Better Algorithms for Individually Fair k-Clustering

Anonymous Author(s)
Affiliation
Address
email

Abstract

We study data clustering problems with ℓ_p-norm objectives (e.g. k-MEDIAN and k-MEANS) in the context of individual fairness. The dataset consists of n points, and we want to find k centers such that (a) the objective is minimized, while (b) respecting the individual fairness constraint that every point v has a center within a distance at most $r(v)$, where $r(v)$ is v's distance to its (n/k)th nearest point. Jung, Kannan, and Lutz [FORC 2020] introduced this concept and designed a clustering algorithm with provable (approximate) fairness and objective guarantees for the ℓ_∞ or k-CENTER objective. Mahabadi and Vakilian [ICML 2020] revisited this problem to give a local-search algorithm for all ℓ_p-norms. Empirically, their algorithms outperform Jung et. al.'s by a large margin in terms of cost (for k-MEDIAN and k-MEANS), but they incur a reasonable loss in fairness. In this paper, our main contribution is to use Linear Programming (LP) techniques to obtain better algorithms for this problem, both in theory and in practice. We prove that by modifying known LP rounding techniques, one gets a worst-case guarantee on the objective which is much better than in MV20, and empirically, this objective is extremely close to the optimal. Furthermore, our theoretical fairness guarantees are comparable with MV20 in theory, and empirically, we obtain noticeably fairer solutions. Although solving the LP *exactly* might be prohibitive, we demonstrate that in practice, a simple sparsification technique drastically improves the run-time of our algorithm.

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.