

Math Handout - Quantum K-Means

Surfers

31/01/2021

Problem

Let \mathcal{P} be a set of N points, e.g. County Districts in the United States.

Let \mathcal{D} be a N by N matrix with the distances between all pairs of points, $\mathcal{D}_{i,j}$ is the distance between points i and j .

We will find K groups Φ_k of equal size (or with equal resource distribution such as population) that minimize the objective function:

$$\sum_j \sum_{p_a \in \Phi_j} \sum_{p_b \in \Phi_j} \mathcal{D}_{i,j}$$

That is, the total sum of distances within points of each cluster.

Our constraint is that we should have an equal target number of points for each cluster, $M = \frac{N}{K}$

$$\sum_j (M - |\Phi_j|)^2$$

This can be modeled as a DQM problem where each point has $\phi_i \in \{0, 1, \dots, K\}$ options and we minimize the previous equations.

On our code however, for simplicity, we preferred to map the DQM problem to a QUBO problem explicitly.

Instead of having $\phi_i \in \{0, 1, \dots, K\}$, we have binary variables $x_{i,j}$ such that $x_{i,j}$ means point i belongs to cluster j .