
Potage of algorithmic designs

MEHRDAD GHADIRI, Georgia Institute of Technology
Socially Fair k-Means Clustering

We show that the popular k-means clustering algorithm (Lloyd's heuristic), used for a variety of scientific data, can result in outcomes that are unfavorable to subgroups of data (e.g., demographic groups). Such biased clusterings can have deleterious implications for human-centric applications such as resource allocation. We present a fair k-means objective and algorithm to choose cluster centers that provide equitable costs for different groups. The algorithm, Fair-Lloyd, is a modification of Lloyd's heuristic for k-means, inheriting its simplicity, efficiency, and stability. In comparison with standard Lloyd's, we find that on benchmark datasets, Fair-Lloyd exhibits unbiased performance by ensuring that all groups have equal costs in the output k-clustering, while incurring a negligible increase in running time, thus making it a viable fair option wherever k-means is currently used.

NICK HARVEY, University of British Columbia
How to make predictions using two experts, forever

A fundamental task in learning theory is to make repeated predictions given advice from experts. A famous algorithm for this task is the multiplicative weight method. It was independently invented in many areas including game theory, computational geometry, online learning, discrete optimization, and information theory.

As an algorithm designer, I am compelled to ask: is the multiplicative weight method the optimal algorithm for this problem? It turns out that it is the optimal algorithm under two assumptions: (1) the number of "experts" is large, and (2) the duration of the task is known in advance. We consider the optimality question after removing both of these assumptions. For the setting with just two experts in which the duration is unknown, we find the optimal algorithm. It is provably better than the multiplicative weight method.

The algorithm is derived in an unusual way. We reformulate the problem so that time progresses continuously rather than discretely. After doing so, the optimal algorithm can be found via a partial differential equation, which we solve analytically. This PDE solution is then used as the algorithm in discrete time. One would expect this approach to incur some discretization error, but bizarrely the discretization error is negative!

Joint work with Chris Liaw, Sikander Randhawa and Ed Perkins.

RICHARD SANTIAGO, ETH Zürich
Non-monotone weakly submodular function maximization subject to a cardinality constraint

Weak submodularity is a natural relaxation of the diminishing return property, which is equivalent to submodularity. Weak submodularity has been used to show that many (monotone) functions that arise in practice can be efficiently maximized with provable guarantees. In this talk we introduce a natural generalization of weak submodularity for non-monotone functions. We show that an efficient randomized greedy algorithm has provable approximation guarantees for maximizing these functions subject to a cardinality constraint.

BRUCE SHEPHERD, UBC
Single Tree Cut Approximators and Disjoint Paths in Outerplanar Graphs

We show the existence of a tree which approximates all cut capacities in an outerplanar within a constant factor. We also show that this can be used to give a constant factor approximation to the maximum weight edge-disjoint path problem (MEDP). Such an algorithm for MEDP was previously only known for trees.