COSC 6364: Advanced Numerical Analysis

Assignment #2

Name: Ashutosh Kumar (1854393) Due on: 11 April 2023

1 Introduction

For this assignment, we are required to implement the gradient descent to find the minimum of a function. The function that we will be using is

$$w_1^2 + w_1 \times w_2 + 2 \times w_2^2$$

2 Method

For gradient descent, we have to set a few parameters.

First is λ, which represents the learning rate. In this assignment, I started with $\lambda = 0.1$. I kept increasing the value of the learning rate until the gradient descent failed at $\lambda = 0.28$.

Further, we have to set a starting point, the maximum number of steps (taken to be 1000 here), and the termination tolerance. I set the value of termination tolerance to $1e - 6$. What this means is that if the difference between our current iteration and the next iteration is smaller than this, we have arrived at the desired solution.

3 Results

3.1 Output for $\lambda = 0.1$

The output and plot for learning rate $\lambda = 0.1$ are given below.

```
Lambda (Learning rate) = 0.1
Termination tolerance = 1e-06
Optimal number of steps = 71
Optimal w = [ 5.63569070e-06 -2.33437952e-06]
Minimum f = 2.95038241867847e-11
```

Figure 1: Result for $\lambda = 0.1$

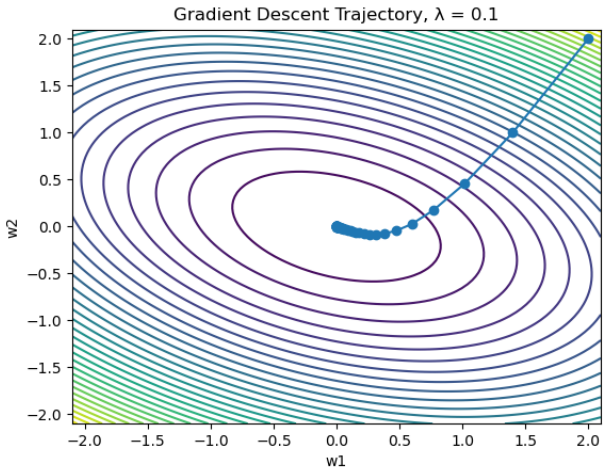
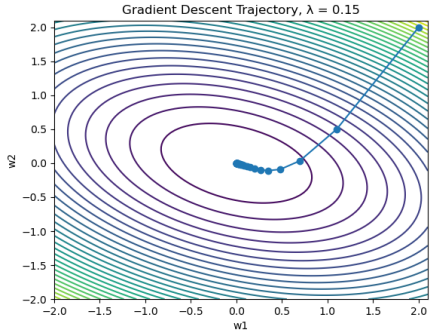
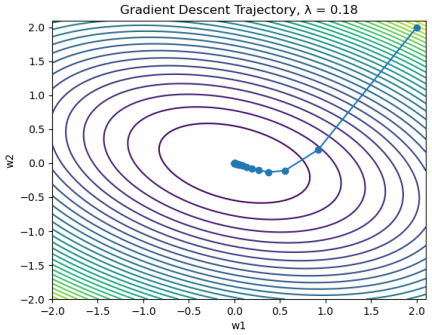


Figure 2: Plot for $\lambda = 0.1$

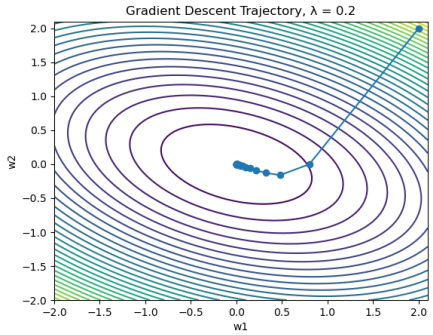
3.2 Plots for other learning rates



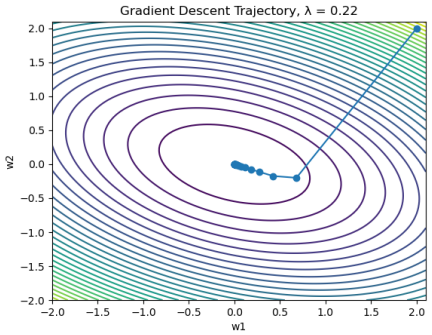
(a) Plot for $\lambda = 0.15$



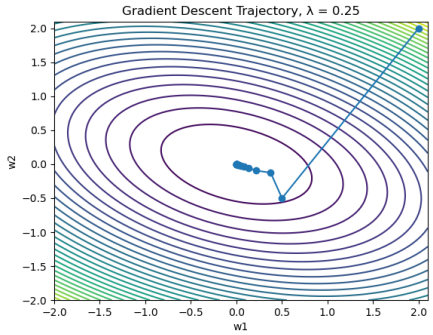
(b) Plot for $\lambda = 0.18$



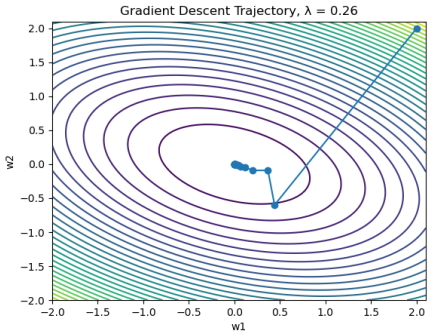
(c) Plot for $\lambda = 0.2$



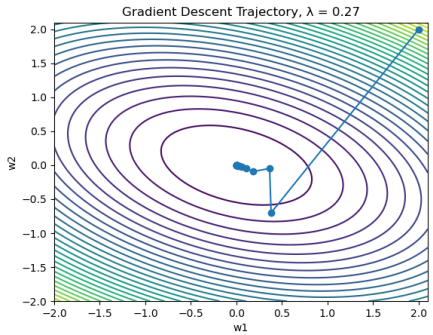
(d) Plot for $\lambda = 0.22$



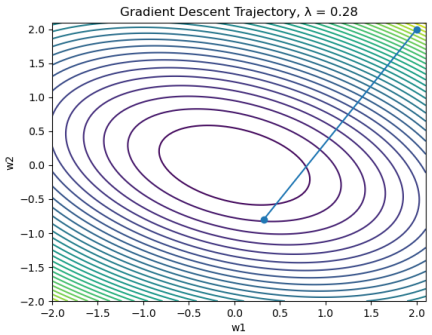
(e) Plot for $\lambda = 0.25$



(f) Plot for $\lambda = 0.26$



(g) Plot for $\lambda = 0.27$



(h) Plot for $\lambda = 0.28$

Figure 3: Plots for various learning rates