Constrained Robust Submodular Optimization

Thomas Powers\textsuperscript{1}, Jeff Bilmes\textsuperscript{1}, Scott Wisdom\textsuperscript{1}, David W. Krout\textsuperscript{2}, and Les Atlas\textsuperscript{1}

\textsuperscript{1} Department of Electrical Engineering, University of Washington
\{tcpowers, bilmes, swisdom, atlas\}@uw.edu
\textsuperscript{2} Applied Physics Laboratory, University of Washington
dkrout@apl.washington.edu

Abstract

In this paper, we consider the problem of constrained maximization of the minimum of a set of submodular functions, in which the goal is to find solutions that are robust to worst-case values of the objective functions. Unfortunately, this problem is both non-submodular and inapproximable. In the case where the submodular functions are monotone, an approximate solution can be found by relaxing the problem. We propose an algorithm called \textsc{Generalized Saturate (GENSat)} that exploits the submodular structure of the problem and, as a result, returns a near-optimal solution with a constant-factor approximation guarantee on the relaxed problem. \textsc{GENSat} can handle any submodular constraint, e.g. matroids, cover, and knapsack, and is compatible with any submodular maximization algorithm.

1 Introduction

Submodular function optimization problems have been well studied. Recent advancements have looked at even harder optimization problems that, while not submodular, have submodular structure. One particular problem is constrained maximization of the minimum of submodular functions, i.e., finding a solution that is robust to worst-case values of the objective functions. While this problem is NP-hard, inapproximable, and non-submodular, a relaxed version can be approximately solved if either the constraint or the problem is relaxed [7, 9]. Both of these methods rely on the saturation trick, in which a submodular surrogate problem is substituted for the original and maximized. The solution to the surrogate problem can then be used to find an approximation to the relaxed problem. Maximizing the minimum of a set of submodular functions is useful in problems like robust sensor selection, robust observation selection, submodular fair allocation, and other worst-case optimization problems [9, 7, 1].

Our main contribution is the proposed optimization algorithm, \textsc{Generalized Saturate (GENSat)}, that is related to \textsc{Saturate} [7]. \textsc{GENSat} maximizes the minimum of a set of submodular functions subject to any submodular constraint, e.g. matroids, cover, and knapsack, rather than just cardinality constraints. As a result, \textsc{GENSat} offers a new strategy for robust submodular optimization over constraint sets (such as matroids) that are not easily expanded as in the case of a cardinality constraints. In section\textsuperscript{2} we review submodularity and matroids. In section\textsuperscript{3} we describe \textsc{GENSat}. In section\textsuperscript{4} we provide experimental results that demonstrate the efficacy of \textsc{GENSat}. We conclude in section\textsuperscript{5}.

2 Submodularity

Submodularity is a property that describes set functions similar to how convexity describes functions in a continuous space. Rather than exhaustively searching over all combinations of subsets, submodular functions provide a fast and tractable framework to compute a solution [8, 4, 6].

Let the set of available objects, known as the ground set, be denoted as \( V \). A submodular function \( f \) maps a set of objects denoted by a binary indicator vector of length \( V \) to a real number. The
binary indicator vector is represented by the expression $2^V$ since the binary elements can take one of two values and are indexed by elements of set $V$. Since a value of 1 or 0 for the $i^{th}$ element of the indicator vector denotes the inclusion or exclusion of the $i^{th}$ element of the ground set $V$, any subset $A \subseteq V$ can be placed in one-to-one correspondence with incidence vectors.

Submodularity can be expressed via the notion of diminishing returns, i.e., the incremental gain of the objective diminishes as the context grows. If we define the incremental gain of adding $v$ to $A$ as $f(A \cup \{v\}) - f(A)$, then submodularity is defined as any function with $f(A \cup \{v\}) \geq f(A) + f(v)$ for all $A \subseteq B \subseteq V$ and $v \notin B$.

