



OPTIMALITY RESULTS OF SUBMODULAR PROGRAMMING PROBLEMS

J. DASMAHAPATRA¹, A. PATTNAIK¹, S. K. PADHAN^{1,*}, AND C. NAHAK²

¹*Department of Mathematics, Veer Surendra Sai University of Technology Burla, Sambalpur 768018, India*

²*Department of Mathematics, Indian Institute of Technology Kharagpur, Kharagpur 721302, India*

ABSTRACT. Submodular functions have emerged as a class of set functions that model the law of diminishing returns in discrete functions. These functions have primarily been applied in discrete optimization problems because they can be minimized exactly and maximized approximately. The present investigation has considered both unconstrained and constrained minimization programming problems involving submodular set functions. It has been observed that the submodular minimization problems has an optimal solution if and only if its Lagrangian has a saddle point. Furthermore, the necessary and sufficient optimality conditions for submodular minimization programs with inequality constraints have been derived, which constitute the discrete analogue of the Karush–Kuhn–Tucker (KKT) conditions for convex programming problems.

Keywords. Submodular functions; Submodular programs; KKT conditions; Optimality conditions.

© Optimization Eruditorum

1. INTRODUCTION

Submodularity has evolved as an important property of set functions that has frequently arisen in various fields of research like matroids theory [5], combinatorial analysis, combinatorial optimization [4, 24], graph theory [15, 14], network computing [25], algebra [9], conversational English speech [19] and many more [3, 17, 26]. In recent years, submodularity has achieved significant popularity in machine learning [7, 28] and in several emerging areas of computer science, such as document summarization [18], image segmentation [16], *etc.* Many researcher have been exploring algorithmic and theoretical aspects of submodular set functions for more than past three to four decades. Considering the concept of polyhedron Edmonods [11], Fujishige [12] has developed the characterization of submodular polyhedron as the base polyhedron and subdifferentials of submodular functions. Fujishige [13] has provided optimality conditions for submodular minimization problem with (only for equality) or without constraints using subdifferentials. Morris [20] has discussed a common theory to optimize set functions and has obtained the necessary conditions for optimality of constrained minimization problems with inequality constraints. Skowron [24] has presented fixed parameter tractable approximation schemes for both the minimization and maximization variants of the problem. Qian *et al.* [23] has studied the general problem classes of maximizing submodular functions with or without a size constraint and maximizing approximately submodular monotone functions with a size constraint.

Submodular functions have been interpreted as the discrete counterparts of convex functions and have seen in various applications. Their structural properties have also been explored algorithmically. Several researchers have employed convexity to optimize the submodular functions. Bach [1] has extended the submodular functions to convex optimization and the probability measures. Pedramfar *et al.* [22] has attempted to maximize the continuous general submodular or DR-submodular functions over

*Corresponding author.

E-mail address: joydevphd14@gmail.com (J. Dasmahapatra), ashapurna.pattnaik@jku.edu.in (A. Pattnaik), skpadhan_math@vssut.ac.in (S. K. Pradhan), and cnahak@maths.iitkgp.ac.in (C. Nahak)

2020 Mathematics Subject Classification: 90C27, 90C26, 90C30, 90C46.

Accepted: August 16, 2026.

a convex set using Frank-Wolfe type offline algorithm. Chakrabarty *et al.* [8] has considered parallel complexity of submodular optimization problem and has provided two methods for query versus depth tradeoffs of a submodular function by using parallel convex optimization. Zhang and Zhong [27] have employed submodular optimization approach for covariate selection in a problem of high-dimensional casual effect estimation.

Submodular maximization problems under various constraints have been characterized as fundamental optimization problems, and significant research has been conducted to develop algorithms that approximately optimize the multilinear extension of the submodular objective function. Oshima [21] has considered maximization of unconstrained k -submodular functions using the improved randomized algorithm for non monotone functions that gives $\frac{k^2+1}{2k^2+1}$ -approximation for $k \geq 3$. Buchbinder and Feldman [6] have developed a solver with 0.401-approximation for constrained maximization of DR-submodular functions. A submodular optimization problem with chance constraints has been studied by Doerr *et al.* [10], who have solved it using a greedy algorithm based on surrogate functions that approximate constraint disruptions through Chernoff bounds. Balcan and Harvey [2] have worked on developing algorithms for the learning theory of submodular functions and have also obtained some novel structural results applicable in machine learning.

In this paper, a submodular function with diminishing return property has been considered. We have defined the Lagrangian for unconstrained and constrained (with equality constraints) submodular minimization problems and have proved that such minimization problems have an optimal solution if and only if their Lagrangians have saddle points. Furthermore, we have developed KKT type conditions for submodular minimization problems with inequality constraints.

