Supplementary Material

Avradeep Bhowmik* Vivek Borkar† Dinesh Garg‡ Madhavan Pallan§

A Basics of Submodular Functions

Consider functions on a given set $V = \{1, 2, 3, \cdots, n\}$ and its subsets. A function $f : 2^V \to \mathbb{R}$ is submodular if and only if it satisfies following equivalent conditions.

- **Definition:** For every $A, B \subset V$, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B)$$

- **First Order Difference:** For every $A, B \subset V$ such that $A \subseteq B$, and every $k \in V$ such that $k \notin A, B$, we have

$$f(A \cup \{k\}) - f(A) \geq f(B \cup \{k\}) - f(B)$$

- **Second Order Difference:** For every $A \subset V$ and every $j, k \in V$ such that $j, k \notin A$, we have

$$f(A \cup \{k\}) - f(A) \geq f(A \cup \{j, k\}) - f(A \cup \{j\})$$

A supermodular function can be defined with all the above inequalities reversed. Following are some useful facts about submodular functions.

- The negative of a submodular function is a supermodular function and vice versa. If a function is both submodular as well as supermodular then it is called a modular function and all the above inequalities become equalities.
- The function f is called a monotone submodular function if for all $A \subset B \in 2^V$ we have either $f(A) \leq f(B)$ or $f(A) \geq f(B)$.

In what follows, we highlight a very useful class of modular and supermodular functions.

Lemma A.1. *Let $f : 2^V \mapsto \mathbb{R}$ be a set function such that for every set $A \subseteq V$, it decomposes into sum of non-negative functions over all the elements of A as shown below*

$$f(A) = \sum_{v \in A} g(v) \tag{A.1}$$

where $g : V \mapsto \mathbb{R}^+ \cup \{0\}$. Then, the function $f(\cdot)$ is a modular function.

Proof. This follows from plugging this decomposition into the definition of submodular or supermodular functions.

Lemma A.2. *Let $f : 2^V \mapsto \mathbb{R}$ be a set function such that $f(A) = 0$ for every set $A \subseteq V$ having $|A| = 1$; and for every set $A \subseteq V$ having $|A| > 1$, it decomposes into sum of non-negative functions over every pair of elements of A as shown below*

$$f(A) = \sum_{(u,v)|u,v \in A} g(u, v) \tag{A.2}$$

where $g : V \times V \mapsto \mathbb{R}^+ \cup \{0\}$. Then, the function $f(\cdot)$ is a supermodular function.

Proof. This follows from the definition of supermodular functions. Consider the sets A, $A \cup \{i\}$ and $A \cup \{i, j\}$. Then,

$$f(A \cup \{i\}) - f(A) = \sum_{u \in A} g(u, i)$$

and

$$\begin{aligned} f(A \cup \{i, j\}) - f(A \cup \{j\}) &= \sum_{u \in A} g(u, i) + g(i, j) \\ &\geq f(A \cup \{i\}) - f(A) \end{aligned}$$

since $g(\cdot)$ is a non-negative function. Hence, the result.

B Proofs

Proof. [of Lemma 2.2] Consider the case when teams A and B both individually satisfy the skill requirements and they are disjoint teams. In such a case, its easy to verify from Equation (2.2) that $f_{skill}(A, P) = f_{skill}(B, P) = f_{skill}(A \cup B, P) = 1$ whereas $f_{skill}(A \cap B, P) = f_{skill}(\phi, P) = 0$, therefore, $f_{skill}(A, P) + f_{skill}(B, P) \geq f_{skill}(A \cup B, P) + f_{skill}(A \cap B, P)$. On the other hand, suppose team A satisfies only some of the requirements of the project, and team B separately satisfies the remaining requirements. Then, teams A and B together satisfy all the skill requirements but

*University of Texas, Austin, avradeep.1@utexas.edu
†IIT Bombay, India, borkar@ee.iitb.ac.in
‡IBM Research, India, garg.dinesh@in.ibm.com
§IBM Research, India, mapallan@in.ibm.com

neither team A nor team B individually meets all the requirements. Therefore, $f_{skill}(A,P) = f_{skill}(B,P) = f_{skill}(A \cap B, P) = 0$ and $f_{skill}(A \cup B, P) = 1$ implying $f_{skill}(A,P) + f_{skill}(B,P) \leq f_{skill}(A \cup B, P) + f_{skill}(A \cap B, P)$. Thus, we see that $f_{skill}(\cdot, P)$ is neither submodular nor supermodular. Note that the condition $|P| > 1$ is being used in the second scenario where team A satisfies only some of the requirements of the project, and team B separately satisfies the remaining requirements.