Submodularity is very closely tied to structures known as matroids, which generalize the notion of linear independence in vector spaces [2]. One can think of matroids as a generalization of matrices, which extend the definition of rank beyond column vectors of a matrix to more general independent subsets over a finite ground set. More importantly, submodular function optimization allows for matroid independence constraints to be placed on the problem, which means complicated variable dependence patterns can be encoded into the problem. Though matroids are combinatorial objects that grow exponentially with ground set size, their submodular nature allows approximate solutions in polynomial time. Given a finite set $V$ and a finite set of subsets $I = \{I_1, I_2, \ldots\}$, the pair $(V, I)$ is said to be a matroid when the family of sets $I$ satisfies the following three properties:

1. $\emptyset \in I$,
2. $I_1 \subseteq I_2 \in I$,
3. $I_1, I_2 \in I, |I_1| < |I_2| \implies \exists v \in I_2 \setminus I_1 : I_1 \cup v \in I$.

3 Proposed algorithm: GENERALIZED SATURATE (GENSAT)

We consider the constrained robust maximization problem given by

$$A^* \in \arg\max_{A \in C} \min_i f_i(A),$$

where $C$ denotes a family of feasible sets that are compatible with submodular optimization, e.g. matroid, a knapsack, cover, or other submodular constraint.

To optimize the Equation (1), we develop a novel algorithm GENSAT, which generalizes the SATURATE algorithm created by Krause et al. [7]. Krause et al. use SATURATE to optimize an objective function of the form

$$A^* \in \arg\max_{|A| \leq k} \min_i f_i(A)$$

where $f_i(A)$ is a set of monotone submodular functions. SATURATE solves this worst-case optimization problem by proposing an alternative formulation and relaxing the cardinality constraint from $|A| \leq k$ to $|A| \leq \alpha k$. As long as $\alpha$ is large enough, the solution $\hat{A}$ from the SATURATE algorithm guarantees that

$$\min_i f_i(\hat{A}) \geq \max_{|A| \leq k} \min_i f_i(A) \quad \text{and} \quad |\hat{A}| \leq \alpha k.$$ 

Krause et al. claim that the only way to achieve a non-trivial guarantee is to relax the constraint, which limits both the types of constraints that can be applied to the problem as well as the values the objective functions can take, i.e. integral or rational valued objective functions. Matroid constraints, for instance, have no immediately obvious relaxation. One way to relax the matroid constraint, however, might be to expand the bases of the matroid, which may lead to undesirable solutions. However, there is another way to achieve non-trivial guarantees, which is to relax the objective itself, leaving the constraints intact, and produce a fractional bound on the objective function, something that is made possible thanks to the use of submodularity. The proposed algorithm, GENSAT, uses this alternative approach to find a solution such that a fraction $\gamma$ of the submodular functions are above a minimum value $\beta$. Moreover, the user can set particular values of $\beta$ or $\gamma$, as long as $\beta$ is less than the submodular guarantee $\alpha$ and $\gamma < 1$. The derivation for the lower bound is given below.

For a fixed value of $c$, which can be thought of as the saturation level, we can determine if $f_i(A) \geq c$ from equation (1) via submodular maximization of the following surrogate function:

$$f^c(A) = \frac{1}{M} \sum_{i=1}^{M} \min\{f_i(A), c\}.$$  (3)
$f^c(A)$ is a submodular function, because it is a non-negatively weighted sum of functions $\min \{ f_i(A), c \}$ which are submodular. At each iteration of GENSAT, we run a submodular maximization algorithm like the one described in Algorithm 2 which selects the element which provides the best incremental gain in the objective and does not violate the constraint.

**Algorithm 1 GENSAT** ($f_1, \ldots, f_M, \alpha, \epsilon$)

1: $c_{\text{min}} \leftarrow 0, c_{\text{max}} \leftarrow \min \{ f_i(V) \}$
2: while $(c_{\text{max}} - c_{\text{min}}) > \epsilon$ do
3:    $c \leftarrow \frac{c_{\text{max}} - c_{\text{min}}}{2}$
4:    $f^c(A) \leftarrow \frac{1}{M} \sum_{i=1}^{M} \min \{ f_i(A), c \}$
5:    $\hat{A} \leftarrow \text{GREEDY}(f^c, c)$
6:    if $f^c(\hat{A}) < \alpha c$ then
7:        $c_{\text{max}} \leftarrow c$
8:    else
9:        $c_{\text{min}} \leftarrow c$,
10:       $A_{\text{best}} \leftarrow \hat{A}$
11: end if
12: end while
13: return $A_{\text{best}}$