2. PRELIMINARIES

Let us discuss some preliminary definitions for submodular functions.

Definition 2.1. Let V be a finite set. The power set of V is denoted by 2^V (throughout this paper). Let $\mathcal{A} \subseteq 2^V$ be constructed with the subsets of V and $\phi, V \in \mathcal{A}$. A set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ defined on $\mathcal{A} \subseteq 2^V$ to the set of all reals $\mathbb{R}$ is said to be a submodular if

$$\mu(M \cup N) + \mu(M \cap N) \leq \mu(M) + \mu(N), \text{ for each subsets } M, N \in \mathcal{A}. \quad (1)$$

W.l.o.g., we assume $\mu(\phi) = 0$ for the rest part of the paper.

Definition 2.2. The marginal value of the element $i \in V$ w.r.t., $M \in 2^V$ is given by

$$\mu_M(i) = \mu(M \cup \{i\}) - \mu(M).$$

Now the set function $\mu : 2^V \rightarrow \mathbb{R}$ is submodular if

$$\mu_M(i) \geq \mu_N(i), \forall N \in 2^V, M \subseteq N, i \notin N. \text{ (Diminishing returns property)} \quad (2)$$

The set of all $|V|$ -real vectors $r := (r(v) \mid v \in V)$ is denoted by $\mathbb{R}^V$, where $|V|$ is the cardinality of the set V . For each $M \in 2^V$ and $r \in \mathbb{R}^V$, we can define

$$r(M) := \sum_{v \in M} r(v) \text{ and } r(\phi) = 0, \phi \in 2^V. \quad (3)$$

Definition 2.3. The submodular polyhedron $\mathcal{P}_\mu$ and base polyhedron $\mathcal{B}_\mu$ of a submodular function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ are defined as

$$\mathcal{P}_\mu = \{r \mid \mu(M) \geq r(M), r \in \mathbb{R}^V, \forall M \in \mathcal{A}\}, \quad (4)$$

$$\mathcal{B}_\mu = \{r \mid \mu(E) = r(E), r \in \mathcal{P}_\mu\}. \quad (5)$$

Note.1: The extreme points of the submodular polyhedron and the base polyhedron of a submodular function are exactly same.

3. SUBGRADIENTS AND SUBDIFFERENTIALS OF SUBMODULAR FUNCTIONS

The fundamental and most important concepts for minimizing a submodular function are subgradients and subdifferentials of a submodular function.

Definition 3.1. A subgradient $r \in \mathbb{R}^V$ of a submodular function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ for a subset(or at) $M \in \mathcal{A}$ is defined as

$$\mu(T) - \mu(M) \geq r(T) - r(M), \forall T \in \mathcal{A} \quad (6)$$

where $\mathcal{A} \subseteq 2^V$ and $\phi, V \in \mathcal{A}$. The set of all the subgradients $r \in \mathbb{R}^V$ of the submodular function μ at M is called the subdifferential $\partial_\mu(M)$ of μ at M , i.e.,

$$\partial_\mu(M) := \{r \in \mathbb{R}^V \mid \mu(T) - \mu(M) \geq r(T) - r(M), \forall T \in \mathcal{A}\}. \quad (7)$$

The above polyhedra can be defined for any set functions (submodularity is not necessary) and also refer to this as the discrete subdifferential $\partial_\mu(M) \neq \phi$.

Let us define the conjugate function and conjugate set with respect to a submodular set function μ to generate the concept of Lagrangian of μ .

Definition 3.2. (Conjugate function)[12] Let $\mu : \mathcal{A} \rightarrow \mathbb{R}$ be a submodular function on $\mathcal{A} \subseteq 2^V$. The conjugate function $\mu^* : \mathbb{R}^V \rightarrow \mathbb{R}$ of μ is defined as

$$\mu^*(r) = \max_{M \in \mathcal{A}} \{r(M) - \mu(M) \mid r \in \mathbb{R}^V\}. \quad (8)$$

Note that $\phi, V \in \mathcal{A}$ and it is not necessary that $\mu(\phi) = 0$.

Now we can define the conjugate set $\mathcal{A}^* \subseteq \mathbb{R}^V$ of $\mathcal{A}$ for a submodular function μ by

$$\mathcal{A}^* = \{r \in \mathbb{R}^V \mid \max_{M \in \mathcal{A}} [r(M) - \mu(M)] < \infty\}. \quad (9)$$

Therefore, the submodular set function μ can be constructed in a new manner for every $M \in \mathcal{A}$ as

$$\mu(M) = \max_{r \in \mathcal{A}^*} \{r(M) - \mu^*(r)\}, \quad (10)$$

where μ^* is a convex conjugate function of μ .

