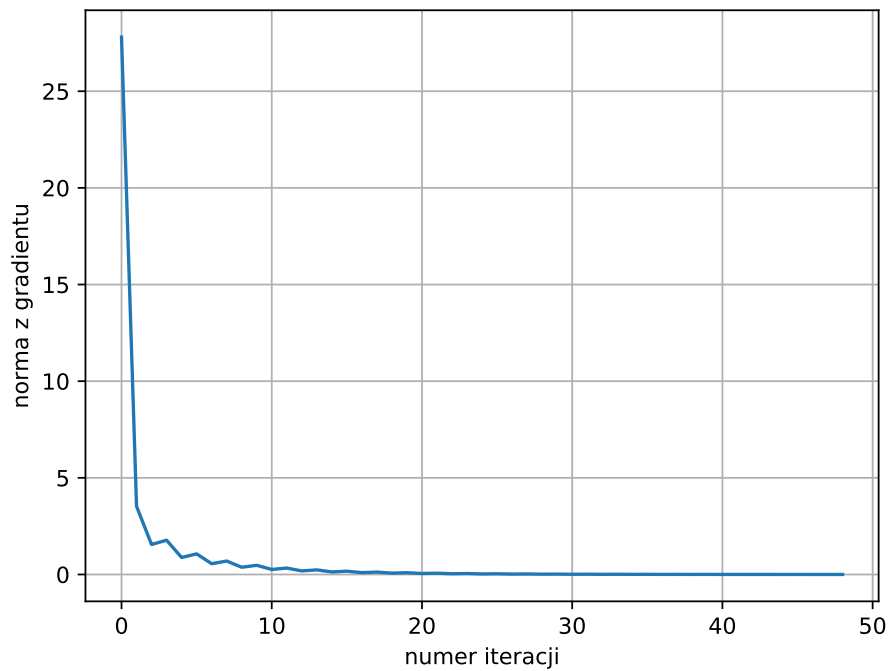
Rysunek 62: wyznaczone punkty



Rysunek 63: różnica między minimum, a aktualną wartością



Rysunek 64: odległość od minimum



Rysunek 65: norma z gradientu

3 Wnioski i obserwacje

Większość metod gradientowych sprawdza się bardzo dobrze w zadaniach optymalizacji. Są one bardzo szybko zbieżne, jednak mogą popaść w zygzakowanie, czemu się przeciwdziała wprowadzając przekształcenia zwiększające ich efektywność. Norma z Gradientu w wielu przypadkach bardzo szybko maleje, od tej

reguły odstaje jednak kilka metod: Metoda McCormicka, Pierwsza Metoda Pearsona, Metoda Broydena-Fletcher-Goldfarba-Shanno oraz Metoda Najszybszego Spadku.

Wyniki te jednak mogą być silnie zależne od punktu startowego, co widać po metodzie najszybszego spadku, która dzięki temu może mieć zastosowanie w pewnych zagadnieniach, redukując w ten sposób konieczną moc obliczeniową (gdyż jest to najprostsza metoda, dla której pozostałe są usprawnieniami).

Metody te mogą się oddalać od minimum w trakcie jego poszukiwań, zmniejszają jednak mimo to wartość funkcji celu. Część metod korzysta z tego, że w miejscach, gdzie funkcja jest bardziej stroma można znaleźć szybciej minimum.