Approximation Algorithm for Maximizing Nonnegative Weakly Monotonic Set Functions

Min Cui[1], Donglei Du[2], Dachuan Xu[3], and Ruiqi Yang[4](✉)

[1] Department of Operations Research and Information Engineering, Beijing University of Technology, Beijing 100124, People's Republic of China
B201840005@emails.bjut.edu.cn

[2] Faculty of Management, University of New Brunswick, Fredericton, NB E3B 5A3, Canada
ddu@unb.ca

[3] Beijing Institute for Scientific and Engineering Computing, Beijing University of Technology, Beijing 100124, People's Republic of China
xudc@bjut.edu.cn

[4] School of Mathematical Sciences, University of Chinese Academy Sciences, Beijing 100049, People's Republic of China
yangruiqi@ucas.ac.cn

Abstract. In recent years, submodularity has been found in a wide range of connections and applications with different scientific fields. However, many applications in practice do not fully meet the characteristics of diminishing returns. In this paper, we consider the problem of maximizing unconstrained non-negative weakly-monotone non-submodular set function. The generic submodularity ratio γ is a bridge connecting the non-negative monotone functions and the submodular functions, and no longer applicable to the non-monotone functions. We study a class of non-monotone functions, define as the weakly-monotone function, redefine the submodular ratio related to it and name it weakly-monotone submodularity ratio $\widehat{\gamma}$, propose a deterministic double greedy algorithm, which implements the $\frac{\widehat{\gamma}}{\widehat{\gamma}+2}$ approximation of the maximizing unconstrained non-negative weakly-monotone function problem. When $\widehat{\gamma} = 1$, the algorithm achieves an approximate guarantee of 1/3, achieving the same ratio as the deterministic algorithm for the unconstrained submodular maximization problem.

Keywords: Non-submodular optimization · Unconstrained · Submodularity ratio

1 Introduction

The research on combination problems with submodular property has received extensive attention in recent years. We say the function $f : 2^N \to R^+$ is submodular on the finite ground set N if and only if for any subsets S, T of N, we have:

D. Mohaisen and R. Jin (Eds.): CSoNet 2021, LNCS 13116, pp. 50–58, 2021.
https://doi.org/10.1007/978-3-030-91434-9_5

$$f(S) + f(T) \geq f(S \cup T) + f(S \cap T) \quad (1)$$

i.e. for any two subsets $S \subseteq T \subseteq N$ and $e \in N \setminus T$

$$f(S \cup \{e\}) - f(S) \geq f(T \cup \{e\}) - f(T).$$

We say the function f is monotone if for any two subsets S, T of N such that $S \subseteq T$, we have

$$f(S) \leq f(T) \quad (2)$$

Submodularity has a very intuitive interpretation in economics, which called the diminishing marginal utility. The diminishing marginal utility enables submodular functions to accurately simulate diversity and information gain in practical applications. At the same time, submodular functions can be solved accurately to minimize and approximately maximize in polynomial time [16]. These make submodular functions getting increasing attention in the field of artificial intelligence [24] and data mining [1], such as: social network influence [19], deep compressed sensing [22], sensor placement [21], targeted marketing [6], to name a few.

Unconstrained submodular problems is one of the most basic problem in submodular optimization. The factor of the approximation algorithm for the unconstrained submodular maximization problem is hardly better than 1/2 in polynomial time [8]. In fact, many basic NP-hard problems are special cases of unconstrained submodular maximization, including undirected cut problems [11], directed cut problems [14], the maximum facility location problems [18], and some limited satisfiability problems. In addition, the approximation algorithms of the unconstrained submodular maximization problem have been used as a subroutine of many other algorithms, such as social network marketing [15], and so on.

The study of unconstrained submodular maximization problems began in the 1960s [5]. Obviously, there are not many results. Feige et al. [9] were the first team to rigorously study general unconstrained submodular maximization issues: they proposed an uniform random subset algorithm and a local search algorithm, then increased the approximation guarantee to 2/5 by adding noise to the local search algorithm; they also showed that it may require exponential query to achieve an approximate ratio of $1/2+\epsilon$ in the value oracle model. Gharan et al. [12] and Feldman et al. [10] used methods such as simulated annealing to further improve the noisy local search technique. Buchbinder et al. [3] showed that a simple random algorithm strategy can be used to achieve the tight 1/2-approximate ratio. Later, their team [4] gave the de-randomized algorithm with the same approximate ratio. Roughgarden et al. [20] studied online unconstrained submodular maximization problem, provided a polynomial-time no-1/2-regret algorithm for this problem.

