AI 512 - Mathematics for Machine Learning
Assignment # 1

Name : Yugaan Mangesh Hagargi
Roll No : IMT2020536

Q1. *Network tomography.* A network consists of n links, labeled $1, \dots, n$. 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

Soln:

Matrix P is a tall matrix $\hookrightarrow P_{N \times n}$ $\quad$ ($\because$ $N > n$ assumption)

N- vector t : $t_{N \times 1}$

To estimate : $\hat{d}$
$= d_{n \times 1}$

Case I

$t_{N \times 1} \in C(P_{N \times n})$

Using RMS deviation we get exact value of $d_{n \times 1}$

Case II

$$t_{N\times 1} \notin C(P_{N\times m})$$

We get approximate solution

Row interpretation

$$\|(Pd-t)\|^2 = (\tilde{P}_1 d - t_1)^2 + \dots (\tilde{P}_N d - t_N)^2$$

$$\frac{\partial}{\partial d}\left(\|Pd-t\|^2\right) = 0$$

$$\frac{\partial}{\partial d}\left((Pd-t)^T(Pd-t)\right) = 0$$

$$\frac{\partial}{\partial d}\left(d^T P^T P d - d^T P^T t - t^T P d + t^T t\right) = 0$$

$$2(P^T P)d - 2P^T t = 0$$

$$(P^T P)d = P^T t$$

Assuming $P^T P$ is invertible,

$$\hat{d} = (P^T P)^{-1} P^T t \longrightarrow \text{best value}$$

$\therefore$ if case I, exact value of $P\hat{d} = t$
if case II, Least squares value of $P\hat{d} = t$.