IEOR 151 – Lecture 13
P-Median Problem

1 Problem Setup

Location planning involves specifying the physical position of facilities that provide demanded services. Examples of facilities include hospitals, restaurants, ambulances, retail and grocery stores, schools, and fire stations. There are a variety of different models to solve this problem, and we will first consider a specific type of a discrete location model. These are models in which the facilities and demands are in discrete positions.

The p-median problem is a specific type of a discrete location model. In this model, we wish to place p facilities to minimize the (demand-weighted) average distance between a demand node and the location in which a facility was placed. This serves as an approximation to total delivery cost. In this model, there are no capacity constraints at the facilities.

1.1 Mathematical Model

We represent the problem formulation using an undirected graph $G = (I, J, E)$, where the demand nodes are represented by a set of vertices $i \in I$, the possible locations of the facilities are given by another set of vertices $j \in J$, and edges $e_{i,j} \in E$ only exist between vertices from $i \in I$ to those in $j \in J$. Furthermore, we assign positive weights to the edges $d_{i,j} \geq 0$, which represents a distance between vertices i and j. Note that it is possible to have zero distance between a demand node and a possible facility location. We also assign positive weights to the demand nodes h_i for $i \in I$, and this represents the amount of demand at a particular node. In this problem, we would like to place p facilities to minimize an average distance between a demand node and its servicing facility.

1.2 ILP Formulation

Because there are no capacity constraints at the facilities, we can assume that each demand node is serviced by a single facility. To formulate an integer linear program (ILP) to solve this problem, we define a decision variable

$$Y_{i,j} = \begin{cases} 1, \text{if demand node } i \in I \text{ assigned to facility located at } j \in J \\ 0, \text{otherwise} \end{cases}$$

that describes which demand nodes are serviced by which facility location. We need to define another decision variable

$$X_j = \begin{cases} 1, \text{if facility located at } j \in J \\ 0, \text{otherwise} \end{cases}$$

that describes the locations at which a facility is placed. Given these decision variables, we can now formulate the p-median problem as the following ILP

$$\begin{aligned} \min & \sum_{j \in J} \sum_{i \in I} h_i d_{i,j} Y_{i,j} \\ \text{s.t.} & \sum_{j \in J} Y_{i,j} = 1, \forall i \in I \\ & Y_{i,j} - X_j \leq 0, \forall i \in I, j \in J \\ & \sum_{j \in J} X_j = p \\ & X_j \in \{0,1\}, \forall j \in J \\ & Y_{i,j} \in \{0,1\}, \forall i \in I, j \in J. \end{aligned}$$

Each of these terms has associated intuition. The objective $\sum_{j \in J} \sum_{i \in I} h_i d_{i,j} Y_{i,j}$ is stating that we wish to minimize the demand-weighted (i.e., weighted by h_i) distance $d_{i,j} Y_{i,j}$ summed over all facilities and demand nodes. The first constraint $\sum_{j \in J} Y_{i,j} = 1$ implies that a demand node i can only be serviced by one facility. The constraint $Y_{i,j} - X_j \leq 0$ says that demand node i can be serviced by a facility at j only if there is a facility at j, because if $X_j = 0$ then we must have that $Y_{i,j} = 0$. The constraint $\sum_{j \in J} X_j = p$ means that we must place exactly p facilities. Lastly, the constraints that $X_j \in \{0,1\}$ and $Y_{i,j} \in \{0,1\}$ force the decision variables to be binary.

1.3 Heuristic Algorithm

The ILP can be difficult to solve, so we discuss one possible heuristic algorithm to solve this problem. There are two stages to the heuristic algorithm that we will discuss: The idea is to begin with a greedy placement of the p facilities in the first stage of the algorithm, and then to refine the placement of the facilities within neighborhoods in the second stage of the algorithm. The first stage is

1. Place the first facility using brute force enumeration to solve the 1-median problem;
2. For $i = 2, \dots, p$
   (a) Keeping the location of already placed facilities fixed, place another facility to minimize $\sum_{j \in J} \sum_{i \in I} h_i d_{i,j} Y_{i,j}$.

The second stage is

1. Find the neighborhood of each facility (meaning an assignment of demand nodes to each facility, such that the distance between a demand node and facility is minimum)

2. Do
    (a) Solve the 1-median problem in each neighborhood;
    (b) Find the neighborhood of each facility
3. While the neighborhoods have changed from the previous iteration