The applicability of submodular function is quite convenient and extensive. Is it possible to apply skills to connect the general function problem with the submodular problem, and then use the algorithms for submodular problem to solve the general problem effectively and securely? Based on one of the equivalent definitions of submodular functions, Das and Kempe [7] proposed the submodularity ratio $\gamma_{N,k}$ with respect to the ground set U and the parameter k, which is a quantity characterizing how close a general set function is to being submodular. Bian et al. [2] combined and generalized the ideas of curvature α and the submodularity ratio $\gamma_{N,k}$. Gong et al. [13] provided a more practical measurement γ which is called generic submodularity ratio which is depend on the monotonicity of the function.

In application, the problem with monotonic and submodularity is an idealized situation. For social network marketing, participants' decision-making are affected by the following conditions in actual social networks [17]: the social trust between participants, the social relationship between participants, and the preference similarity between participants. An appropriate number of "Big V", "Opinion Leader", "KOL" and "Online Celebrity" recommendations could make things widely spread without causing too much disgust, so as to ensure the positive growth of marketing effect. For the facility location problems, suppose the objective function is to evaluate the overall income of supermarkets in a city. The scope of the city would not change in a short time, and the construction and daily operation of the supermarket is a fixed cost. When the supermarket supply meets the urban demand [23], the newly opened supermarkets and previous supermarkets have to carry out price reduction and promotion in order to survive, which reduces the overall income.

Our Results. In this paper, we study a class of non-monotone functions which is a relaxed version of the monotonicity called the weakly-monotone, and discuss the problem of maximizing unconstrained weakly-monotone functions.

- We first give the definition of weakly-monotone function, then define the weakly-monotone submodularity ratio $\widehat{\gamma}$.
- Second, we present a deterministic double greedy algorithm for the unconstrained weakly-monotone maximization problem, and prove the algorithm achieves $\frac{\widehat{\gamma}}{\widehat{\gamma}+2}$-approximation ratio in $2n+2$ times query and in $O(n)$ computing time.

In addition, when the $\widehat{\gamma}$ reaches 1(i.e. the function is submodular), the approximation guarantee of the algorithm recovers the tight ratio as deterministic algorithm for the unconstrained submodular maximization problem.

Organization. The rest of this paper is structured as follows. In Sect. 2, we introduce the basic definitions and symbols used throughout this article, and give new definitions. We provide a deterministic algorithm for the unconstrained weakly-monotone functions maximization problem in Sect. 3, and give the approximate ratio analysis. Section 4 offers direction of future work.

2 Preliminaries

In this paper, we consider the problem of maximizing an unconstrained nonnegative weakly-monotone function, the objective is to select a subset S of the ground set N to maximize $f(S)$. The set function $f : 2^N \to R^+$ is a non-negative weakly-monotone set function with $f(\emptyset) = 0$. This problem can be stated as:

$$\max_{S \subseteq N} f(S). \tag{3}$$

$f(S \cup T) - f(S)$ denotes the marginal gain of adding the set $T \subseteq N$ to the set $S \subseteq N$. Specially, when the set $T = \{e\} \in N \setminus S$, the marginal gain of adding the single element e to the set S is defined as $f(S \cup \{e\}) - f(S)$.

Then, we defined the weakly-monotone function in following:

Definition 1. Weakly-monotone: *Let $f : 2^N \to R^+$ be a non-negative set function with for any two subsets of different sizes S, T of N that $f(S) \neq f(T)$. We say f is weakly-monotone if for any subset $A \subseteq N$ and any $e \in N \setminus A$,*

1. *if $f(A \cup \{e\}) > f(A)$, for any $S \subseteq N$ such that $A \subsetneqq S$, $f(A) \neq f(S)$;*
2. *if $f(A \cup \{e\}) < f(A)$, for any $S \subseteq N$ such that $A \subsetneqq S$, $f(A) > f(S)$.*

Example 1. $N = \{1, 2, 3, 4, 5\}$, $f(S) = \min\{|S|, |N| - |S| + 0.5\}$.

