Untitled5

May 28, 2019

Assignment09
Assignment08: Build a binary classifier to classify digit 0 against all the other digits at MNIST dataset
Software Engineering
20154652 Lee Dong Jae

In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
train = pd.read_csv("mnist_train.csv")
```

In [3]:
```
test = pd.read_csv("mnist_test.csv")
```

In [4]:
```
train = np.array(train)
test = np.array(test)
```

In [5]:
```
tr_ans = train.T[0]
ts_ans = test.T[0]
```

In [6]:
```
tr_data = train.T[1:].T
ts_data = test.T[1:].T
```

In [7]:
```
#get traing_set samples - feature matrix
A_tr = np.zeros((len(tr_data), 1+(28*28)))
for i in range(1+(28*28)):
    for j in range(len(tr_data)):
        if i == 0:
            A_tr[j][i] = 1
        else:
            A_tr[j][i] = tr_data[j][i-1]
```

In [8]:
```
#make a label
Y_tr = np.zeros(len(tr_ans))
for i in range(len(tr_ans)):
    if tr_ans[i] == 0:
        Y_tr[i] = 1
    elif tr_ans[i] != 0:
        Y_tr[i] = -1
```

1. Compute an optimal model parameter using the training dataset

```
In [9]: #do a QR decomposition
        q, r = np.linalg.qr(A_tr)
        r_inverse = np.linalg.pinv(r)
        temp_z = np.dot(r_inverse, q.T)

        #obtain co-efficients
        z = np.dot(temp_z, Y_tr)
        print("Optimal model parameter: ")
        print(z)
        print('\n\n')
```

```
Optimal model parameter:
[-6.84400968e-01  1.11701697e-13 -7.61260730e-14 -2.93018364e-16
  1.27180086e-14  2.88731070e-14 -8.98414109e-15 -2.19830054e-14
 -1.16844612e-14  8.56132789e-15 -1.28623547e-14 -2.89889739e-15
 -1.00479639e-14  6.04224036e-04  7.77694995e-04 -6.06867193e-04
 -2.52861331e-05 -8.43537696e-15  7.80738467e-15 -3.99570850e-15
  2.48076499e-15 -1.56839138e-15 -6.84229080e-15  3.76632641e-15
 -5.38203321e-15  1.79404603e-15 -4.67698387e-15  1.90358025e-15
 -5.60909480e-15 -1.28111383e-14 -1.17340584e-14  6.25289137e-15
 -1.72928718e-14 -1.60085305e-02  4.14782420e-03  4.74742394e-04
  8.35280484e-04  2.50300414e-04  1.04852791e-05  2.15417605e-04
  4.56768356e-04 -2.85486946e-04  1.07011808e-03 -6.39508565e-04
  2.42213332e-04  7.47703281e-07  3.02386386e-04 -2.43053099e-04
  6.00939168e-05  3.35408731e-04 -5.56090571e-04  9.32911561e-04
  1.42636538e-04 -3.05186672e-15  8.95025199e-15  1.24070150e-14
 -1.12841198e-16  4.87054571e-16  3.05324336e-15 -3.27712236e-02
  3.83099672e-02  2.09577326e-03 -3.13999704e-03 -7.42404557e-04
  3.95818084e-04 -2.06938646e-04 -3.20870801e-04 -2.93432463e-04
  1.96949608e-04 -1.93353307e-04 -5.59288703e-05 -1.33128304e-04
  1.43402841e-05 -4.34807292e-04 -1.01570997e-04 -1.71799070e-04
 -1.84830730e-04 -3.87192748e-04  3.97266627e-04 -6.62298079e-04
  4.05916256e-04 -9.63828148e-04  1.60516686e-03 -1.82478446e-15
  3.75162484e-15  1.46062479e-14 -8.76204224e-15  1.55014298e-02
 -7.03335314e-03 -3.33841241e-04 -3.20255844e-04 -2.49182460e-04
 -8.01922055e-06  9.22067161e-05 -2.43918635e-04 -1.32643674e-04
  2.14730022e-04 -2.33072418e-04  1.80956359e-04 -1.79764653e-04
 -1.20499943e-04 -6.89585049e-05 -1.60575701e-04 -4.05885161e-04
 -1.80111333e-04 -4.23902938e-04 -4.11906670e-04 -2.71575063e-04
 -2.63010061e-04  4.24163632e-04 -4.38257265e-04  8.96389621e-04
  5.18715485e-17  1.15161857e-14  2.42768369e-03 -1.12423326e-03
 -2.24680600e-04  2.86969942e-04 -4.64259754e-05 -8.05305596e-05
 -1.95398498e-04  1.46199376e-04 -1.70769036e-04 -8.78205792e-05
 -1.10984060e-04  6.53701041e-06  1.20338032e-04 -1.36623153e-04
  1.53343975e-04 -6.51752230e-05  1.48555860e-05 -9.54927169e-06
  1.01073752e-04 -2.71232557e-05  1.08854588e-04 -9.68759148e-05
```