Definition 3.3. A subgradient $M \in \mathcal{A}$ of a conjugate function $\mu^* : \mathcal{A}^* \rightarrow \mathbb{R}$ of submodular set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ at a point $r^* \in \mathcal{A}^*$ is defined by

$$\mu^*(r) \geq \mu^*(r^*) + (r - r^*)(M), \forall r \in \mathcal{A}^* \quad (11)$$

the set of all the subgradients $M \in \mathcal{A}$ of the conjugate function μ^* at a point $r^* \in \mathcal{A}^*$ is called the subdifferential $\partial_\mu^*(r^*)$ of μ^* at r^* , i.e.,

$$\partial_\mu^*(r^*) = \{M \in \mathcal{A} \mid \mu^*(r) \geq \mu^*(r^*) + (r - r^*)(M), \forall r \in \mathcal{A}^*\}. \quad (12)$$

From equation (8), the Young's type inequality for the submodular set function μ is defined as

$$\mu^*(r) + \mu(M) \geq r(M), \quad (13)$$

for any $M \in \mathcal{A}$ and $r \in \mathcal{A}^*$.

The following proposition will provide us information about the equality of the inequality (13).

Proposition 3.1. Suppose $\mu : \mathcal{A} \rightarrow \mathbb{R}$ is a submodular set function and $\mu^* : \mathcal{A}^* \rightarrow \mathbb{R}$ is the conjugate function of μ then,

- (i) $r \in \partial_\mu(M^*)$ iff $\mu^*(r) + \mu(M^*) = r(M^*)$, for $M^* \in \mathcal{A}$,
- (ii) $M \in \partial_\mu^*(r^*)$ iff $\mu^*(r^*) + \mu(M) = r^*(M)$, for $r^* \in \mathcal{A}^*$.

Proof: (i) For any $M \in \mathcal{A}$ and $r \in \partial_\mu(M^*)$, from (7)

$$\begin{aligned} \mu(M) - \mu(M^*) &\geq r(M) - r(M^*), \\ \Rightarrow r(M^*) &\geq \mu(M^*) + r(M) - \mu(M), \forall M \in \mathcal{A} \\ \Rightarrow r(M^*) &\geq \mu(M^*) + \max_{M \in \mathcal{A}} [r(M) - \mu(M)], \\ &\Rightarrow r(M^*) \geq \mu(M^*) + \mu^*(r). \end{aligned}$$

Now using Young's type inequality (13) for $M^* \in \mathcal{A}$ and $r \in \mathcal{A}^*$, we get

$$\mu(M^*) + \mu^*(r) \geq r(M^*),$$

therefore, we have

$$\mu(M^*) + \mu^*(r) = r(M^*).$$

Conversely, let $\mu(M^*) + \mu^*(r) = r(M^*)$ then by the definition of conjugate function μ^* of a submodular set function μ , we obtain

$$\begin{aligned} r(M^*) &\geq \mu(M^*) + [r(M) - \mu(M)], \\ \Rightarrow \mu(M) - \mu(M^*) &\geq r(M) - r(M^*), \forall M \in \mathcal{A}, \\ &\Rightarrow r \in \partial_\mu(M^*). \end{aligned}$$

(ii) For $M \in \partial\mu^*(r^*)$, we have

$$\begin{aligned} \mu^*(r) &\geq \mu^*(r^*) + (r - r^*)(M), \forall r \in \mathcal{A}^* \\ \Rightarrow r^*(M) &\geq \mu^*(r^*) + r(M) - \mu^*(r), \forall r \in \mathcal{A}^*, \\ \Rightarrow r^*(M) &\geq \mu^*(r^*) + \max_{r \in \mathcal{A}^*} [r(M) - \mu^*(r)], \\ &\Rightarrow r^*(M) \geq \mu^*(r^*) + \mu(M). \end{aligned} \tag{14}$$

By using (14) and inequality (13) for $M \in \mathcal{A}$ and $r^* \in \mathcal{A}^*$, we get $\mu(M) + \mu^*(r^*) = r^*(M)$. Conversely, if $\mu(M) + \mu^*(r^*) = r^*(M)$ then

$$\begin{aligned} \mu(M) + \mu^*(r^*) &= r^*(M) \\ \Rightarrow r^*(M) &= \mu(M) + \mu^*(r^*), \\ \Rightarrow r^*(M^*) &\geq [r(M) - \mu^*(r)] + \mu^*(r^*), \\ \Rightarrow \mu^*(M^*) &\geq \mu^*(r^*) + (r - r^*)(M), \forall r \in \mathcal{A}^*. \end{aligned}$$

