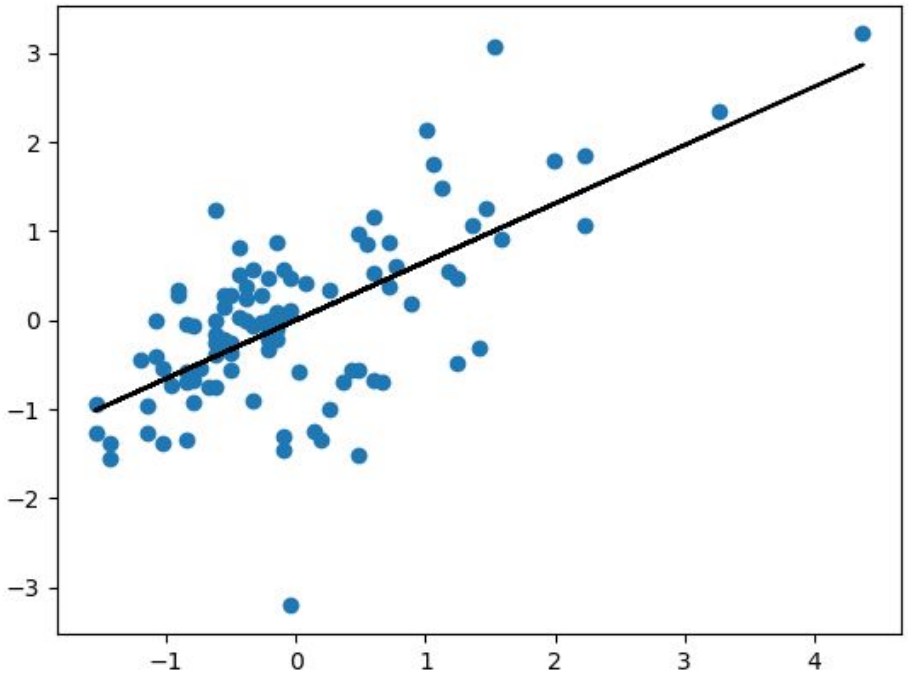
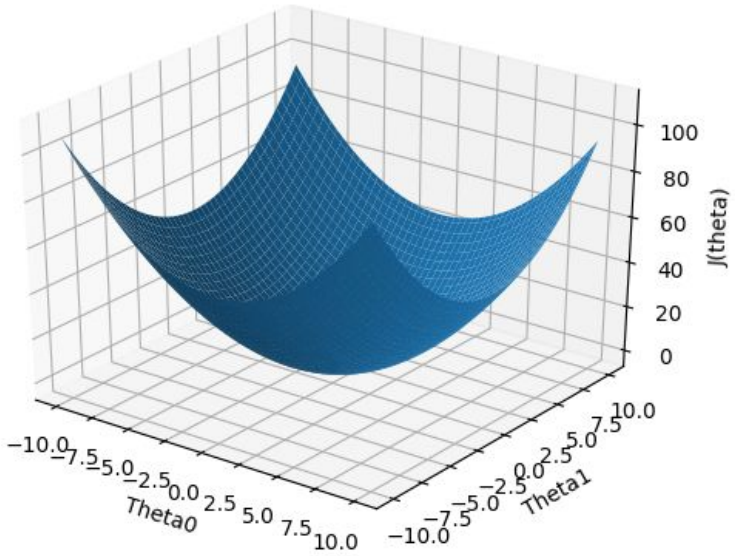
Answer 1