```
-2.60142099e-04 -2.52258868e-04 -1.40885721e-04 -9.52449235e-04
 5.20580291e-03  3.25018296e-15 -3.66924116e-16 -3.60252921e-03
-2.95571736e-04 -5.47510570e-05 -1.58318533e-04  1.51122350e-04
-5.96120105e-05 -1.15184630e-04 -9.61900776e-05 -7.71163346e-05
 3.88636461e-06 -8.81553516e-05  1.40337290e-04  4.19722537e-07
 7.59965538e-05  3.84584292e-05  5.53567663e-05 -5.76252157e-05
 7.40889084e-05 -9.97501487e-05 -7.13934914e-05 -2.00047191e-04
-2.46008264e-04 -9.34831208e-05 -1.12335896e-04 -4.38311381e-04
-3.51463671e-03  4.62646368e-15 -1.00197161e-02  6.95105265e-04
 2.61982777e-05  7.42336304e-05 -1.13436098e-04  5.79293538e-05
-6.56200696e-05  7.29116197e-05 -1.05803139e-04 -8.68666663e-05
-4.16012001e-05 -6.91525054e-06  8.97830282e-05  1.15840325e-05
-1.95883955e-05  4.71497816e-05 -4.44262214e-06  5.90680442e-05
-5.43022455e-05  1.11516545e-05  7.53528288e-05  2.30177229e-05
-1.50160171e-04 -4.07623681e-04 -1.23745233e-04 -2.99174876e-04
-3.38756010e-04 -1.18920320e-01 -2.16168914e-04 -1.22840757e-03
 1.61827031e-04 -5.06820763e-05  5.88959629e-05 -3.07658542e-05
 1.71980819e-05 -1.01102783e-04  1.40747444e-04 -6.22014243e-05
 5.58456094e-05  5.65817677e-05  5.38613813e-05  1.44731875e-04
 1.25967991e-04  7.09691459e-06  1.50089467e-04  8.41693544e-05
 7.41706237e-05 -8.23604266e-05 -2.76992820e-05  1.42835510e-05
-6.63554463e-05 -4.28539211e-04 -1.84046661e-04 -1.32330941e-04
 4.67837221e-04  3.03440947e-02  3.15405904e-04  1.25831293e-04
-1.49064645e-04 -1.80269456e-05 -4.24415016e-05  1.48385753e-04
-4.04608377e-05  1.95154327e-05 -1.81566276e-05  4.81381598e-06
-5.06451269e-05  5.91288182e-05  1.61917763e-05  7.84972947e-05
 7.11794355e-05  2.04697949e-04  5.53001204e-05  6.24711420e-05
 3.93525035e-06  1.22539188e-04 -1.06067150e-04  4.85332682e-05
-3.65148485e-06 -4.81393599e-04 -4.97089004e-04 -1.04321196e-04
-2.94270824e-04 -2.47585939e-03  4.97396634e-04 -2.08510445e-04
 1.75184899e-05 -6.74425128e-05 -6.87211433e-05 -9.01380423e-06
 8.86386309e-05  4.59166397e-05  1.98105409e-05  4.28698903e-05
-7.02707207e-05 -2.50049534e-05  1.14767945e-04  1.01212849e-04
 1.28827537e-04  1.65800931e-05  4.62182846e-05  3.43164446e-05
 1.62801679e-04  4.13981901e-05  8.23636646e-05 -1.37752193e-04
 1.04846921e-04 -1.30484697e-04 -8.62622596e-04  1.41033723e-05
-2.91145590e-04  1.65298731e-03 -8.57831323e-04 -2.36513851e-05
-8.30771926e-05  1.72219022e-05  8.82497605e-05 -4.86922904e-06
-5.35242473e-05 -6.40568082e-05  1.12133472e-04 -1.56203600e-05
 4.30349525e-05  6.80255392e-05  3.54863988e-05  3.49831090e-05
 1.19530520e-04  7.49047535e-05  1.41341604e-04  1.34573336e-04
 3.09750869e-05  2.61811876e-04  1.29486711e-04  1.70250637e-04
 2.94936328e-04 -2.99190840e-04 -8.75858681e-04 -6.61370154e-04
 1.52598188e-03 -1.49421271e-03  1.16892706e-03 -2.20319309e-04
-1.34414244e-04  1.09688268e-05 -2.73439561e-04  3.76319506e-06
 7.58984966e-05  4.02621445e-05 -5.06868900e-05  9.23189330e-05
 5.72046648e-05 -8.06084254e-05 -5.61842781e-06 -1.66146266e-04
-1.00146108e-04 -8.18143910e-05  2.87507684e-06  3.86427988e-05
```