**Algorithm 2 GREEDY** ($f^c, c$)

1: $A \leftarrow \emptyset$
2: while $\exists a \in V \setminus A$ s.t. $A \cup \{ a \} \in C$ and $f^c(a|A) > 0$ do
3:    $S \leftarrow \{ a \in V \setminus A : A \cup \{ a \} \in C \}$
4:    $s^* \leftarrow \text{argmax}_{s \in S} f^c(a|A)$
5:    $A \leftarrow A \cup \{ s^* \}$
6: end while
7: return $A$

**Theorem 1.** Given a value $\beta < \alpha$, GENSAT finds a solution $\hat{A}$ that guarantees the following fraction $\gamma$ of the $M$ functions $\min \{ f_i(\hat{A}), c \} \geq \beta c$:

$$\gamma \geq \frac{\alpha - \beta}{1 - \beta}$$

where $\alpha$ is the approximation guarantee for matroid independence set constrained submodular maximization problem.

**Proof:** $f^c(A) \geq c$ only if $\min_i f_i(A) \geq c$. When all $f_i(A) \geq c$, then $f^c(A) = c$. Likewise, when any $f_i(A) < c$ then the $f^c(A) < c$, since the maximum value of $f^c(A)$ is $c$. The greedy solution $\hat{A}$ for maximizing a monotone submodular function subject to a constraint is $f^c(\hat{A}) \geq \alpha f^c(A^*)$, where $\alpha$ depends on the algorithm chosen and the constraint. If line 6 of Algorithm 1 is true, then $f^c(\hat{A}) < \alpha c$ which implies that $\min_i f_i(A^*) < c$. Line 6 being true also implies that $c$ is too large, so $c_{\text{max}} \leftarrow c$. At line 12 of Algorithm 1, $(c_{\text{max}} - c_{\text{min}}) \leq \epsilon$ and the true optimal value $\min_i f_i(A^*)$ is in the interval $[c_{\text{min}}, c_{\text{max}}]$. The submodular approximation guarantee ensures that $f^c(A) \geq \alpha f^c(A^*)$ and $f^c(A) \geq \alpha c$. Given a value for $\beta$, the terms of $f^c(A)$ can be split into two groups, one that have value less than $\beta c$ and the other greater than or equal to $\beta c$:

$$f^c(\hat{A}) = \frac{1}{M} \sum_{i : \min \{ f_i(A), c \} < \beta c} \min \{ f_i(A), c \} + \frac{1}{M} \sum_{i : \min \{ f_i(A), c \} \geq \beta c} \min \{ f_i(A), c \}.$$

Let $\gamma$ be the fraction of terms that meet the $\beta c$ threshold. Then, the two summation terms become $f^c(\hat{A}) = (1 - \gamma) \beta c + \gamma c \geq \alpha c$. Rearranged, the expression becomes $\gamma \geq \frac{\alpha - \beta}{1 - \beta}$. $\square$

A visualization of the lower bound $\beta c$ over values of the saturation level $c$ and fractional number of functions $\lfloor \gamma M \rfloor$ for $M = 8$ is shown in figure 1. In this case, $\alpha = 1/2$, which is the submodular guarantee for a monotone submodular function constrained by a single matroid via the greedy algorithm [2]. Notice that the bound becomes trivial, $(\beta c = 0)$, when $\gamma \geq \alpha$, i.e. where $\lfloor \gamma M \rfloor \geq 5$. 

3
Constrained maximization of the minimum of a set of monotone submodular functions has several applications in machine learning, including robust sensor selection, robust observation selection, and submodular fair allocation. While this problem is non-submodular, NP-hard, and inapproximable, it is not entirely intractable. In this paper, we show that through relaxation of the minimum, a fraction of the functions in the minimum are lower-bounded. We present the algorithm GENSAT that achieves the lower bound and derive the formula theoretical lower bound. We then demonstrate that GENSAT can obtain near optimal performance on a matroid-constrained sensor placement problem.
References