Hence, $M \in \partial\mu^*(r^*)$. □

Theorem 3.1. [12] Suppose $\mu_1 : \mathcal{A}_1 \rightarrow \mathbb{R}$ and $\mu_2 : \mathcal{A}_2 \rightarrow \mathbb{R}$ are two submodular set functions, where $\mathcal{A}_1, \mathcal{A}_2 \in 2^V$ with $\phi, V \in \mathcal{A}_1 \cap \mathcal{A}_2$. Then for each $M \in \mathcal{A}_1 \cap \mathcal{A}_2$, we have

$$(i) \partial_{(\mu_1 + \mu_2)}(M) = \partial_{\mu_1}(M) + \partial_{\mu_2}(M), \tag{15}$$

$$(ii) \partial_{(\lambda\mu_1)}(M) = \lambda\partial_{\mu_1}(M), \forall \lambda > 0, M \in \mathcal{A}_1, \tag{16}$$

$$(iii) \partial_{(0, \mu_1)}(M) = \{r | r \in \mathbb{R}^V, \forall N \in \mathcal{A}_1, r(N) - r(M) \leq 0\} = 0^+ \partial_{\mu_1}(M), \tag{17}$$

where $0^+ \partial_{\mu_1}(M)$ is the characteristic cone of $\partial_{\mu_1}(M)$. □

The following Corollary 3.1 is an immediate outcome of the Theorem 3.1.

Corollary 3.1. Let $\mu_i : \mathcal{A}_i \rightarrow \mathbb{R}, i = 1, 2, \dots, k$, be the submodular set functions, where $\mathcal{A}_i \in 2^V, i = 1, 2, \dots, k$, with $\phi, V \in \cap_{i=1}^k \mathcal{A}_i$, then for every $M \in \cap_{i=1}^k \mathcal{A}_i$, we get

$$\partial_{(\mu_1 + \mu_2 + \dots + \mu_k)}(M) = \partial_{\mu_1}(M) + \partial_{\mu_2}(M) + \dots + \partial_{\mu_k}(M). \tag{18}$$

4. OPTIMALITY CONDITIONS FOR SUBMODULAR MINIMIZATION

In this section, we have considered the optimality conditions for both unconstrained and constrained submodular minimization problems. That means the involved objective function and constraint functions in the minimizing programs are all submodular.

4.1. Unconstrained submodular minimization problems. Let us consider the minimization problem

$$\min_{M \in \mathcal{A}} \mu(M), \quad (19)$$

where $\mathcal{A} \subseteq 2^V$ and $\mu : \mathcal{A} \rightarrow \mathbb{R}$.

Define the Lagrangian function $L(M; r) : \mathcal{A} \times \mathbb{R}^V \rightarrow \mathbb{R}$ by

$$L(M; r) = r(M) - \mu^*(r). \quad (20)$$

Now it can be easily prove that the well known saddle point property holds for $L(M; r)$.

Theorem 4.1. If $M^* \in \mathcal{A}$ and $r^* \in \mathcal{A}^* \subseteq \mathbb{R}^V$, then the following two statements are equivalent,

(i) M^* minimizes the program (19), that is

$$\mu(M^*) = \min_{M \in \mathcal{A}} L(M; r^*),$$

(ii) (M^*, r^*) is a saddle point of the Lagrangian $L(M; r)$, so that

$$L(M^*; r) \leq L(M^*; r^*) \leq L(M; r^*), \forall r \in \mathcal{A}^*, M \in \mathcal{A}.$$

Therefore $L(M^*; r^*) = \mu(M^*)$.

Proof: Let us assume (i) holds, then by the definition (20), we have

$$\begin{aligned} L(M^*; r^*) &= r^*(M^*) - \mu^*(r^*) \\ &\leq \max_{r \in \mathcal{A}^*} [r(M^*) - \mu^*(r)] \\ &= \mu(M^*) \\ &= \min_{M \in \mathcal{A}} L(M; r^*) \\ &\leq L(M; r^*) \forall M \in \mathcal{A}. \end{aligned}$$

Since, we have $\mu(M^*) = \min_{M \in \mathcal{A}} L(M; r^*) \leq r^*(M^*) - \mu^*(r^*)$ this implies,

$$\mu(M^*) + \mu^*(r^*) \leq r^*(M^*).$$

Using Young's type inequality (13) for $M^* \in \mathcal{A}$ and $r^* \in \mathcal{A}^*$, we get

$$\mu(M^*) + \mu^*(r^*) = r^*(M^*).$$

$\Rightarrow M^* \in \partial_{\mu^*}(r^*)$. by (ii) of Proposition 3.1.

