A Survey on
Distributed/Streaming Submodular Optimization

Jiecao Chen

February 18, 2016

1 Introduction

Submodularity is a property of set functions with deep theoretical and practical consequences. Submodular functions occur in a variety of applications, including representative skyline selection [10], network structure learning [3], influence maximization [4], document summarization [6], image segmentation [1, 5] and many others.

There is a large body of research in submodular optimization which we can not cover thoroughly here. In this survey, we focus on the recent advances in optimizing (mostly, maximizing) submodular functions in distributed and streaming setting.

1.1 Notations

Through out this survey,

we use V to represent the ground set we consider.

2^V is the power set of V (i.e. the set of all subsets of V).

For simplicity, we sometime write $A \cup \{x\}$ as $A + x$.

2 Submodularity

In this section, we first give several equivalent definitions of submodularity, and then we introduce several fundamental properties of submodular functions. We also discuss various constraints that occur frequently in submodular optimization problems. In the last part of this section, we cover algorithms that solve constrained submodular maximization problems with theoretical approximation guarantee.

2.1 Definitions

There are many equivalent definitions, and we will discuss three of them in this section.

Definition 1 (submodular concave) A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A, B \subseteq V$, we have that:

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B). \tag{1}$$

An alternate equivalent definition is more interpretable in many situations,

Definition 2 (diminishing returns) A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A \subseteq B \subset V$, and $v \in V \backslash B$, we have that:

$$f(A + v) - f(A) \geq f(B + v) - f(B). \tag{2}$$

Intuitively, this definition requires that the incremental "gain" of adding a new element v decreases (diminishes) as the base set grows from A to B. We will see that this property is actually shared by many real-world phenomenons.

It turns out that a stronger but equivalent statement can also serve as the definition of a submodular function,

Definition 3 (group diminishing returns) A function $f : 2^V \to \mathbb{R}$ is submodular if for any $A \subseteq B \subset V$, and $C \subseteq V \backslash B$, we have that:

$$f(A \cup C) - f(A) \geq f(B \cup C) - f(B). \tag{3}$$

2.2 Modularity and Supermodularity

We also briefly mention modularity and supermodularity here. These two concepts are closely related to submodularity.

A function $f : 2^V \to \mathbb{R}$ is modular if we replace inequality by equality in Definition 2 (or any of other two). Formally,

Definition 4 (Modularity) A function $f : 2^V \to \mathbb{R}$ is modular if for any $A \subseteq B \subset V$, and $v \in V \backslash B$, we have that:

$$f(A + v) - f(A) = f(B + v) - f(B). \tag{4}$$

Notably, a modular function f can always be written as

$$f(S) = f(\emptyset) + \sum_{v \in S} (f(\{v\}) - f(\emptyset))$$

for any $S \subseteq V$.

If we further assume $f(\emptyset) = 0$ (in this case, we call f normalized), we have a simplified expression,

$$f(S) = \sum_{v \in S} f(\{v\}).$$

Modularity can be useful in our discussion of submodularity, because one can use modular functions to construct submodular functions with desired properties in their applications. Examples can be found in e.g. [6,7].

A supermodular function is defined by flipping the inequality sign in the definition of a submodular function. Formally,

Definition 5 (Supermodularity) A function $f : 2^V \to \mathbb{R}$ is modular if for any $A \subseteq B \subset V$, and $v \in V \backslash B$, we have that:

$$f(A + v) - f(A) \leq f(B + v) - f(B). \tag{5}$$

We will focus on submodular functions because a function is supermodular if and only if its negative is submodular.

2.3 Properties

Like convex and concave functions, submodular functions have many nice properties. Lovász's description of convex functions [8] can be viewed as accurate comments on submodularity:

> - Convex functions occur in many mathematical models in economy, engineering, and other sciences. Convexity is a very natural property of various functions and domains occurring in such models; quite often the only non-trivial property which can be stated in general.
>
> - Convexity is preserved under many natural operations and transformations, and thereby the effective range of results can be extended, elegant proof techniques can be developed as well as unforeseen applications of certain results can be given.
>
> - Convex functions and domains exhibit sufficient structure so that a mathematically beautiful and practically useful theory can be developed.
>
> - There are theoretically and practically (reasonably) efficient methods to find the minimum of a convex function.

2.4 Constraints

2.5 Algorithms for Submodular Maximization

Greey algorithm guarantee is for normalized, monotone submodular maximization under cardinality constraint. Moreover, no polynomial time algorithm can provide a better approximation guarantee unless P = NP [2].

3 Applications

In this section, we discuss several representative applications in great detail. We will see from those examples that submodularity is such a natural property that many real-world problems can be cast in to the framework of submodular optimization.

- [9] distributed k-centers using submodular optimization.

4 Distributed Submodular Optimization

5 Streaming Submodular Maximization

References

[1] Y. Y. Boykov and M.-P. Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 1, pages 105–112. IEEE, 2001.

[2] U. Feige. A threshold of ln n for approximating set cover. *Journal of the ACM (JACM)*, 45(4):634–652, 1998.

[3] M. Gomez Rodriguez, J. Leskovec, and A. Krause. Inferring networks of diffusion and influence. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1019–1028. ACM, 2010.

[4] D. Kempe, J. Kleinberg, and É. Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 137–146. ACM, 2003.

[5] P. Kohli, M. P. Kumar, and P. H. Torr. P^3 & beyond: Move making algorithms for solving higher order functions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(9):1645–1656, 2009.

[6] H. Lin and J. Bilmes. A class of submodular functions for document summarization. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, pages 510–520. Association for Computational Linguistics, 2011.

[7] H. Lin and J. Bilmes. Word alignment via submodular maximization over matroids. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2*, pages 170–175. Association for Computational Linguistics, 2011.

[8] L. Lovász. Submodular functions and convexity. In *Mathematical Programming The State of the Art*, pages 235–257. Springer, 1983.

[9] G. Malkomes, M. J. Kusner, W. Chen, K. Q. Weinberger, and B. Moseley. Fast distributed k-center clustering with outliers on massive data. In *Advances in Neural Information Processing Systems*, pages 1063–1071, 2015.

[10] A. D. Sarma, A. Lall, D. Nanongkai, R. J. Lipton, and J. Xu. Representative skylines using threshold-based preference distributions. In *Data Engineering (ICDE), 2011 IEEE 27th International Conference on*, pages 387–398. IEEE, 2011.