Property Demonstration on the Greedy Algorithm regarding Center Selection Problem with Various Examples

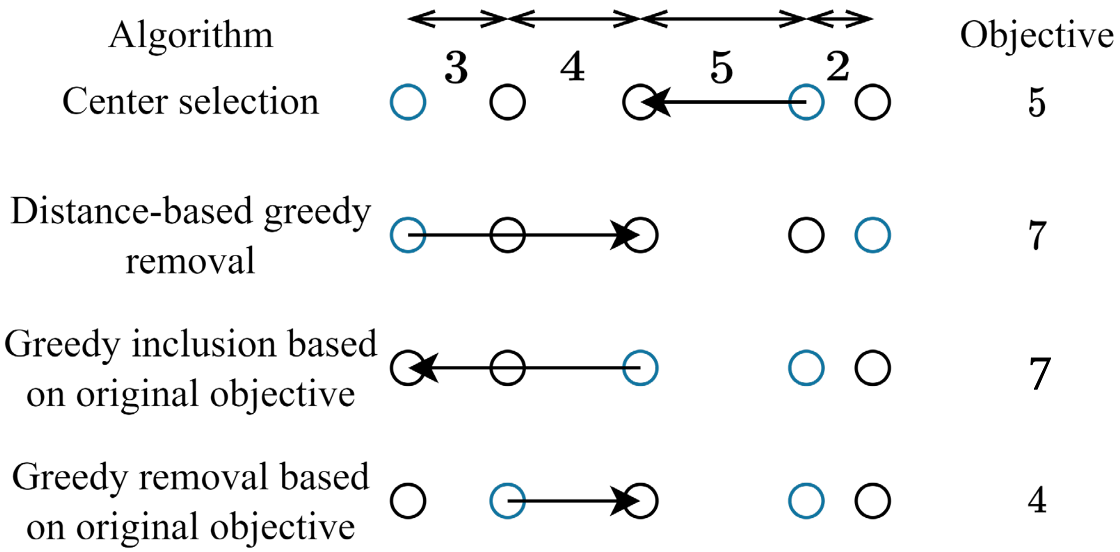
陈子蔚 Chen Ziwei

Department of Computer Science and Engineering

Recent Research Topic: Generative Model

Supervisor: 郑锋

Task 6-1



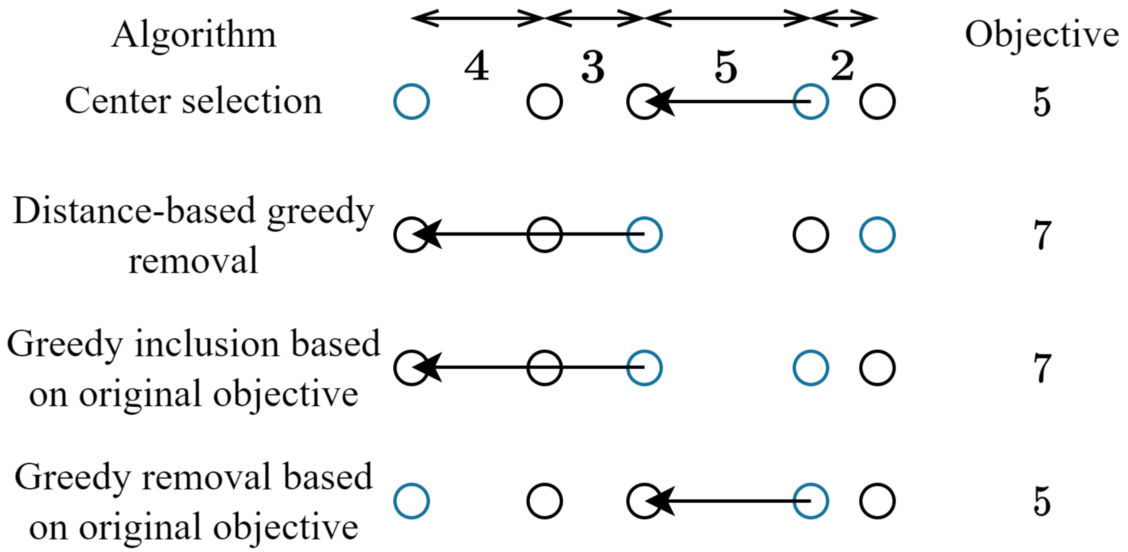
Center selection algorithm could never get the optimal solution in this case.

Distance-based greedy removal algorithm might obtain a quite bad solution with a bad initial center.

The selection of the second center in Greedy inclusion algorithm based on original objective is meaningless since it helps nothing on objective, which means there are many sites with the same best evaluation.

Greedy removal algorithm based on the original objective obtains the optimal solution when it has the greatest time cost among them.

Task 6-2



By changing the sample a little bit, some algorithms comes up with a different solution than before when some objectives remain.

We find that greedy removal algorithm based on the original objective would obtain a worse objective than the optimal now since there are two sites with the same best evaluation in the second step.