This implies

$$\begin{aligned} \mu^*(r) &\geq \mu^*(r^*) + (r - r^*)(M^*), \forall r \in \mathcal{A}^*, \\ \Rightarrow r^*(M^*) - \mu^*(r^*) &\geq r(M^*) - \mu^*(r), \forall r \in \mathcal{A}^*, \\ \Rightarrow L(M^*; r^*) &\geq L(M^*; r). \end{aligned}$$

Therefore, (M^*, r^*) is a saddle point of the Lagrangian $L(M; r)$, so that

$$L(M^*; r) \leq L(M^*; r^*) \leq L(M; r^*), \forall r \in \mathcal{A}^*, M \in \mathcal{A}.$$

Thus we have proved (i) $\Rightarrow$ (ii).

To prove (ii) $\Rightarrow$ (i), suppose (M^*, r^*) is a saddle point of the Lagrangian $L(M; r)$, then

$$\mu(M) = \max_{r \in \mathcal{A}^*} [r(M) - \mu^*(r)] \geq r^*(M) - \mu^*(r^*) = L(M; r^*) \geq L(M^*; r^*) \geq L(M^*; r),$$

$$\Rightarrow \mu(M) \geq r(M^*) - \mu^*(r) \geq \max_{r \in \mathcal{A}^*} [r(M^*) - \mu^*(r)] = \mu(M^*), \forall M \in \mathcal{A}.$$

Therefore, M^* minimizes the program (19). Since $(M^*; r^*)$ is a saddle point, we obtain

$$\min_{M \in \mathcal{A}} L(M; r^*) \geq L(M^*; r^*) \geq \max_{r \in \mathcal{A}^*} L(M^*; r) = \mu(M^*).$$

Hence, $\min_{M \in \mathcal{A}} L(M; r^*) \geq \mu(M^*)$. Again, we have

$$\mu(M^*) \geq r^*(M^*) - \mu^*(r^*) = L(M^*; r^*) \geq \min_{M \in \mathcal{A}} L(M; r^*).$$

Thus

$$\mu(M^*) \geq \min_{M \in \mathcal{A}} L(M; r^*).$$

□

The optimality conditions for unconstrained submodular minimization programs is established in Theorem 4.2.

Theorem 4.2. A set $M^* \in \mathcal{A}$ minimizes the set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ iff,

$$0 \in \partial_\mu(M^*). \quad (21)$$

Proof: Suppose $M^* \in \mathcal{A}$ minimizes the set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ then by Theorem 4.1, (M^*, r^*) is a saddle point of the Lagrangian $L(M; r)$, where $r^* \in \mathcal{A}^*$.

$$\text{Hence, } L(M^*; r^*) \leq L(M; r^*)$$

$$\Rightarrow r^*(M^*) - \mu^*(r^*) \leq r^*(M) - \mu^*(r^*)$$

$$\Rightarrow \max_{r^* \in \mathcal{A}^*} [r^*(M^*) - \mu^*(r^*)] \leq \max_{r^* \in \mathcal{A}^*} [r^*(M) - \mu^*(r^*)]$$

$$\Rightarrow \mu(M) - \mu(M^*) \geq 0, \forall M \in \mathcal{A} \text{ by equation (10)}$$

$$\Rightarrow 0 \in \partial_\mu(M^*).$$

Conversely, if $0 \in \partial_\mu(M^*)$ holds, then the Definition 7, the set $M^* \in \mathcal{A}$ is the minimizer of $\mu : \mathcal{A} \rightarrow \mathbb{R}$.

□

Lemma 4.1. [12] A set $M^* \in \mathcal{A}$ is a minimizer of a submodular set function $\mu : \mathcal{A} \rightarrow \mathbb{R}$ iff $\mu(M^*) \leq \mu(M)$, for every set M such that $M \subseteq M^*$ or $M^* \subseteq M$.

4.2. Constrained submodular minimization problems. In this section, we have considered the optimality conditions for constrained submodular minimization problems with equality and inequality constraints separately.