```
 1.63935092e-04  7.86168642e-05  2.52021227e-04  5.80783180e-05
 2.77174293e-04  4.20337560e-04 -1.33797169e-03  5.73508791e-05
-2.63249869e-03  5.22823534e-04 -9.31908228e-04 -2.94184669e-04
 8.96450071e-05 -2.94283459e-04 -3.92191296e-05  7.30361408e-05
-6.05908748e-05  7.75072940e-05  3.76887827e-05  6.86988257e-05
-9.09451853e-05  5.80188873e-05 -4.31157285e-05 -3.00175448e-04
-1.13787034e-04 -1.67762068e-04 -1.31629115e-04 -1.46034159e-04
-1.12843838e-04  1.04948175e-04  6.40749343e-05  1.10824464e-04
 5.04151852e-04  3.27244848e-04 -1.47701901e-04 -1.28881507e-03
 2.76732710e-03  1.42919072e-03  2.79457726e-03  9.58726054e-05
-8.15918764e-06 -5.02735103e-05 -4.23085169e-05 -8.17561096e-05
 7.65539214e-05 -8.73616936e-05  6.03146204e-05  3.95411281e-05
 3.08176670e-05 -1.10737042e-04 -1.62585792e-04 -2.03913957e-04
-1.03098645e-04 -9.02739304e-05 -1.90641840e-04 -1.95876678e-04
-9.08181500e-05 -1.78189484e-04  7.18457045e-05  3.46135725e-04
 2.82068282e-04 -3.77451389e-04  3.42668598e-07  4.67355979e-04
-1.61637586e-03 -1.31566197e-03 -1.71279758e-03 -2.43564277e-04
-2.36119754e-04 -5.44568976e-04  8.20173168e-05  1.54090764e-04
 9.46606026e-05  5.68806834e-05  8.06169511e-05  1.96579864e-04
 4.19469862e-07 -1.68776582e-04  2.00127806e-05 -2.73220169e-04
 2.19178110e-05 -1.56780495e-04 -1.68270089e-04 -1.15873811e-04
-1.50115446e-04  7.27922041e-05  3.15485065e-04  1.22475089e-06
 9.32309824e-05  1.15641585e-04 -2.04919582e-04 -2.33507328e-04
 2.59871465e-03 -1.68351354e-03  9.87502009e-04  6.87826683e-05
-1.22618501e-04  1.90863724e-04  2.23758553e-04 -1.43977368e-04
 1.37972115e-04  1.29990556e-04  1.46584676e-04  1.24328320e-04
-5.64428937e-05 -1.26804058e-04 -8.67367083e-05 -2.26861905e-04
-3.16756424e-05 -1.94500764e-04 -1.57340358e-04 -1.16335873e-04
-5.59351942e-05  1.24425591e-04 -1.96621324e-05 -5.35631793e-06
 1.00841830e-04 -9.97006425e-05 -2.01614342e-04 -1.22991990e-04
 5.81662602e-04 -7.15840898e-03  2.04802383e-04  1.17847543e-03
-9.27632567e-04  2.94532149e-05  3.93339441e-05  1.15420910e-04
 2.71381877e-05  1.06860155e-04  3.75278268e-05  8.28189439e-05
-3.99871108e-05 -2.20403500e-04 -1.44869141e-04 -1.38796518e-04
-1.17608415e-04 -1.60443229e-04 -1.75902475e-04 -2.18275733e-05
 6.05697256e-05  5.42856676e-05  4.22279621e-05 -2.14727287e-04
 9.65605914e-05  9.22781119e-05 -4.95289345e-05 -2.50394320e-04
-1.66821002e-03  2.39528143e-15 -5.04346677e-04 -7.33845144e-04
-1.07810242e-04  6.57957793e-05  9.51421260e-05  1.11093653e-04
 9.76588849e-05 -1.64717860e-05  5.98027438e-05  6.62151252e-05
-5.90031511e-05 -5.67031163e-05 -1.73888496e-04 -2.19388950e-04
-1.88120445e-04 -7.16595782e-05  4.29807702e-05  3.48074445e-06
 8.54289328e-05 -5.64099603e-05 -6.33233601e-05  2.12415311e-04
-9.48163312e-05  1.98201620e-05 -7.09775668e-05  6.46907170e-04
 6.46969483e-03 -8.58548561e-04 -5.25377709e-04  9.97545878e-04
-9.27280894e-04 -1.88372248e-05  1.34145573e-04 -3.29963028e-05
 1.59347901e-04  1.30806130e-04  1.56817583e-05 -3.06996729e-05
 3.34488307e-05 -1.04053404e-04 -7.18694769e-05 -1.20203916e-04
```

