

## Assignment

### Question 1:

- A) The value after running 1 iteration of gradient descent is 1.1718750000000004 which is quite close to the optimal value 0;
- B) We see that on reducing the value of beta the has increased by some amount this is because by doing so we are performing a very crude/basic search which is not offering us a that good step size value however the backtracking parameters don't affect the f-value with much difference while approaching to convergence
- C) We notice that that f-value we got from newton descent is nearly double the f-value of gradient descent with same parameters also on comparing the two we find gradient was sufficient in gradient descent but in this we also need to calculate hessian then its inverse in every iteration which is much more than the gradient descent algorithm but this comes to be beneficial as it finishes much faster when running complete loop.
- D) For both algorithms parameters used are:
  - Alpha=0.1
  - Beta=0.7
  - Epsilon=0.001
  - No. of iterations of gradient descent algorithm: 11911
  - No. of iterations of Newton descent algorithm: 6
  - Newton descent finishes much faster

### Question 2:

#### Dual variables defined in .iypnb