4.2.1. Submodular minimization problems with equality constraints. Let us consider a submodular minimization problem with equality constraints

$$\begin{aligned} & \min_{M \in \mathcal{A}} \mu(M) \\ & \text{subject to } \nu_j(M) = \alpha_j, \quad j = (1, 2, \dots, k) \end{aligned} \quad (22)$$

where $\mu, \nu_j : \mathcal{A} \rightarrow \mathbb{R}$, $j = (1, 2, \dots, k)$. Now define $\mathcal{A}_0 \subseteq \mathcal{A}$ as the constraint set such that

$$\mathcal{A}_0 = \{M \in \mathcal{A} \mid \nu_j(M) = \alpha_j, j = 1, 2, \dots, k\}.$$

Therefore, the minimization problem (22) will be $\min_{M \in \mathcal{A}_0} \mu(M)$.

Now, the Lagrangian is defined as

$$L(M; \lambda) = \mu(M) + \sum_{j=1}^k \lambda_j (\nu_j(M) - \alpha_j), \quad (23)$$

where $\lambda \in \mathbb{R}_+^k$ is said to be a Lagrangian multiplier if

$$\min_{M \in \mathcal{A}} L(M; \lambda) = \min_{M \in \mathcal{A}_0} \mu(M). \quad (24)$$

Theorem 4.3. A set $M^* \in \mathcal{A}_0$ minimizes μ on $\mathcal{A}_0$ and (M^*, λ^*) is a saddle point of the Lagrangian $L(M^*; \lambda^*)$ iff $0 \in \partial_\mu(M^*) + \sum_{j=1}^k \lambda_j^* \partial_{\nu_j}(M^*)$, where $\mu, \nu_j : \mathcal{A} \rightarrow \mathbb{R}$, $j = (1, 2, \dots, k)$ are submodular set functions and $\lambda^* \in \mathbb{R}_+^k$ is a Lagrangian multiplier and $\partial_\mu(M^*)$ is the subdifferential taken on $\mathcal{A}$.

Proof: Since the Lagrangian $L(M^*; \lambda^*)$ has the saddle point (M^*, λ^*) , we have

$$\max_{\lambda \in \mathbb{R}_+^k} L(M^*; \lambda) \leq L(M^*; \lambda^*) \leq \min_{M \in \mathcal{A}} L(M; \lambda^*).$$

Hence, $M^* \in \mathcal{A}_0$. So, $L(M^*; \lambda^*) \leq \min_{M \in \mathcal{A}} L(M; \lambda^*)$. Now, by the equation (24), we obtain $L(M^*; \lambda^*) \leq \min_{M \in \mathcal{A}_0} \mu(M)$. Therefore, M^* is a minimizer on $\mathcal{A}_0$ and λ^* is the Lagrangian multiplier. This implies $0 \in \partial_{[\mu + \sum_{j=1}^k \lambda_j^* \nu_j]}(M^*)$. By using (16) and (18), we get

$$0 \in \partial_\mu(M^*) + \sum_{j=1}^k \lambda_j^* \partial_{\nu_j}(M^*). \quad \square$$

4.2.2. Submodular minimization problems with inequality constraints. Now, we consider a submodular minimization problem with inequality constraints

$$\begin{aligned} & \text{(P)} \quad \min_{M \in \mathcal{A}} \mu(M) \\ & \text{subject to } \nu_j(M) \leq 0, \quad j = 1, 2, \dots, k \end{aligned} \quad (25)$$

where $\mu, \nu_j : \mathcal{A} \rightarrow \mathbb{R}$, $j = (1, 2, \dots, k)$. Now define $\mathcal{A}_0 \subseteq \mathcal{A}$ as the constraint set such that

$$\mathcal{A}_0 = \{M \in \mathcal{A} \mid \nu_j(M) \leq 0, j = 1, 2, \dots, k\}.$$

Therefore, the minimization problem (25) will be $\min_{M \in \mathcal{A}_0} \mu(M)$. Now, the Lagrangian is defined as

$$L(M; \lambda) = \mu(M) + \sum_{j=1}^k \lambda_j \nu_j(M), \quad (26)$$

where $\lambda \in \mathbb{R}_+^k$ is said to be a Lagrangian multiplier if

$$\min_{M \in \mathcal{A}} L(M; \lambda) = \min_{M \in \mathcal{A}_0} \mu(M). \quad (27)$$

The following Lemma 4.2 will give us the necessary condition for a constrained minimum.