```
 1.13407702e-05 -5.10979830e-05 -4.44284438e-05 -5.62806426e-05
-4.83714606e-06 -1.08260158e-04 -2.07217381e-05 -2.62884626e-05
-2.78190906e-04  2.54214670e-04 -1.36038744e-04 -7.52903654e-04
-4.84638714e-03 -4.40620623e-03  9.52348971e-04 -3.58789775e-04
 4.93293107e-05 -2.51910168e-05 -1.06511209e-04  6.07861929e-05
 9.78136797e-05  4.82248783e-05  1.35553815e-04 -2.06118107e-05
 7.08122775e-05  5.17568227e-05  4.54970782e-05  9.38555346e-06
-4.68941427e-05 -1.44966496e-05  8.83951919e-06  2.29367030e-07
-5.19683505e-05  2.81666015e-05 -3.40512115e-05  3.14146343e-06
 5.03184637e-05 -7.12562684e-05 -2.07529234e-04  1.60737607e-04
 3.37020586e-03 -7.59920831e-16  3.88346555e-03  3.96949753e-04
-4.03311059e-04 -1.08071680e-04  6.71651507e-05 -6.26045264e-05
-1.41853925e-06  4.37424820e-05  6.44177017e-06  8.20815536e-05
 2.26156715e-06  8.58978296e-05  4.96036314e-05  5.49793468e-07
 2.61072403e-05 -1.01045973e-04  3.26762355e-05 -7.70514284e-05
 6.71248957e-06 -7.31179885e-05  1.26249338e-04 -7.80303990e-05
-2.26830299e-04  3.24377345e-05  3.83561206e-04 -4.27956297e-04
-1.19972531e-03  1.40936459e-03 -4.07060569e-03 -1.21058478e-04
-2.55441040e-05  2.21052283e-04 -1.00348744e-04 -3.06664657e-05
-3.86937165e-05 -3.64594610e-05  1.07680326e-04  6.69128206e-05
 8.33113159e-05 -3.13460775e-06 -4.91934206e-05  1.25904097e-05
 1.05336069e-04  2.25596613e-05 -9.25331847e-05  2.47086465e-05
-3.36176639e-05 -1.50065451e-05 -6.34616937e-05  4.65247137e-05
-1.71135202e-04  8.62713194e-05 -4.25044704e-05  1.02725110e-03
 2.24941488e-03  1.36532194e-03 -4.02793010e-03 -6.86125350e-05
 4.05502980e-05 -4.55582690e-04  9.97235916e-05  1.27351040e-04
 3.98203844e-05  2.17996891e-05  6.69196363e-05  1.09480808e-05
 7.57646867e-05  5.80004625e-05  1.65486050e-04  4.27098596e-05
-6.01286811e-05 -2.36237299e-05 -2.26474499e-05 -1.37505989e-04
 6.41721804e-05 -1.90764302e-04  7.98497302e-05  5.15799603e-05
-5.83782529e-05  7.15119757e-05  4.01655342e-05  1.38868161e-03
 4.15276594e-03  6.57098115e-16 -8.16954708e-16  4.92274799e-04
-9.16471229e-05 -2.34028703e-04 -6.64174628e-06 -2.36486430e-04
 9.97523347e-05 -7.48830648e-05  5.15667758e-05  1.22980089e-05
 8.41765505e-05  1.00725927e-04  1.32503509e-05  1.18756935e-04
 1.48374738e-04  8.51896914e-05  8.33727927e-05 -7.02841395e-05
-8.67091657e-05 -4.36901508e-05 -2.07732613e-04  3.59789011e-05
 7.58593328e-06  6.39834097e-05 -5.74935495e-04 -2.39562431e-03
 3.31171376e-16  1.12048015e-16 -1.22122649e-16 -2.63414579e-03
 6.42010194e-04  1.00020026e-04 -3.45254330e-04  2.84157903e-04
-1.40978026e-04  2.00236676e-05 -1.07442169e-04 -1.88173542e-05
-1.79543949e-04 -5.46494266e-05 -7.07448290e-05 -1.24430903e-04
-5.40243808e-05 -1.58693031e-05  1.62290006e-05  1.52095747e-04
-3.50758450e-05  1.74867130e-04  1.25070051e-05 -3.97461921e-04
 1.06202276e-04 -5.52916113e-04  9.74012420e-04  1.74317765e-02
-6.69918294e-17 -1.44258354e-17 -2.70999936e-17  2.26068414e-03
 2.87582359e-04 -1.29841970e-05 -5.36913857e-04 -2.16377266e-04
-4.13091862e-04 -2.16774719e-04 -1.88425965e-04 -2.61355929e-04
```

