Instructions: Make sure your questions and answers are on different pages. Do not include your name or any other identifying information. I will know that information from Canvas.

Question 1: Why is Newton's method more computationally expensive than Gradient Descent?

Question 2: What is the advantage of using gradient descent on convex objectives?

Question 3: What does the condition number convey?

Answer Question 1: Newton's method must calculate a matrix of second partial derivatives, where gradient descent only considers the first partial derivatives.

Answer Question 2: It's a simple case to find global optimum.

Answer Question 3: The difficulty / number of iterations needed to get close to the optimum of a problem.