When $A = \emptyset$, $e \in N \setminus A$, $f(A \cup \{e\}) = 1 > 0 = f(A)$: for any $A \subsetneqq S \subseteq \{1, 2, 3, 4, 5\}$, $f(S) \in \{1, 2, 2.5, 1.5, 0.5\} \neq 0 = f(A)$;

When $A \in \{\{1\}, \{2\}, \{3\}, \{4\}, \{5\}\}$, $e \in N \setminus A$, $f(A \cup \{e\}) = 2 > 1 = f(A)$: for any $A \subsetneqq S \subseteq \{1, 2, 3, 4, 5\}$, $f(S) \in \{2, 2.5, 1.5, 0.5\} \neq 1 = f(A)$;

When $A \in \{\{1, 2\}, \{1, 3\}, \{1, 4\}, \{1, 5\}, \{2, 3\}, \{2, 4\}, \{2, 5\}, \{3, 4\}, \{3, 5\}, \{4, 5\}\}$, $e \in N \setminus A$, $f(A \cup \{e\}) = 2.5 > 2 = f(A)$: for any $A \subsetneqq S \subseteq \{1, 2, 3, 4, 5\}$, $f(S) \in \{2.5, 1.5, 0.5\} \neq 2 = f(A)$;

When $A \in \{\{1, 2, 3\}, \{1, 2, 4\}, \{1, 2, 5\}, \{1, 3, 4\}, \{1, 3, 5\}, \{1, 4, 5\}, \{2, 3, 4\}, \{2, 3, 5\}, \{2, 4, 5\}, \{3, 4, 5\}\}$, $e \in N \setminus A$, $f(A \cup \{e\}) = 1.5 < 2.5 = f(A)$: for any $A \subsetneqq S \subseteq \{1, 2, 3, 4, 5\}$, $f(S) \in \{1.5, 0.5\} < 2.5 = f(A)$;

When $A \in \{\{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \{1, 2, 4, 5\}, \{1, 3, 4, 5\}, \{2, 3, 4, 5\}\}$, $e \in N \setminus A$, $f(A \cup \{e\}) = 0.5 < 1.5 = f(A)$: for any $A \subsetneqq S \subseteq \{1, 2, 3, 4, 5\}$, $f(S) = 0.5 < 1.5 = f(A)$.

For the function that does not conform to the strictly property, we only need to add a small disturbance to the value of the function. For example, if there are two equal function values Q, we could add $\frac{Q}{n \cdot 10^k}$ to the second function value, and so on.

Then, we give the definition of the weakly-monotone submodularity ratio $\widehat{\gamma}$.

Definition 2. Weakly-monotone submodularity ratio: *For a non-negative weakly-monotone set function $f : 2^N \to R^+$, the weakly-monotone submodularity ratio of f is the largest scalar $\widehat{\gamma}$ satisfied the corresponding inequailities under these two cases, for any $S \subseteq N$ and any $e \in N \setminus S$:*

1. *When $f(S \cup \{e\}) - f(S) > 0$, for any $T \subseteq N \setminus \{e\}$ such that $S \subseteq T$:*

$$f(S \cup \{e\}) - f(S) \geqslant \widehat{\gamma} \mid f(T \cup \{e\}) - f(T) \mid;$$

2. When $f(S \cup \{e\}) - f(S) < 0$, for any $T \subseteq N \setminus \{e\}$ such that $S \subseteq T$:

$$f(S \cup \{e\}) - f(S) \geqslant \frac{1}{\widehat{\gamma}}[f(T \cup \{e\}) - f(T)].$$

Note 1. The weakly-monotone submodularity ratio $\widehat{\gamma}$ of f in Example 1 is 0.5.

For non-negative weakly-monotone set functions, the following lemma always holds.

Lemma 1. *Given a non-negative weakly-monotone set function $f : 2^N \to R^+$ with weakly-monotone submodularity ratio $\widehat{\gamma}$, it always holds that*

a. $\widehat{\gamma} \in (0, 1]$.
b. If the function f is submodular, then $\widehat{\gamma} = 1$.