```
-1.05784707e-04 -2.54150975e-04 -1.20888140e-04 -1.72924144e-04
-4.30644297e-05 -1.18862030e-04 -1.34521255e-04 -8.61849195e-05
 4.39551326e-05 -2.82225464e-04  3.20356520e-05  5.59236443e-04
-6.65973371e-04 -7.14285735e-04 -1.94892780e-03 -1.66504895e-02
 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
-3.63105735e-03  1.02883061e-03 -8.08323979e-04  2.99705351e-05
-3.59593664e-04 -1.85698914e-04 -4.57183311e-04 -1.69009710e-04
-3.97441951e-04 -1.32280398e-04 -3.53758907e-04 -1.83737820e-04
-3.83355388e-04 -2.12310032e-04 -1.28365261e-04 -2.92094401e-04
-2.96822599e-05  3.35314208e-05 -6.11071596e-04 -2.49705875e-04
-1.03381522e-03  8.08493376e-03 -1.07060000e-03  0.00000000e+00
 0.00000000e+00  0.00000000e+00  0.00000000e+00  0.00000000e+00
 0.00000000e+00  3.13085578e-03  9.90099899e-05 -1.48174445e-03
 2.36696946e-04 -5.05756538e-04 -4.44842004e-04 -1.36379294e-04
-7.04668796e-05 -5.33133241e-04 -3.00925463e-05 -5.32371122e-04
-9.62959848e-05 -1.69472960e-04 -4.51187202e-04 -1.75872357e-04
 4.26376949e-05 -4.03413496e-04  7.23514264e-04 -1.37287972e-03
-8.17930980e-03  0.00000000e+00  0.00000000e+00  0.00000000e+00
 0.00000000e+00]
```

In [10]:
```
#get test_set samples - feature matrix
A_ts = np.zeros((len(ts_data), 1+(28*28)))
for i in range(1+(28*28)):
    for j in range(len(ts_data)):
        if i == 0:
            A_ts[j][i] = 1
        else:
            A_ts[j][i] = ts_data[j][i-1]
```

In [11]:
```
A_mul_x_ts = np.dot(A_ts, z)
A_mul_x_tr = np.dot(A_tr, z)
Y_ts = np.zeros(len(ts_ans))
label_tr = np.zeros(len(tr_ans))
```

In [12]:
```
for i in range(len(ts_data)):
    if A_mul_x_ts[i] > 0:
        Y_ts[i] = 0
    else:
        Y_ts[i] = 1

