Homework 4 for "Algorithms for Big-Data Analysis"

Acknowledgement:

May 7, 2015

Note: Please write up your solutions independently. If you get significant help from others, write down the source of references. A formal mathematical proof for all your claims is required.

1. (选做题) (The Original Bellman-Ford Method). Given a directed graph (N, A) with nodes numbered $1, \ldots, n$. Each arc $(i, j) \in A$ has a cost or "length" a_{ij} associated with it. Consider the single origin/all destinations shortest path problem. The Bellman-Ford method, as originally proposed by Bellman and Ford, updates the labels of all nodes simultaneously in a single iteration. In particular, it starts with the initial conditions

$$d_1^0 = 0, d_j^0 = \infty, \forall j \neq 1,$$

and generates d_j^k, $k = 1, 2, \ldots$, according to

$$d_1^k = 0, d_j^k = \min_{(i,j) \in A} \{d_i^{k-1} + a_{ij}\}, \forall j \neq 1.$$

   (a) Show that for all $j \neq 1$ and $k \geq 1$, d_j^k is the shortest distance from 1 to j using paths with k arcs or less, where $d_j^k = \infty$ means that all the paths from 1 to j have more than k arcs.

   (b) Assume that all cycles have nonnegative length. Show that the algorithm terminates after at most n iterations, in the sense that for some $k \leq n$ we have $d_j^k = d_j^{k-1}$ for all j. Conclude that the running time of the algorithm is $O(n \cdot |A|)$.

2. (选做题) Let $G = (V, E)$ be a directed network. Let t be a designated sink, and let $S \subseteq V$ be a set of sources. Each source $s \in S$ needs to transmit data to t, and quality of service requirements necessitate that it exclusively reserve an entire path to the exclusion of all other sources. Formally, a routing for a set of sources $R \subseteq S$ is a family of edge-disjoint paths $\{p_s : s \in R\}$, where p_s is a path from s to t.

   (a) Given a set $R \subseteq S$, show how to compute in polynomial time a routing for R if one exists, or certify that such a routing does not exist. (Hint: you may use the fact that the Ford-Fulkerson algorithm can compute an integral max flow from one source to one sink, so long as the network has integer capacities)

   (b) Call a set $R \subseteq S$ feasible if a routing for R exists. Describe the convex hull of indicator vectors of feasible sets as a polytope in inequality form. (Hint: use an inequality for each cut in the graph).

   (c) Given weights $w : S \to \mathbb{R}_+$, show how to use your answer to part (b) to compute a maximum weight feasible set, and a corresponding routing, in polynomial time. Your algorithm should show how to solve the LP you described in (b).

3. Let F be a nondecreasing submodular set function on X, and q be a real number.

(a) Prove that the function

$$G(U) = \min(q, F(U)), \forall U \subseteq X$$

is submodular.

(b) Is monotonicity required? If so, show that monotonicity cannot be removed. If not, show that monotonicity is not critical for this property to be true.

(c) What about non-increasing (decreasing) functions in min? That is, is submodularity preserved in this case as well when F is monotone non-increasing? Prove or give a counterexample.

4. Properties of Submodular Functions

   (a) Prove that any non-negative submodular function is also subadditive, i.e. if $F : 2^X \to \mathbb{R}_+$ is submodular then $F(S \cup T) \leq F(S) + F(T)$ for any $S, T \subseteq X$. Here, $\mathbb{R}_+ = \{x \in \mathbb{R}, x \geq 0\}$.

   (b) Prove that a function $F : 2^X \to \mathbb{R}_+$ is submodular if and only if for any $S, T \subseteq X$, the marginal contribution function $F_S(T) = F(S \cup T) - F(S)$ is subadditive.

5. Given finite ground set X , and given $w_d \in [0, 1]$ for all $d \in X$, define

$$F(S) = \Pi_{d \in S} w_d,$$

   where $F(\emptyset) = 1$. Is this submodular, supermodular, modular, or neither?