Supplement to problem 5 of Homework 5

This problem is not correct. Notice that the loss function

$$J(x, y) = x + y - 100(x^2 + y^2 - 1)^2 \tag{1}$$

can be a very large negative number (unbounded!) as $(x^2 + y^2 - 1)^2 \geq 0$. Therefore, for such a loss function, we can only find the **maximum** value of the loss. Therefore, we have to use the gradient **ascent** algorithm to find argmax $J(x, y)$. By using $\eta = 0.0005$, we have the solution as

$$x = y = 0.70835$$

It is also possible to analytically solve this problem to verify whether the computed solution is correct or not. The candidates of extreme points are

$$\frac{\partial f}{\partial x} = 1 - 400x(x^2 + y^2 - 1) = 0 \tag{2}$$

$$\frac{\partial f}{\partial y} = 1 - 400y(x^2 + y^2 - 1) = 0 \tag{3}$$

With some algebraic work, we have

$$(x - y)(x^2 + y^2 - 1) = 0$$

Therefore,

$$x = y \ \textbf{or} \ (x^2 + y^2 - 1) = 0 \ \text{(or both)}$$

If we let $x = y$ and $x^2 + y^2 = 1$, we can find $x = y = \frac{\sqrt{2}}{2}$. These values are not answers, because they cannot satisfy Eq. (2) or (3). So, only one of

$$x = y \ \text{or} \ x^2 + y^2 = 1$$

is true. It is easy to find answers if we use $x = y$. Putting $x = y$ into Eq. (2) and finding the polynomial roots numerically, we have

$$x = -0.0025, -0.7059, 0.70835$$

The third root is exactly the answer we have by using the gradient ascent method. Recall the analytic solution to the original problem is $x = y = \frac{\sqrt{2}}{2} \approx 0.707$. The penalty term introduces a small error when compared with the exact solution. We can increase the constant in the penalty from 100 to, say, 1000 to reduce the error. However, there is a shortcoming if increasing the penalty constant.

Question 1: What is the shortcoming to use a very large penalty constant?

Question 2: We know only one of $x = y$ or $x^2 + y^2 = 1$ is true. But why not using $x^2 + y^2 = 1$ to compute the optimal solution? (Hint: $J(x, y)$ will be unbounded)

Question 3: How to modify the loss function if we want to find the minimum value of $(x + y)$ subject to $x^2 + y^2 = 1$?