Throughout this paper, denotes S as the output solution by the algorithm; denote O and OPT (i.e. $f(O) = OPT$) as the optimal set and the value of the optimal sets respectively. We assume that a single query on the oracle value requires $O(1)$ time.

3 The Deterministic-Greedy Algorithm

In this section, we propose a deterministic algorithm for maximizing unconstrained weakly-monotone functions. The algorithm runs in n iterations. In the i-th iteration, we only consider whether to keep the element e_i in the solution. The algorithm always maintains two feasible solutions S and T. The initial setting of S is an empty set and the T is the ground set N. The algorithm in arbitrary sequence checks each element $e_i \in N$ one by one to decide on adding it to S or deleting it from T. The decision is greedy depend on the size of marginal gain a_i of adding e_i to S and marginal gain b_i of abandoning e from S. If a_i is not less than b_i, adding e_i to S; otherwise, removing e_i from T. After traversing all the elements in N, we get $S = T$ which as the output of the algorithm. The principle of double greedy is intuitive as the operation for the element that we decide brings greater marginal benefits. A formal description of the algorithm appears as Algorithm 1.

In order to prove the approximate ratio of the algorithm, we give some additional notations. According to the construction of S and T, S staring with the empty set denotes S_0, T staring with the ground set N denotes T_0; in the $i-$th iteration, the algorithm either adds e_i to S_{i-1} or removes e_i from T_{i-1}. Record the S as S_i and record T as T_i when we have finished the operation.

First, we introduce an intermediate function $f((O \cup S_i) \cap T_i)$, using the change of intermediate function to bound the loss of $f(S)$ and $f(T)$ in each iteration.

Lemma 2. *For all $i = 1, 2, \cdots, n$:*

$$f((O \cup S_{i-1}) \cap T_{i-1}) - f((O \cup S_i) \cap T_i) \leq \frac{1}{\widehat{\gamma}}[f(S_i) - f(S_{i-1}) + f(T_i) - f(T_{i-1})]$$

Algorithm 1. Deterministic-Greedy

Input: evaluation oracle $f : 2^N \to R^+$
Output: the set S

Initialize $S \leftarrow \emptyset$, $T \leftarrow N$
for $i = 1$ to n **do**
  Initialize $a_i \leftarrow f(S \cup \{e_i\}) - f(S)$
  Initialize $b_i \leftarrow f(T \setminus \{e_i\}) - f(T)$
  if $a_i \geq b_i$ **then**
    $S \leftarrow S \cup \{e_i\}$
    $T \leftarrow T$
  else
    $S \leftarrow S$
    $T \leftarrow T \setminus \{e_i\}$
return S

Due to the length limitation, we only give the main idea of proof: from the relationship between a_i and b_i, the proof is divided into two cases: $a_i \geq b_i$ or $a_i < b_i$. In each case, we need to find out the relationship between S_{i-1}, T_{i-1} and S_i, T_i, and show one of a_i and b_i must more than 0; then we discuss whether e_i is in the optimal solution O and whether the value of $f((O \cup S_{i-1}) \cap T_{i-1})$ increases; in the end, using the definition of $\widehat{\gamma}$ to find the relationship between the $f((O \cup S_{i-1}) \cap T_{i-1}) - f((O \cup S_i) \cap T_i)$ and $f(S_i) - f(S_{i-1}) + f(T_i) - f(T_{i-1})$.

Then, we use the total change of intermediate function $f((O \cup S_i) \cap T_i)$ to bound the total loss of $f(S)$ and $f(T)$ at the algorithm.