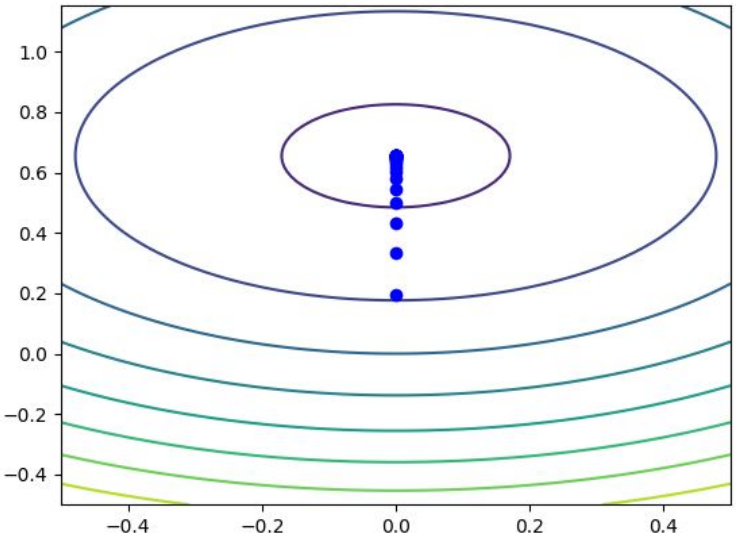
lr=0.001

The stopping criteria I have used is that the value of the gradient is very close to 0 (less than 10^-4)

Theta=(-5.991690033013666e-14, 0.6550634275627529)



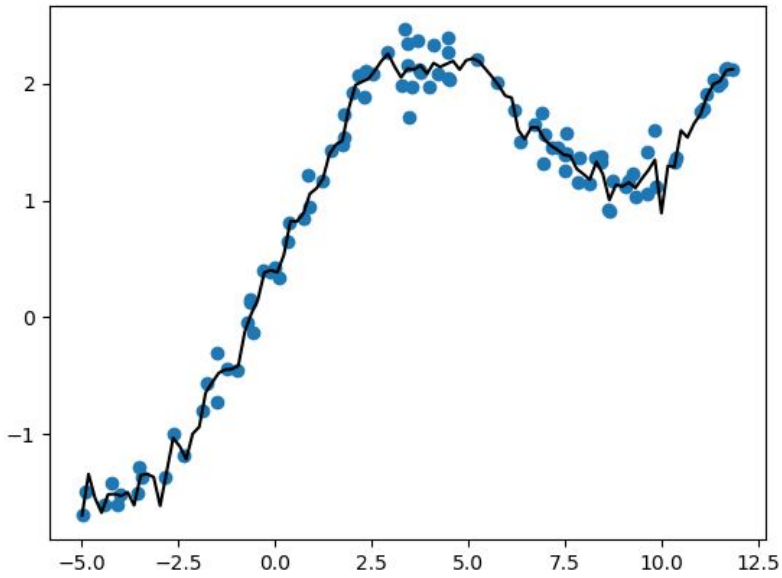

100
80
60
40
20
0
J(theta)
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
Theta0
Theta1


1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.4
−0.2
0.0
0.2
0.4

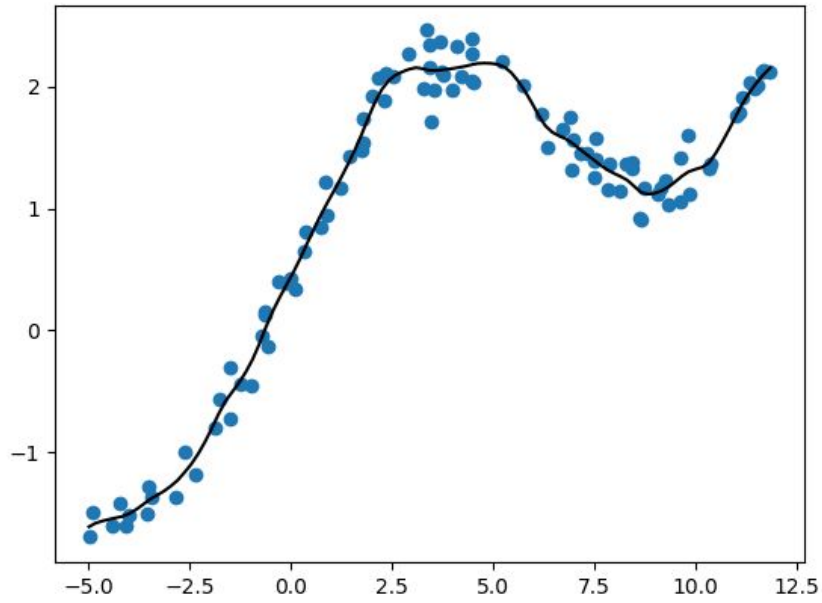
lr=0.3

Answer 2:

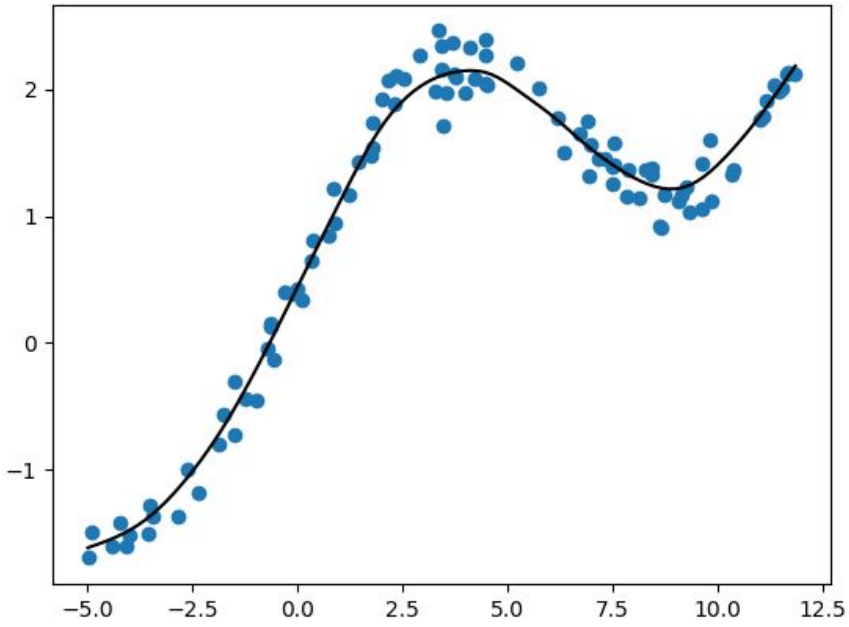
Tau=0.1



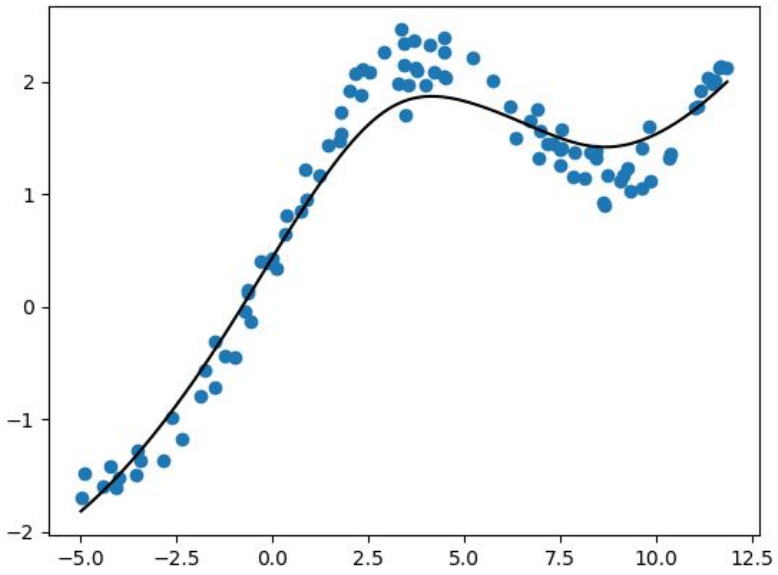
Tau=0.3



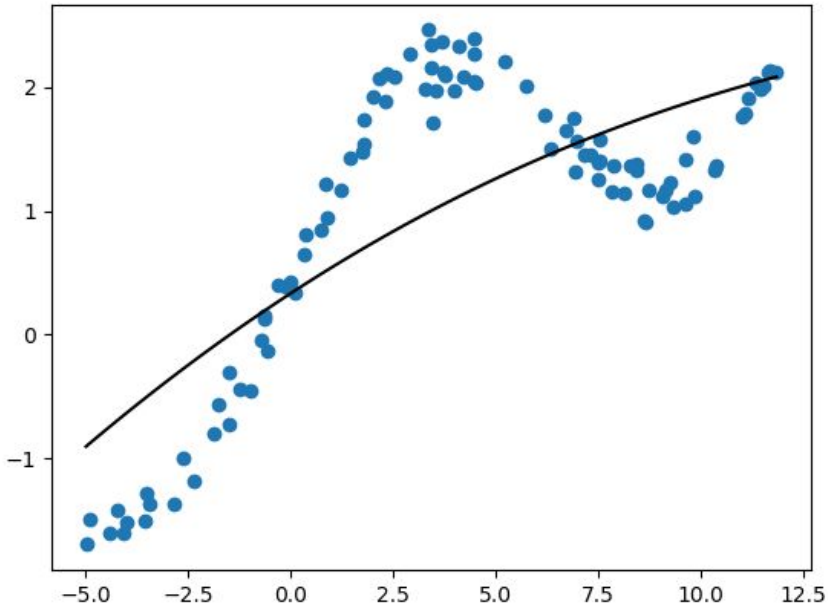
Tau=0.8



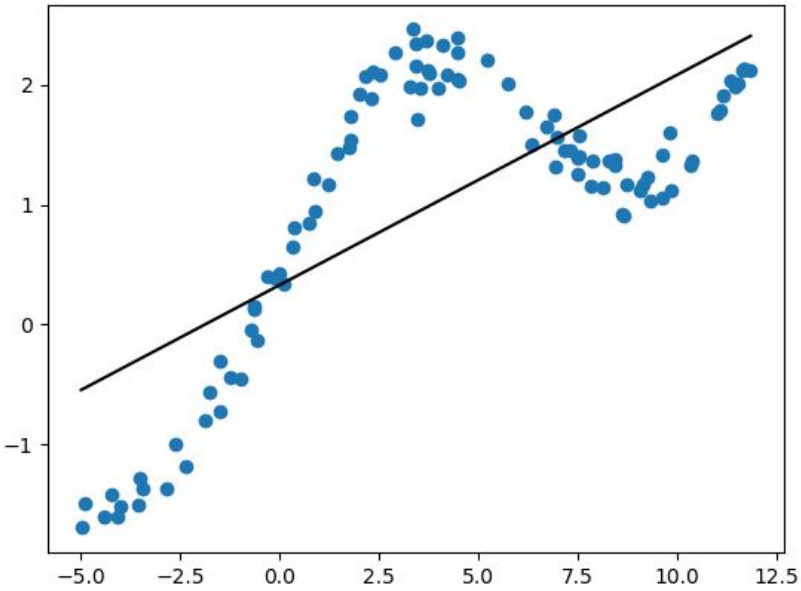
Tau=2



