

Problem Statement 4: Efficient and Fair Line Construction

Consider a set of n houses with coordinates given by latitude (x) and longitude (y), defined as a set $P = \{p_1, p_2, \dots, p_n\}$. The task is to construct a gas pipeline in a straight line that serves all of them, where a straight line $\ell = \{a, b\}$ is defined by a direction vector a of unit norm and a point b on the line.

Distance Definition

The distance from a point p to a line $\ell = \{a, b\}$ is given by:

$$\text{dist}(p, \ell) := \|(I - aa^T) \cdot (p - b)\|^2$$

Objective 1: Efficient Line

A line is efficient if it minimizes the following cost:

$$\sum_{i \in [n]} \|\text{dist}(p_i, \ell)\|$$

Design an algorithm that computes an efficient line or a line which is almost efficient, meaning the cost is a local minimum.

Objective 2: Fair Line

A line is fair if it minimizes the maximum distance to any house:

$$\min_{\ell} \max_{i \in [n]} \|\text{dist}(p_i, \ell)\|$$

Design an algorithm that computes a fair line or a line which is almost fair, meaning the above cost is a local minimum.

Objective 3: Multiple Efficient Lines

For a set of k lines $\mathbf{L} = \{\ell_1, \dots, \ell_k\}$, the set is efficient if it minimizes:

$$\sum_{i \in [n]} \min_{\ell \in \mathbf{L}} \|\text{dist}(p_i, \ell)\|$$

Design an algorithm that computes k efficient lines or a set of lines which is almost efficient, meaning the above cost is a local minimum.

Dataset

Use the latitude and longitude from the California dataset provided by ‘sklearn’ to solve these problems.

[California Housing Dataset](#)