Lemma 3.

$$f((O \cup S_0) \cap T_0) - f((O \cup S_n) \cap T_n)) \leq \frac{1}{\widehat{\gamma}}[f(S_n) + f(T_n) - f(S_0) - f(T_0)]$$

Proof. Summing up the inequalities in Lemma 2 for all $i = 1, 2, \cdots, n$, we get

$$\begin{aligned} &\sum_{i=1}^{n}(f((O \cup S_{i-1}) \cap T_{i-1}) - f(O \cup S_i) \cap T_i))) \\ \leq &\frac{1}{\widehat{\gamma}}\sum_{i=1}^{n}[f(S_i) - f(S_{i-1}) + f(T_i) - f(T_{i-1})] \end{aligned} \tag{4}$$

Combine the similar items, we have

$$\begin{aligned} &f((O \cup S_0) \cap T_0) - f((O \cup S_n) \cap T_n)) \\ \leq &\frac{1}{\widehat{\gamma}}[f(S_n) + f(T_n) - f(T_0)] \end{aligned} \tag{5}$$

Notice, at the begining of the algorithm the intermediate function $f((O \cup S_0) \cap T_0)$ is $f(O)$.

Theorem 1. *For any non-negative weakly-monotone function $f : 2^N \to R^+$, Algorithm Deterministic-Greedy is a $\frac{\widehat{\gamma}}{\widehat{\gamma}+2}$-approximation algorithm, the query complexity is $2n+2$, the computing time is $O(n)$.*

Proof. From the setting of S_i, T_i $(i = 0, 1, \cdots, n)$, we can easily get $S_n = T_n = S$; $(O \cup S_n) \cap T_n = S_n = S$. So

$$f(O) - f(S) \leq \frac{2}{\widehat{\gamma}} f(S)$$

Thus,

$$f(S) \geq \frac{\widehat{\gamma}}{\widehat{\gamma}+2} OPT.$$

Then, we consider the queries of the algorithm by two parts. The first part is to compute the value of a_i when every element e_i arrive, the number of queries in this part is $n+1$. The second part is to compute the value of b_i whenevery element e_i arrive, the number of queries in this part is also $n+1$. Consequently, the number of queries is $2n+2$. Thus, the computing time of the algorithm is $O(n)$ oracle queries plus $O(n)$ other operations.

From [3], we know that for any $\epsilon > 0$, $(1/3+\epsilon)$-approximation is tight for deterministic algorithm for unconstrained submodular maximization problem. When $\widehat{\gamma} = 1$, the algorithm has an approximation ratio of $1/3$. Therefore, we can say our algorithm is tight.

4 Discussion

Today, we would face the increasingly large data sets that are ubiquitous in modern machine learning and data mining applications. Greedy algorithm is highly continuous and would no longer have advantages in large-scale data sets. Various algorithms have been proposed to solve numerous submodular problems including large-scale problems according to application requirements, which can be roughly divided into centralized algorithms, streaming algorithms, distributed algorithms and decentralized framework. My recent interest is parallel algorithms for submodular problems. One future work is the research of parallel algorithms for maximizing an unconstrained non-negative weakly-monotone function on large-scale data sets or streaming setting.

Acknowledgements. The first and third authors are supported by National Natural Science Foundation of China (No. 12131003) and Beijing Natural Science Foundation Project No. Z200002. The second author is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) grant 06446, and Natural Science Foundation of China (Nos. 11771386, 11728104). The fourth author is supported by the Fundamental Research Funds for the Central Universities (No. E1E40108).