Task 6-3

Consider a simple situation with $k = 1$, then

$$\mathrm{dist}(s, C) = \mathrm{dist}(s, c) = \|c - s\|$$

As for the second objective function,

$$\begin{aligned}\frac{\partial \sum_{s\in\mathcal{S}} \|c-s\|^2}{\partial c} &= \frac{\partial \sum_{s\in\mathcal{S}} (c-s)^\top (c-s)}{\partial c} \\ &= \frac{\sum_{s\in\mathcal{S}} \partial (c-s)^\top (c-s)}{\partial c} \\ &= \sum_{s\in\mathcal{S}} 2(c-s) \\ &= 2\left(|\mathcal{S}| \cdot c - \sum_{s\in\mathcal{S}} s\right)\end{aligned}$$

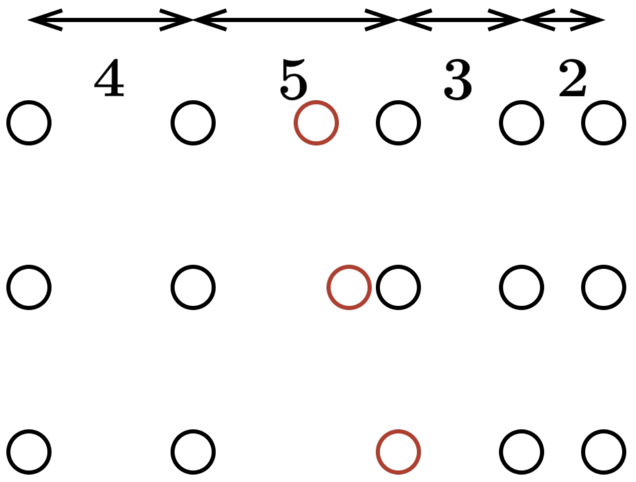
$$\frac{\partial \sum_{s\in\mathcal{S}} \|c-s\|^2}{\partial c} = 0 \Rightarrow c = \frac{\sum_{s\in\mathcal{S}} s}{|\mathcal{S}|}$$

Objective Funtion

$\min_c \max_{s\in\mathcal{S}} ||c - s||$

$\min_c \sum_{s\in\mathcal{S}} ||c - s||^2$

$\min_c \sum_{s\in\mathcal{S}} ||c - s||$



Task 6-3

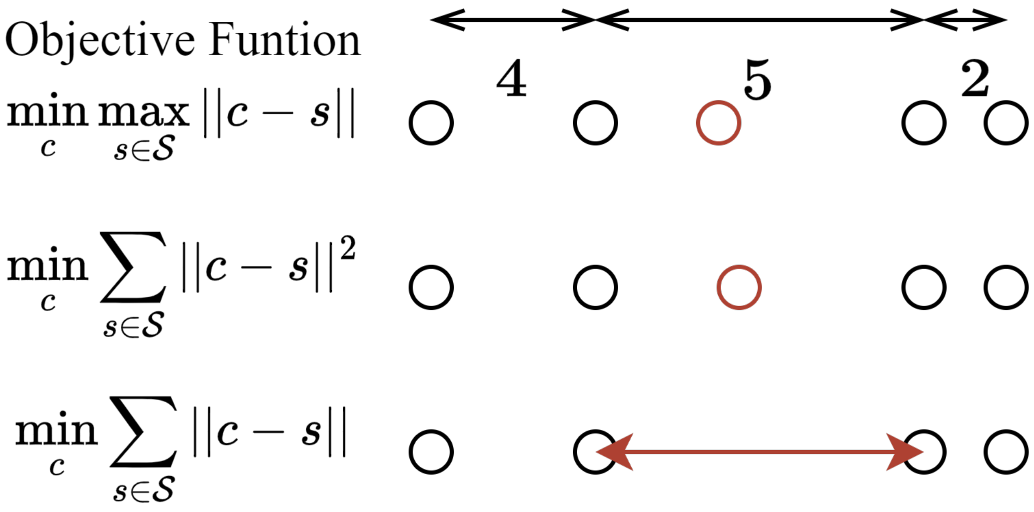
As for the third objective function,

$$\frac{\partial \sum_{s\in\mathcal{S}} \|c - s\|}{\partial c} = \frac{\sum_{s\in\mathcal{S}} \partial \|c - s\|}{\partial c}$$
$$= \sum_{s\in\mathcal{S}} \frac{c - s}{\|c - s\|}$$

It is not extractable when $c \in \mathbb{R}^d$ and d is quite large.
Assume $d = 1$,

$$\frac{\partial \sum_{s\in\mathcal{S}} \|c - s\|}{\partial c} = \sum_{s\in\mathcal{S}} \frac{c - s}{\|c - s\|}$$
$$= \sum_{s\in\mathcal{S}} (2 \cdot \mathbf{1}\{\mathrm{c} > \mathrm{s}\} - 1)$$

$\sum_{s\in\mathcal{S}} \|c - s\|$ would reach minimal when its differential is 0 or -1 or 1.



In the example, the solution of the third formulation could be arbitrary point on the red bidirectional arrow.