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: What conditions need to be true for gradient descent to decrease the value of the objective at each iteration?

Question 2: What does the condition number do to gradient descent on strongly convex objectives?

Question 3: What does it mean for a function to be convex and why are convex functions significant to gradient descent?

Answer Question 1: the learning variable has to be small enough and the gradient has to be non-zero.

Answer Question 2: the condition number (L/μ) encodes the difficulty level of solving a convex problem. The larger the condition number the, the more difficult it is to solve, meaning that multiple iterations may be required to get close to the optimum.

Answer Question 3: It means that for a any line drawn between two points in the graph of the function, this line would be above the function, but more importantly it only has a global minimum. This is significant for gradient descent, because the point of gradient descent is ideally converging to the global minimum.