Applied Artificial Intelligence Institute (A2I2)
Deakin University
February 24, 2025

Problem Notes

Long Van Tran

s224930257@deakin.edu.au

1 Problem Statement

In a conventional setting, we have a time-series dataset $\{\mathbf{x}_t\}_{t=1}^{T}$ describing patient records across T consecutive time steps (e.g., daily or weekly measurements). However, due to various practical reasons (such as data anonymization, incomplete data logs, etc.), the temporal ordering is lost. Instead of a temporally-labeled sequence $\big(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_T\big)$, we only have a collection of N unlabeled points $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N\} \subset \mathbf{R}^d$. Our goal is to recover a plausible temporal sequence from this unordered set.

To do this, we assume that certain constraints exist which a valid time-series ordering should hold, such as:

1.1. Trajectory Distance.

The total distance traversed by the sequence should be "small", reflecting the fact that patient measurements taken throughout a period of time often don't change too significantly. Thus, we aim to minimize the sum of distances between consecutive points:

$$\min_{\pi} \sum_{t=1}^{N-1} d\Big(\mathbf{x}_{\pi(t)}, \mathbf{x}_{\pi(t+1)}\Big),$$

where $d(\cdot,\cdot)$ is a distance function (e.g., the Euclidean norm $\|\cdot\|$ in $\mathbf{R}^d$).

1.2. Consecutive Step.

Also from the "patient measurements" intuition stated previously, any two consecutive measurements should not differ by more than a threshold ϵ:

$$d\Big(\mathbf{x}_{\pi(t)}, \mathbf{x}_{\pi(t+1)}\Big) \;\leq\; \epsilon \quad \text{for all } t \in \{1, \ldots, N-1\}.$$

This enforces that the reconstructed time series does not have abrupt, large jumps in the data.

1.3. Directional Consistency.
Another possible constraint to be used is the assumption that direction of changes should be fairly consistent. Even though the noise terms are independent, the drift terms make the relative directional changes between consecutive points not too significant. Thus, we can filter out less likely sequences that contain too abrupt changes in direction, such that:

$$\theta_{\mathbf{x}_t,\mathbf{x}_{t+1}} \geq \theta \quad \text{for all } t \in \{1, \ldots, N-1\}.$$

where $\theta_{\mathbf{x}_t,\mathbf{x}_{t+1}}$ is the angle between two consecutive vectors $\mathbf{x}_t$ and $\mathbf{x}_{t+1}$.

2 Current Experimental Results and Issues