Lemma 4.2. [20] Let $\mu, \nu_1, \nu_2, \dots, \nu_k$ be submodular set functions on $\mathcal{A}$ for the minimization problem (P). Assume the Slater's constraint qualification: $\exists$ a set $M^* \in \mathcal{A}$ s.t., $\nu_j(M^*) \leq 0, j = 1, 2, \dots, k$. If $M^* \in \mathcal{A}$ is an optimal solution of the problem (P) then $\exists \lambda^* \in \mathbb{R}_+^k$ such that $\sum_{j=1}^k \lambda_j^* \nu_j(M^*) = 0$. Also, (M^*, λ^*) is a saddle point of the Lagrangian function $L(M; \lambda) = \mu(M) + \sum_{j=1}^k \lambda_j \nu_j(M)$ so that

$$\mu(M^*) + \sum_{j=1}^k \lambda_j \nu_j(M^*) \leq \mu(M^*) + \sum_{j=1}^k \lambda_j^* \nu_j(M^*) \leq \mu(M) + \sum_{j=1}^k \lambda_j^* \nu_j(M), \quad (28)$$

for all $\lambda \in \mathbb{R}_+^k$ and $M \in \mathcal{A}$.

The following Theorem 4.4 is a Kuhn-Tucker type conditions for submodular minimization problems, which involves inequality constraints.

Theorem 4.4. (KKT type conditions for submodular minimization)

Let $\mu, \nu_1, \nu_2, \dots, \nu_k$ be submodular set functions on $\mathcal{A}$ in the minimization problem (P) and satisfy the Slater's constraint qualification given in Lemma (4.2). If $M^* \in \mathcal{A}$ be an optimal solution of the problem (P) then $\exists \lambda^* \in \mathbb{R}_+^k$ such that

$$(i) \sum_{j=1}^k \lambda_j^* \nu_j(M^*) = 0, \quad (29)$$

and

$$(ii) 0 \in \partial_\mu(M^*) + \sum_{j=1}^k \lambda_j^* \partial_{\nu_j}(M^*). \quad (30)$$

Proof: Let $M^* \in \mathcal{A}$ be an optimal solution of the submodular minimization problem (P). By Lemma (4.2), $\exists \lambda^* \in \mathbb{R}_+^k$ such that $\sum_{j=1}^k \lambda_j^* \nu_j(M^*) = 0$, and the Lagrangian $L(M; \lambda) = \mu(M) + \sum_{j=1}^k \lambda_j \nu_j(M)$ has a saddle point (M^*, λ^*) . Therefore,

$$\begin{aligned} \mu(M^*) + \sum_{j=1}^k \lambda_j^* \nu_j(M^*) &\leq \mu(M) + \sum_{j=1}^k \lambda_j^* \nu_j(M), \forall M \in \mathcal{A} \\ \Rightarrow \mu(M^*) + \sum_{j=1}^k \lambda_j^* \nu_j(M^*) &= \min_{M \in \mathcal{A}} [\mu(M) + \sum_{j=1}^k \lambda_j^* \nu_j(M)], \forall M \in \mathcal{A} \\ &\Rightarrow 0 \in \partial_{[\mu + \sum_{j=1}^k \lambda_j^* \nu_j]}(M^*). \end{aligned}$$

Now, by using (16) and Corollary 3.1, we get $0 \in \partial_\mu(M^*) + \sum_{j=1}^k \lambda_j^* \partial_{\nu_j}(M^*)$. $\square$

5. CONCLUSION

The optimality theory for submodular minimization programs has been extended and revealed the similarity between convex functions and submodular set functions. The Lagrangian formulations for unconstrained and constrained submodular minimization programming problems have been presented, and KKT optimality conditions have also been developed. In light of the present investigation, one can establish the duality theory for submodular programming problems.

STATEMENTS AND DECLARATIONS

The authors declare that they have no conflict of interest, and the manuscript has no associated data.

REFERENCES

- [1] F. Bach. Submodular functions: from discrete to continuous domains. *Mathematical Programming*, 175:419-459, 2019.
- [2] M.-F. Balcan and N. J. A. Harvey. Submodular functions: Learnability, structure, and optimization. *SIAM Journal on Computing*, 47(3):703-754, 2018.
- [3] E. Balkanski and Y. Singer. The adaptive complexity of maximizing a submodular function. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing*, pages 1138-1151, Los Angeles, USA, 2018. Association for Computing Machinery, New York.
- [4] F. Baumann, S. Berckey, and C. Buchheim. Exact algorithms for combinatorial optimization problems with submodular objective functions. In M. Jünger and G. Reinelt, editors, *Facets of Combinatorial Optimization*, pages 271-294. Springer, Berlin, Heidelberg, 2013.
- [5] K. Bérczi, A. Imolay, L. Lovász, and T. Schwarcz. Monotonic decompositions of submodular set functions. *SIAM Journal on Discrete Mathematics*, 40(1):308-335, 2026.
- [6] N. Buchbinder and M. Feldman. Constrained submodular maximization via new bounds for DR-submodular functions. In *Proceedings of the 56th Annual ACM Symposium on Theory of Computing*, pages 1820-1831, Vancouver, Canada, 2024. Association for Computing Machinery, New York.

