HOMEWORK 3

Problem 1. The dataset sales.csv contains the daily sales of a product for 1000 days.Visually demonstrate that the daily sales follows a normal distribution.

In this problem, you will be fitting a normal distribution to the observed dataset by minimizing the negative log-likelihood function. Using the probability distribution function for a normal distribution $N(\mu, \sigma^2)$ with two parameters: mean μ and variance σ^2, obtain the negative log-likelihood function. Using gradient descent, minimize the negative log-likelihood function to estimate the parameters μ and σ that best explain the data.

Problem 2. Consider two objective functions $f(x, y) = (x - 5)^2 + 2(y + 3)^2 + xy$ and $g(x, y) = (1 - (y - 3))^2 + 10((x + 4) - (y - 3)^2)^2$.

Starting with $(x, y) = (0, 2)$ run the gradient descent algorithm to minimize each function. Plot the value of the objective function after each gradient descent step. Use a fixed learning rate of $\gamma = 0.05$ for objective function f and $\gamma = 0.0015$ for objective function g.

Then, employ an exponential decay learning rate and an inverse decay learning rate. Compare their performance with a constant learning rate. Can you tune the parameters to achieve faster optimization with either of the decaying learning rates? What is the least number of iterations within which you are able to optimize the two functions?