Wydział Elektroniki i Technik Informacyjnych
Politechnika Warszawska

Wprowadzenie do sztucznej inteligencji

Sprawozdanie z ćwiczenia nr 1

Tymon Kobylecki

Warszawa, 2022

Spis treści

1. Wstęp

W niniejszym sprawozdaniu opisane zostało rozwiązanie zadania oraz eksperymenty dotyczące zadania nr 1. Całość ćwiczenia została wykonana w języku Python.

2. Ćwiczenie

2.1. Eksperymenty

Podczas eksperymentów zmieniane były wartości parametru β oraz punktu początkowego. Tolerancja w każdym eksperymencie była nastawiona na 0,000 01. Jest ona tożsama z minimalną długością kroku wymaganą do kontynuacji działania algorytmu. Funkcja f posiada minimum w punkcie $x = 0$, zaś funkcja g posiada dwa minima lokalne: jedno w punkcie $x_1 = 0, x_2 = 0$, a drugie w punkcie $x_1 = 1, x_2 = -2$. Globalnym minimum jest to pierwsze.

Podczas eksperymentów parametr β był manipulowany w zakresie $< 0{,}01; 0{,}3 >$, zaś punkt początkowy był manipulowany zależnie od postaci funkcji.

2.2. Wyniki

2.2.1. Funkcja f

Pierwszą funkcją poddaną algorytmowi była funkcja f. Warto nadmienić, że nie zawsze osiągnięcie minimum jest możliwe - wynika to ze zebyt dużej miary parametru β w porównaniu do odległości punktu poczatkowego od poszukiwanego minimum. Wówczas zachodzi rozbieżność algorytmu, wobec czego rozwiązanie zadania znalezienia minimum funkcji jest niemożliwe.

β, punkt początkowy	1	3	10
0,01	$x = 0{,}01$, 67843 kroki	$x = 0{,}01$, 67852 kroki	BRAK
0,1	$x = 0{,}01$, 6779 kroków	BRAK	BRAK
0,3	$x = -0{,}01$, 2250 kroków	BRAK	BRAK

2.2.2. Funkcja g

W kolejnym etapie eksperymentom poddana została funkcja g. Jej wykres jednak cechuje się wartością stałą w wielu punktach tworzących poziomą płaszczyznę, wobec czego gradient w tych miejscach wynosi zawsze 0. Metoda gradientowa zwraca wówczas najbliższe minimum lokalne, czyli punkt początkowy. Dodatkowo, funkcja ma jeszcze dwa minima lokalne: w punkcie (0;0) oraz (1;-2).

β, punkt początkowy	(1, 1)	(2, -3)	(10, 9)
0,01	$x_1 = x_2 = 0{,}00$ 705 kroków	$x_1 = 0{,}99, x_2 = -1{,}97$ 1496 kroków	$x_1 = 10, x_2 = 9$ 1 krok
0,1	$x_1 = x_2 = 0{,}00$ 67 kroków	$x_1 = 0{,}98, x_2 = -1{,}97$ 147 kroków	$x_1 = 10, x_2 = 9$ 1 krok
0,3	$x_1 = x_2 = 0{,}00$ 19 kroków	$x_1 = 0{,}98, x_2 = -1{,}97$ 46 kroków	$x_1 = 10, x_2 = 9$ 1 krok

2.3. Wnioski

Metoda gradientowa jest skuteczną metodą znajdowania ekstremów funkcji. Istotnym jednak jest poprawny dobór parametru β oraz punktu początkowego.

Ten pierwszy odpowiada za długość skoku między kolejnymi iteracjami. Jeśli jest więc za wysoki, skoki również są zbyt duże, wobec czego może dojść do rozbieżności algorytmu, wskutek czego minimum nie zostaje znalezione - algorytm zwraca liczby o coraz większym module, aż przekroczy zakres zmiennej typu float. Nie należy jednak ustalać parametru β na najmniejszą możliwą wartość, gdyż skutkuje to zwiększeniem (często znacznym) liczby kroków, a zarazem czasu działania algorytmu.

Dobór punktu początkowego jest istotny, ponieważ determinuje, które minimum zostaje znalezione przez algorytm w przypadku, gdzie mamy do czynienia z funkcją o wielu minimach lokalnych. Taką funkcją jest g. Metoda gradientowa opiera się na znajdowaniu minimum poprzez szukanie pierwiastków pochodnej analizowanej funkcji, wobec czego na odcinkach funkcji o stałej wartości algorytm odnajdzie minimum lokalne w punkcie początkowym (jeśli ten znajdował się na takim odcinku). Z definicji minimum lokalnego wynika bowiem, że w otoczeniu minimum lokalnego nie może być punktów o wartości większej niż wartość w potencjalnym minimum. Punkty na płaskim odcinku, lub na płaszczyźnie (w przypadku funkcji dwuwymiarowej) spełniają ten warunek - zatem algorytm „uznaje", że zakończył swoje zadanie. Tak samo działa to w każdym minimum lokalnym.