ECE 490 – Spring 2022
Midterm
Instructor: Bin Hu

Name: ______________________________

UIN: ______________________________

Instructions:

- Please carefully read all questions.
- For maximum credit - attempt all questions and carefully elucidate your reasoning and answers.
- You can upload your solutions in one of three ways:
  1. You can download this PDF and write on it using a tablet (in MS OneNote for example).
  2. You can print this exam and write in the space provided by hand.
  3. You can write the solutions on your own worksheets/pages provided each question starts on a new page.
- The upload link to Gradescope will close at 12:30 p.m. C.D.T on 10th March. Make sure you allot time to scan/save and upload.
- Good luck!

1. Consider the quadratic function.

$$f(x_1, x_2) = x_1^2 + 3x_1x_2 + x_2^2 + 5x_1 - 10x_2 + 4$$

   (a) (10 points) Show that (x_1, x_2) can be chosen to make $f(x_1, x_2)$ approach ∞ and $-\infty$.
   Hint: Think of choose $x_1 = x_2 = t$ and $x_1 = -x_2 = t$.

(b) (10 points) Find (Q, p, r) to write $f(x_1, x_2)$ in the form of $f(x) = f(x_1, x_2) = \frac{1}{2}x^T Qx + p^T x + r$. Here, Q is required to be symmetric. Is f a convex function?

(c) (10 points) Find the stationary point for f. Prove that this stationary point is neither a local min or a local max, but a saddle point.

2. (a) (10 points) Consider the problem of minimizing the function of two variables $f(x, y) = 3x^2 + y^4$. Calculate one iteration of the steepest descent method with $(1, 2)$ as the starting point and with the stepsize chosen by the Armijo rule with $\alpha_0 = 1$, $\sigma = 0.05$, and $\beta = 0.25$.

(b) (10 points) Consider a simple two-layer neural network problem on a single data point. Given scalars x and y, we want to minimize $f(w^{(1)}, w^{(2)}) = (y - w^{(2)}\sigma(w^{(1)}x))^2$. What is the gradient of this function?

3. Consider Newton's method $x_{k+1} = x_k - \alpha(\nabla^2 f(x_k))^{-1}\nabla f(x_k)$, where α is the stepsize.

   (a) (10 points) Suppose we apply Newton's method to the function $f(x) = x^6$. Identify the range of α for which the method converges. Is the convergence linear?

(b) (10 points) Consider the use of Newton method to minimize $f(x) = -\cos(x)$. Choose the stepsize to be 1. Write x_1 and x_2 as a function of the initial point x_0. Show that there exists $x_0 \in [\frac{\pi}{4}, \frac{\pi}{2})$ such that $x_2 = x_0$. For this x_0, does Newton's method converge?

4. (30 points) (True or False). State whether each of the statements is true or false. Briefly explain your choice. No points will be awarded without an explanation.

   (a) The function $f(x) = x^6$ is strongly convex.

(b) For a smooth strongly convex function, the gradient descent method with well chosen stepsize always converges. The larger the condition number is, the slower the convergence is.

(c) The gradient descent algorithm with an exact line search always finds the minimum of a strictly convex quadratic function in exactly one iteration.

(d) Let C_1 and C_2 be two convex subsets of $\mathbb{R}^n$. Then, $C_3 = \{x - y : x \in C_1, y \in C_2\}$ is always a convex set.

(e) If $\nabla f(x) = 0$, then x must either be a local maximizer or a local minimizer of f.

(f) A non-convex real-valued function always admits a local minimum that is not a global minimum.