

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 is the difference between “empirical risk” and “loss”?

Question 2: How is Newton’s method different from gradient descent?

Question 3: Can we always assume that a loss function is convex?

Answer Question 1: Empirical risk is the loss function divided by number of data points.

Answer Question 2: Newton's method uses the second derivative

Answer Question 3: No, but the math does work out well if the loss function is convex.