- [7] V. Cevher and A. Krause. Greedy dictionary selection for sparse representation. *IEEE Journal of Selected Topics in Signal Processing*, 5(5):979-988, 2011.
- [8] D. Chakrabarty, A. Graur, H. Jiang, and A. Sidford. Parallel submodular function minimization. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, pages 68639-68654. Curran Associates, United States, 2023.
- [9] A. Clark, Q. Hou, L. Bushnell, and R. Poovendran. Maximizing the smallest eigenvalue of a symmetric matrix: A submodular optimization approach. *Automatica*, 95:446-454, 2018.
- [10] B. Doerr, C. Doerr, A. Neumann, F. Neumann, and A. M. Sutton. Optimization of chance-constrained submodular functions. In: *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence*, 34(2):1460-1467.
- [11] J. Edmonds. Submodular functions, matroids, and certain polyhedra. In R. Guy, H. Hanani, N. Sauer, and J. Schönheim, editors, *Combinatorial Structures and Their Applications*, pages 69-87. Gordon and Breach, New York, 1970.
- [12] S. Fujishige. *Submodular Functions and Optimization*, volume 58 of *Annals of Discrete Mathematics*. Elsevier, Amsterdam, 2nd edition, 2005.
- [13] S. Fujishige. Theory of submodular programs: a Fenchel-type min-max theorem and subgradients of submodular functions. *Mathematical Programming*, 29(2):142-155, 1984.
- [14] D. S. Hochbaum. Complexity and approximations for submodular minimization problems on two variables per inequality constraints. *Discrete Applied Mathematics*, 250:252-261, 2018.
- [15] S. Jegelka and J. Bilmes. Graph cuts with interacting edge weights: Examples, approximations, and algorithms. *Mathematical Programming*, 162:241-282, 2017.
- [16] P. Kohli, P. Kumar, and P. H. S. Torr. $\mathcal{P}^3$ and beyond: move making algorithms for solving higher order functions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(9):1645-1656, 2009.
- [17] A. Kulik, R. Schwartz, and H. Shachnai. A refined analysis of submodular Greedy. *Operations Research Letters*, 49:507-514, 2021.
- [18] H. Lin and J. Bilmes. A class of submodular functions for document summarization. In D. Lin, Y. Matsumoto, Rada Mihalcea, editors, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, pages 510-520, Portland, USA, 2011. Association for Computational Linguistics, United States.
- [19] Y. Liu, R. Iyer, K. Kirchhoff, and J. Bilmes. SVitchboard-II and FiSVer-I: Crafting high quality and low complexity conversational English speech corpora using submodular function optimization. *Computer Speech & Language*, 42:122-142, 2017.
- [20] R. J. T. Morris. Optimal constrained selection of a measurable subset. *Journal of Mathematical Analysis and Applications*, 70(2):546-562, 1979.
- [21] H. Oshima. Improved randomized algorithm for k -submodular function maximization. *SIAM Journal on Discrete Mathematics*, 35(1):1-22, 2021.
- [22] M. Pedramfar, C. Quinn, and V. Aggarwal. A unified approach for maximizing continuous DR-submodular functions. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, pages 61103-61114, 2023. Curran Associates, United States .
- [23] C. Qian, Y. Yu, K. Tang, X. Yao, and Z.-H. Zhou. Maximizing submodular or monotone approximately submodular functions by multi-objective evolutionary algorithms. *Artificial Intelligence*, 275:279-294, 2019.
- [24] P. Skowron. FPT approximation schemes for maximizing submodular functions. *Information and Computation*, 257:65-78, 2017.
- [25] A. Testa, F. Farina, and G. Notarstefano. Distributed submodular minimization via block-wise updates and communications. *IFAC-PapersOnLine*, 53(2):2678-2683, 2020.
- [26] R. Yang, S. Gu, C. Gao, W. Wu, H. Wang, and D. Xu. A constrained two-stage submodular maximization. *Theoretical Computer Science*, 853:57-64, 2021.
- [27] Q. Zhang and H. Zhong. A mutual information-driven submodular optimization approach to covariate selection in causal effect estimation. *Knowledge-Based Systems*, 314:Article ID 115808, 2026.
- [28] L. Zhao, H. Nagamochi, and T. Ibaraki. Greedy splitting algorithms for approximating multiway partition problems. *Mathematical Programming*, 102(1):167-183, 2005.