for i in range(len(tr_data)):
    if A_mul_x_tr[i] > 0:
        label_tr[i] = 0
    else:
```

```
            label_tr[i] = 1

    for i in range(len(ts_ans)):
        if ts_ans[i] != 0:
            ts_ans[i] = 1

    for i in range(len(tr_ans)):
        if tr_ans[i] != 0:
            tr_ans[i] = 1
```

```
In [13]: #calculate each scores
         def get_score(actual, pred):
             TP = 0
             FP = 0
             TN = 0
             FN = 0

             for i in range(len(pred)):
                 if actual[i]==pred[i]==0:
                    TP += 1
                 if pred[i]==0 and actual[i]!=pred[i]:
                    FP += 1
                 if actual[i]==pred[i]==1:
                    TN += 1
                 if pred[i]==1 and actual[i]!=pred[i]:
                    FN += 1

             return(TP, FP, TN, FN)
```

```
In [14]: TP_tr, FP_tr, TN_tr, FN_tr =  get_score(tr_ans, label_tr)
         TP_ts, FP_ts, TN_ts, FN_ts =  get_score(ts_ans, Y_ts)

         TP_ratio_tr = TP_tr / (TP_tr + FN_tr)
         FP_ratio_tr = FP_tr / (FP_tr + TN_tr)
         TN_ratio_tr = TN_tr / (TN_tr + FP_tr)
         FN_ratio_tr = FN_tr / (FN_tr + TP_tr)


         TP_ratio_ts = TP_ts / (TP_ts + FN_ts)
         FP_ratio_ts = FP_ts / (FP_ts + TN_ts)
         TN_ratio_ts = TN_ts / (TN_ts + FP_ts)
         FN_ratio_ts = FN_ts / (FN_ts + TP_ts)
```

2. Compute (1) True Positive, (2) False Positive, (3) True Negative, (4) False Negative based on the computed optimal model parameter using (1) training dataset and (2) testing dataset.

```
In [15]: print("Train Set: ")
         print("True Positive Ratio: ", TP_ratio_tr)
         print("False Positive Ratio: ", FP_ratio_tr)
```

```
        print("True Negative Ratio: ", TN_ratio_tr)
        print("False Negative Ratio: ", FN_ratio_tr)
        print("\n\n")
        print("Train Set: ")
        print("True Positive Ratio: ", TP_ratio_ts)
        print("False Positive Ratio: ", FP_ratio_ts)
        print("True Negative Ratio: ", TN_ratio_ts)
        print("False Negative Ratio: ", FN_ratio_ts)
```

```
Train Set:
True Positive Ratio:  0.8725308120884687
False Positive Ratio:  0.003310156076632887
True Negative Ratio:  0.9966898439233671
False Negative Ratio:  0.1274691879115313



Train Set:
True Positive Ratio:  0.8836734693877552
False Positive Ratio:  0.004767712606719148
True Negative Ratio:  0.9952322873932808
False Negative Ratio:  0.11632653061224489
```