Proof. [of Lemma 2.4]

(1) This we prove by giving counterexamples. First we show that diameter function is neither a submodular nor a supermodular function. For this, consider the Figure 1 where we have shown two different scenarios corresponding to how a given set of experts might be connected via a social graph G. The edge weights represent communication cost between two nodes. In the first scenario, we consider the team

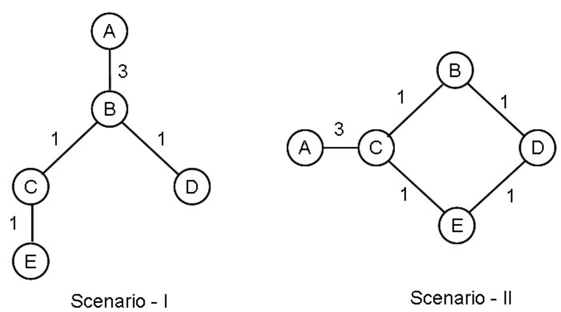

Figure 1: Counter Examples for Diameter being neither a submodular nor a supermodular

$T = \{B, C, E\}$. It is easy to verify that $dia(T \cup \{D\}) - dia(T) \geq dia(T \cup \{A, D\}) - dia(T \cup \{A\})$, where $dia(\cdot)$ represents the diameter of the given team. For the second scenario, this inequality reverses and hence the result follows.

For the MST also we take a similar approach and consider two different scenarios as shown in Figure 2. In the first scenario, we consider the team $T = \{A, D\}$. It is easy to verify that $MST(T \cup \{C\}) - MST(T) \geq MST(T \cup \{B, C\}) - MST(T \cup \{B\})$, where $MST(\cdot)$ represents the length of MST of the given team. For the second scenario, this inequality reverses and hence the result follows.

(2) This follows from Lemma A.1.

(3) This follows from Lemma A.2 with $g(u,v)$ being the distance function $d(u,v)$, and from Lemma A.1 with $g(v)$ being the leader distance $d(v,l)$.

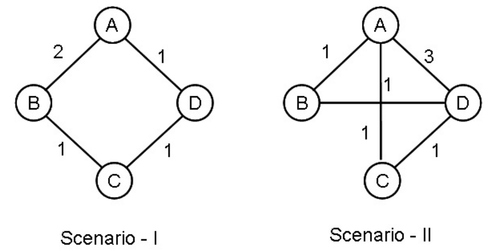

Figure 2: Counter Examples for MST being neither a submodular nor a supermodular

(4) We first prove the supermodularity of the modified sum of distances. For this, we show that for any two skills s_i and s_j, the function $d(s_i, s_j, \cdot)$ is a supermodular function over the team T.
To show that $d(s_i, s_j, \cdot)$ is a supermodular function, note that if $A \subset B$ then by definition, we must have $d(s_i, s_j, B) \leq d(s_i, s_j, A)$. Also, let $v \in V$ be some expert. If inclusion of v in the team A does not decrease the shortest distance between two skills s_i and s_j then inclusion of the same expert in the set B will also not decrease this distance. In such a case we would be having $f(A \cup \{v\}) - f(A) = f(B \cup \{v\}) - f(B)$.
However, if inclusion of v in the team A does decrease the shortest distance between skills s_i and s_j then inclusion of the same expert in the team B may not necessarily decrease the shortest distance, and even if it does decrease it the decrement cannot be more than that obtained through set A. Therefore, in this scenario, we must have $f(A \cup \{v\}) - f(A) \leq f(B \cup \{v\}) - f(B)$. This proves the supermodularity of the $d(s_i, s_j, \cdot)$ function.
Therefore the modified sum of the distances function, in both the weighted and the unweighted version, is also supermodular since a non-negative weighted sum of supermodular functions is still supermodular.
For the modified leader distance, in both the weighted and the unweighted version, one can use similar arguments but with the distance $d(s, l, \cdot)$ instead of the distance $d(s_i, s_j, \cdot)$.

Proof. [of Lemma 2.5] If the teaming cost is defined as the team size then its modularity follows from Lemma A.1. Coming to the personnel cost, for a given project P and a given expert v, we can define a function $g(v) = \sum_{s|s \in P \cap S_v} cost(v, s)$. In view of this function, the personnel cost is now decomposable into the elements of its argument set and hence by virtue of Lemma A.1 this is also a modular function.

Proof. [of Lemma 2.6] Let $A \subset B \subseteq V$ and $v \in V$ then its easy to verify that

$$\begin{aligned} f_{red}(B \cup \{v\}) - f_{red}(B) &= \sum_{v_i \in B} r(v, v_i) \\ f_{red}(A \cup \{v\}) - f_{red}(A) &= \sum_{v_i \in A} r(v, v_i) \end{aligned}$$

Given the fact that $r(v_i, v_j) \geq 0 \ \forall v_i, v_j \in V$, since $A \subset B$, the supermodularity of the function $f_{red}(\cdot)$ follows from the definition based on first order differences. Regarding $f_{expInc}(\cdot)$, this function satisfies the requirements of the Lemma A.1 and hence it is a modular function.