Assignment -1 (10 Marks)

Q1.

Network tomography. A network consists of n links, labeled $1, \ldots, n$. A *path* through the network is a subset of the links. (The order of the links on a path does not matter here.) Each link has a (positive) *delay*, which is the time it takes to traverse it. We let d denote the n-vector that gives the link delays. The total travel time of a path is the sum of the delays of the links on the path. Our goal is to estimate the link delays (*i.e.*, the vector d), from a large number of (noisy) measurements of the travel times along different paths. This data is given to you as an $N \times n$ matrix P, where

$$P_{ij} = \begin{cases} 1 & \text{link } j \text{ is on path } i \\ 0 & \text{otherwise,} \end{cases}$$

and an N-vector t whose entries are the (noisy) travel times along the N paths. You can assume that $N > n$. You will choose your estimate $\hat{d}$ by minimizing the RMS deviation between the measured travel times (t) and the travel times predicted by the sum of the link delays. Explain how to do this, and give a matrix expression for $\hat{d}$. If your expression requires assumptions about the data P or t, state them explicitly.

Remark. This problem arises in several contexts. The network could be a computer network, and a path gives the sequence of communication links data packets traverse. The network could be a transportation system, with the links representing road segments.

3 Marks

Q2. *Polynomial classifier with one feature.* Generate 200 points $x^{(1)}, \ldots, x^{(200)}$, uniformly spaced in the interval $[-1, 1]$, and take

$$y^{(i)} = \begin{cases} +1 & -0.5 \leq x^{(i)} < 0.1 \text{ or } 0.5 \leq x^{(i)} \\ -1 & \text{otherwise} \end{cases}$$

for $i = 1, \ldots, 200$. Fit polynomial least squares classifiers of degrees $0, \ldots, 8$ to this training data set.

(a) Evaluate the error rate on the training data set. Does the error rate decrease when you increase the degree?

(b) For each degree, plot the polynomial $\tilde{f}(x)$ and the classifier $\hat{f}(x) = \mathbf{sign}(\tilde{f}(x))$.

(c) It is possible to classify this data set perfectly using a classifier $\hat{f}(x) = \mathbf{sign}(\tilde{f}(x))$ and a cubic polynomial

$$\tilde{f}(x) = c(x + 0.5)(x - 0.1)(x - 0.5),$$

for any positive c. Compare this classifier with the least squares classifier of degree 3 that you found and explain why there is a difference.

7 Marks

Additional Information: Relevant information for solving Q2 can be found in Section 14.1 and 14.2 in the textbook – "Introduction to Applied Linear Algebra" by S.Boyd.

Additional Instructions:

1- For Q1, you can upload a printed or a handwritten answer .

2. For Q2 code in python and use plots to showcase your results. You are allowed to use python packages in your code. Upload a report explaining your answers along with your python code for Q2.