HW #5 Due: 5/17/2023

1. For the problem of finding the fastest way to the island on pp 2 of 19_Gradient_descent_basis, (a) formulate it as a constrained optimization problem with variables $x$ and $y$. (b) Analytically solve this problem by using the Lagrange multiplier method.
2. Follow the treatment of the L2 regularization (pp. 45 – 48) to find out the influence of the L1 regularization on the updating of weights.
3. We mention in the lecture about how to use a penalty function to convert a constrained optimization problem to an unconstrained one. However, the obtained solution is not exactly the same as that from the original problem. The following illustrates this situation. (a) Analytically find the values of $x$ and $y$ to maximize the cost function

$$J(x, y) = x + y - 100(x^2 + y^2 - 1)^2.$$

(b) Is your answer the same as the solution of the original problem (the example in the lecture notes)?

(c) How to change $J(\mathrm{x}, \mathrm{y})$ if we want to find the minimum of the function

$$f(x, y) = x + y \text{ subject to } x^2 + y^2 = 1?$$

4. Repeat problem 3 (a), but use the gradient **ascent** algorithm to find the solution. To have a unique answer, use $x_0 = y_0 = 1$ and $\eta = 0.0005$. (a) Compute the solution with 1,000 iterations and print out the final values of ($x$, $y$). (b) Is your numerical solution close to the analytic solution in problem 3? (c) If we increase the penalty constant from 100 to 10,000, what side effect(s) may result in your program?
5. Use the LDA to reduce the feature dimension from 4 to 2 for the Iris data set. As usual, take 70% of the samples as the training set to perform dimensionality reduction. Use 5-NN to classify the test set and then report the average accuracy after 10 trials.