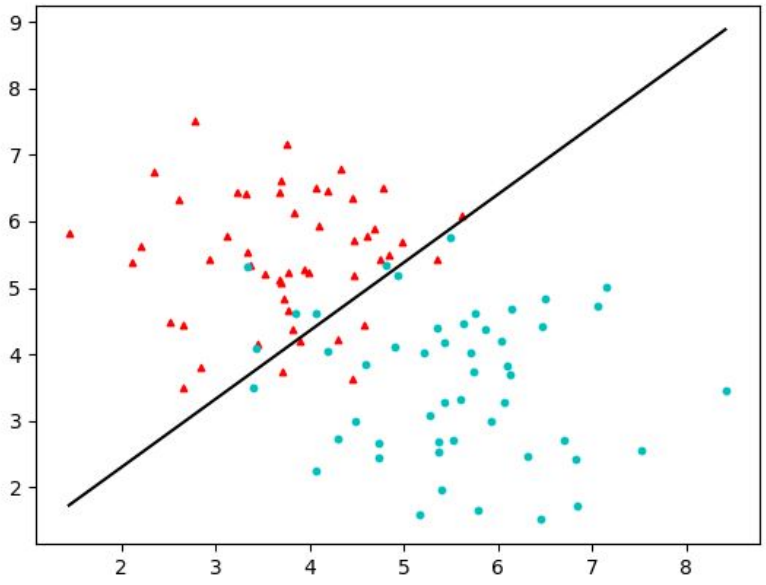
Tau=10



Unweighted

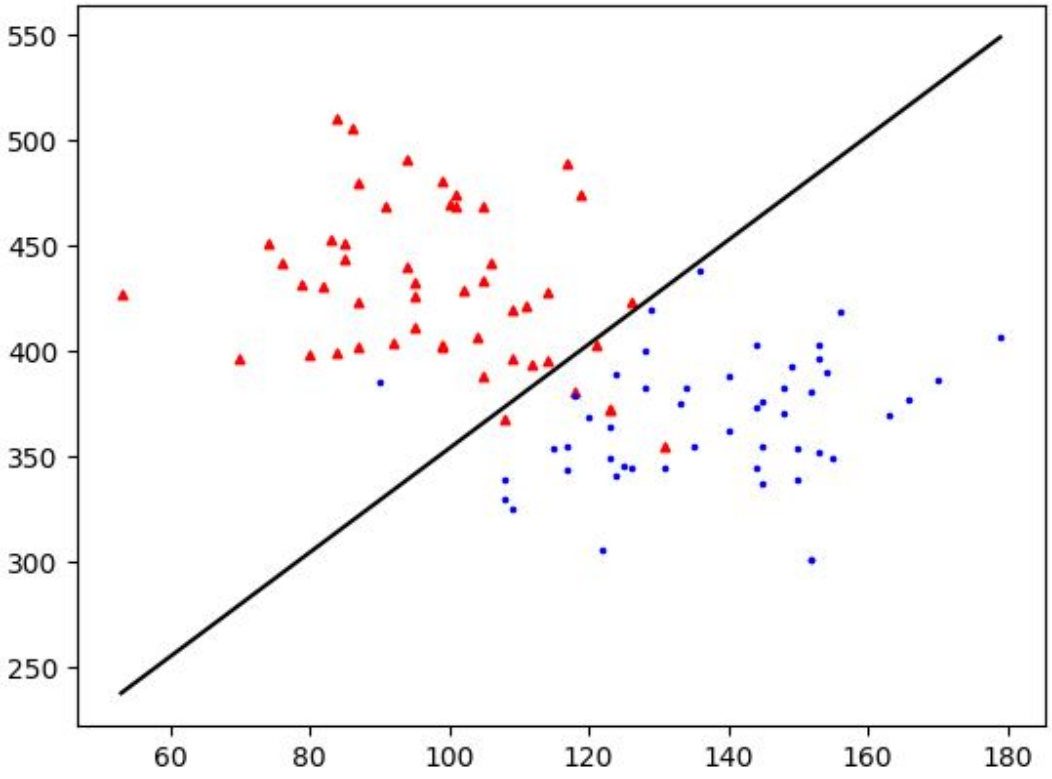


Answer 3:
[[-0.17591389]
 [-0.70462465]
 [0.68742883]]



Answer 4:

0.5
mu0=[[98.38 429.66]]
mu1=[[137.46 366.62]]
sigma=[[287.482 -26.748]
 [-26.748 1123.25]]



phi=0.5
mu0=[[98.38 429.66]]
mu1=[[137.46 366.62]]

sigma0=[[319.5684 130.8348]
 [130.8348 875.3956]]
sigma1=[[287.482 -26.748]
 [-26.748 1123.25]]