References

1. Badanidiyuru, A., Mirzasoleiman, B., Karbasi, A., Krause, A.: Streaming submodular maximization: massive data summarization on the fly. In: 20th International Proceedings on Knowledge Discovery and Data Mining, New York, NY, USA, pp. 671–680. Association for Computing Machinery (2014)
2. Bian, A.A., Buhmann, J.M., Krause, A., Tschiatschek, S.: Guarantees for Greedy maximization of non-submodular functions with applications. In: 35th International Conference on Machine Learning, NSW, Australia, vol. 70, pp. 498–507. Proceedings of Machine Learning Research (2017)
3. Buchbinder, N., Feldman, M., Naor, J.S., Schwartz, R.: A tight linear time (1/2)-approximation for unconstrained submodular maximization. SIAM J. Comput. **44**(5), 1384–1402 (2015)
4. Buchbinder, N., Feldman, M.: Deterministic algorithms for submodular maximization problems. In: 27th International Symposium on Discrete Algorithms, Arlington, VA, USA, pp. 392–403. Society for Industrial and Applied Mathematics (2016)
5. Cherenin, V.: Solving some combinatorial problems of optimal planning by the method of successive calculations. In: Conference of Experiences and Perspectives of the Applications of Mathematical Methods and Electronic Computers in Planning, Russian, Mimeograph, Novosibirsk (1962)
6. Coelho, V.N., et al.: Generic Pareto local search metaheuristic for optimization of targeted offers in a bi-objective direct marketing campaign. Comput. Oper. Res. **78**, 578–587 (2017)
7. Das, A., Kempe, D.: Submodular meets spectral: Greedy algorithms for subset selection, sparse approximation and dictionary selection. In: 28th International Conference on Machine Learning, Bellevue, Washington, USA, pp. 1057–1064. Omnipress (2011)
8. Dobzinski, S., Vondrak, J.: From query complexity to computational complexity. In: 44th International Symposium on Theory of Computing, New York, NY, USA, pp. 1107–1116. Society for Industrial and Applied Mathematics (2012)
9. Feige, U., Mirrokni, V.S., Vondrák, J.: Maximizing non-monotone submodular functions. SIAM J. Comput. **40**(4), 1133–1153 (2011)
10. Feldman, M., Naor, J.S., Schwartz, R.: Nonmonotone submodular maximization via a structural continuous greedy algorithm. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) ICALP 2011. LNCS, vol. 6755, pp. 342–353. Springer, Heidelberg (2011). https://doi.org/10.1007/978-3-642-22006-7_29
11. Galbiati, G., Maffioli, F.: Approximation algorithms for maximum cut with limited unbalance. Theor. Comput. Sci. **385**(1), 78–87 (2007)
12. Gharan, S.O., Vondrák, J.: Submodular maximization by simulated annealing. In: 22nd International Symposium on Discrete Algorithms, San Francisco, California, USA, pp. 1098–1116. Society for Industrial and Applied Mathematics (2011)
13. Gong, S., Nong, Q., Liu, W., Fang, Q.: Parametric monotone function maximization with matroid constraints. J. Global Optim. **75**(3), 833–849 (2019). https://doi.org/10.1007/s10898-019-00800-2
14. Halperin, E., Zwick, U.: Combinatorial approximation algorithms for the maximum directed cut problem. In: 20th International Symposium on Discrete Algorithms, Washington, DC, USA, pp. 1–7. Association for Computing Machinery/Society for Industrial and Applied Mathematics (2001)
15. Hartline, J., Mirrokni, V., Sundararajan, M.: Optimal marketing strategies over social networks. In: 17th International Proceedings on World Wide Web, Beijing, China, pp. 189–198. Association for Computing Machinery (2008)

16. Iwata, S., Fleischer, L., Fujishige, S.: A combinatorial strongly polynomial algorithm for minimizing submodular functions. J. Assoc. Comput. Mach. **48**(4), 761–777 (2001)
17. Mansell, R., Collins, B.S.: Introduction: Trust and Crime in Information Societies. Edward Elgar (2005)
18. Melo, M.T., Nickel, S., Saldanha-da-Gama, F.: Facility location and supply chain management - A review. Eur. J. Oper. Res. **196**(2), 401–412 (2009)
19. Mossel, E., Roch, S.: On the submodularity of influence in social networks. In: 30th International Symposium on Theory of Computing, San Diego, California, USA, pp. 128–134. Association for Computing Machinery (2007)
20. Roughgarden, T., Wang, J.: An optimal learning algorithm for online unconstrained submodular maximization. In: 31st International Conference on Learning Theory, Stockholm, Sweden, pp. 1307–1325. Proceedings of Machine Learning Research (2018)
21. Tran, A.K., Piran, M.J., Pham, C.: SDN controller placement in IoT networks: an optimized submodularity-based approach. Sensors **19**(24), 5474 (2019)
22. Tsai, Y.-C., Tseng, K.-S.: Deep compressed sensing for learning submodular functions. Sensors **20**(9), 2591 (2020)
23. Tsiang, S.C.: A note on speculation and income stability. Economica **10**(40), 27–47 (1943)
24. Zhang, H., Vorobeychik, Y.: Submodular optimization with routing constraints. In: 30th International Proceedings on Artificial Intelligence, Phoenix, Arizona, USA, pp. 819–825. AAAI Press (2016)