HW4

CISC648010 - Spring 2022

Due Date: March 11 at 11 PM

1 Gradient Descent (5 pts each part)

a) Consider function $J(\theta)$

$$J(\theta) = \frac{1}{2}\theta^T A\theta$$
$$\text{where } \theta = \begin{bmatrix}\theta^{(1)}\\ \theta^{(2)}\end{bmatrix} \in \mathbb{R}^2$$
$$\text{and } A = \begin{bmatrix}2 & 0\\ 0 & 3\end{bmatrix}$$

Let θ_t be the value of θ at iteration t of gradient descent. Assume that learning rate is $\alpha = \frac{1}{5}$. Write the gradient descent update in form of $\theta_{t+1} \leftarrow g(\theta_t)$ (you should find g).

b) If $\theta_0 = \begin{bmatrix}1\\ 1\end{bmatrix}$, find value of θ at iteration 2 and iteration 10.

c) What is $\lim_{t\to\infty} \theta_t$?

d) Repeat part (a) with learning rate 1.

e) Assume that $\theta_0 = \begin{bmatrix}1\\ 1\end{bmatrix}$. Does $\lim_{t\to\infty} \theta_t$ exist in this case?
If yes, find it.
If no, explain why gradient descent does not converge.

2 Handwritten Image Classification 20 pts

Download the file mnist_49_3000.mat from Canvas. This is a subset of the MNIST handwritten digit database, which is a well-known benchmark

database for classification algorithms. This subset contains examples of the digits 4 and 9. The data file contains variables x and y, with the former containing patterns and the latter labels. The images are stored as vectors. To load the data use the following code:

```
import scipy.io
import numpy as np
data = scipy.io.loadmat('mnist_49_3000.mat')
x = np.array(data['x'])
y = np.array(data['y'][0])
y[y==-1] = 0
```

To visualize an image, type the followings:

```
from matplotlib import pyplot as plt
index = 0 #change the index to show different images
image = x[:,index].reshape(28,28)
plt.imshow(image, interpolation='nearest')
plt.show()
```

Implement Newton's method (a.k.a. Newton-Raphson) to find a minimizer of the regularized logistic regression. Try setting $\lambda = 10$. Use the first 2000 examples as training data, and the last 1000 as test data. Please report the following:

a) (5 points) The test error

b) (2 points) Your termination criterion (multiple options here)

c) (3 points) The value of the objective function at the optimum

d) (5 points) In addition, generate a plot with 5 images. These 5 images should be the 5 misclassified images for which the logistic regression classifier was most confident about its prediction (you will have to define a notion of confidence in a reasonable way – explain what this is). In the title of each subplot, indicate the true label of the image. What you should expect to see is a bunch of 4s that look kind of like 9s and 9s that look like kind of like 4s. Include the plot in your report

e) (5 points) To receive credit for this problem, please submit your code via Canvas, in a single file named prob3_lastname.py