Lecture 9: submodular function minimization and chance-constrained programming

Dabeen Lee

Industrial and Systems Engineering, KAIST

2025 Winter Lecture Series on Combinatorial Optimization

January 16, 2025

Outline

- Submodular functions
- Submodular function minimization
- Chance-constrained programs

Submodular functions

- Let E be a set of elements. We say that a set function $f : 2^E \to \mathbb{R}$ is **submodular** if it satisfies

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

  for all $A, B \subseteq E$.
- An equivalent definition of submodularity for set functions is the notion of **diminishing marginal returns** property.
- That is, a set function $f : 2^E \to \mathbb{R}$ is submodular if and only if it satisfies

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

  for all $A \subseteq B \subseteq E$ and $e \notin B$.

Submodular functions

Examples

- **Linear function:** For any $w \in \mathbb{R}^{|E|}$, f with

$$f(S) = \sum_{e \in S} w_e$$

for $S \subseteq E$ is submodular.

- **Concave utility:** For any concave function $g : \mathbb{R}_+ \to \mathbb{R}$ and $w \in \mathbb{R}_+^{|E|}$, f with

$$f(S) = g(\sum_{e \in S} w_e)$$

for $S \subseteq E$ is a submodular function.

- **Coverage function:** Suppose that each element $e \in E$ corresponds to some area A_e. Then f with

$$f(S) = |\cup_{e \in S} A_e|$$

for $S \subseteq E$ is submodular.

Submodular functions

Examples

- **Success probability:** Let $p_e \in [0, 1]$ for $e \in E$. Then f with

$$f(S) = 1 - \prod_{e \in S} (1 - p_e)$$

  for $S \subseteq E$ is submodular.
- **Graph cuts:** Let $G = (V, E)$ be an undirected graph. Then f with

$$f(S) = |\delta(S)|$$

  for $S \subseteq V$ is submodular, where $\delta(S)$ is the set of edges crossing the partition $(S, V \setminus S)$ of the vertex set V.
- **Directed cuts:** Let $D = (N, A)$ be a directed graph. Then f with $f(S) = |\delta^+(S)|$ for $S \subseteq V$ is submodular, where $\delta^+(S)$ is the set of arcs leaving S.
- **Matroid rank functions:** Let $\mathcal{M} = (E, \mathcal{I})$ be a matroid. Then its rank function r given by $r(S) = \max\{|A| : A \in \mathcal{I}\}$ for $S \subseteq E$ is submodular.

Submodular functions

- As this wide range of examples suggests, submodular functions provide a useful framework for modeling discrete-valued decision variables.
- For utility, coverage, and success probability functions, the problem of **maximizing** a submodular function is relevant.
- For cut functions, submodular function **minimization** is relevant.
- As a first step, in this lecture, we consider the minimization problem.

Submodular function minimization

- Let us consider the problem of minimizing a submodular function.
- Given a submodular function $f : 2^E \to \mathbb{R}$ over the element set E, we consider
$$\text{minimize} \quad f(S) \quad \text{subject to} \quad S \subseteq E. \tag{1}$$
- Since f is a set function, we can interpret the function over the set of binary vectors $\{0,1\}^{|E|}$.
- Any $S \subseteq E$ can be represented by its characteristic vector $1_S \in \{0,1\}^{|E|}$ that takes 1 for the elements in S and 0 for the other elements.
- Similarly, any vector $z \in \{0,1\}^{|E|}$ corresponds to a subset $S_z = \{e \in E : z_e = 1\}$.
- Then, with a slight abuse of notation, we may define
$$f(z) := f(S_z).$$
- Then the problem can be rewritten as the following binary optimization problem:
$$\text{minimize} \quad f(z) \quad \text{subject to} \quad z \in \{0,1\}^{|E|}. \tag{2}$$

Submodular function minimization

- Note that with an auxiliary variable y to make the objective linear, (2) is equivalent to

$$\text{minimize} \quad y \quad \text{subject to} \quad (y, z) \in Q_f \tag{3}$$

where Q_f is the **epigraph** of f given by

$$Q_f = \left\{(y, z) \in \mathbb{R} \times \{0, 1\}^{|E|} : \; y \geq f(z)\right\}.$$

- Since y is a linear function, it follows that (3) is equivalent to

$$\text{minimize} \quad y \quad \text{subject to} \quad (y, z) \in \mathrm{conv}(Q_f) \tag{4}$$

where $\mathrm{conv}(Q_f)$ is the convex hull of Q_f.

- By the equivalence between optimization and separation, the optimization problem (4) is equivalent to separation over $\mathrm{conv}(Q_f)$.

Submodular function minimization

- Next we will characterize the convex hull of Q_f and provide a linear description of it.
- To do so, we need to define the **extended polymatroid** of f, given by

$$EP_f := \left\{ \pi \in \mathbb{R}^{|E|} : \sum_{e \in S} \pi_e \leq f(S) \quad \text{for all } S \subseteq E \right\}.$$

- Note that the extended polymatroid is nonempty if and only if $f(\emptyset) \geq 0$.
- In general, a submodular function f does not have to satisfy $f(\emptyset) \geq 0$.
- Nevertheless, we may take $f - f(\emptyset)$, instead of f, which is submodular if f is submodular.
- Henceforth, we assume that $f(\emptyset) = 0$.

Submodular function minimization

- Having defined the extended polymatroid, we are ready to characterize the convex hull of Q_f.

Theorem (Edmonds, Lovász)

Let $f : \{0,1\}^{|E|} \to \mathbb{R}$ be a submodular function with $f(\emptyset) = 0$, and let Q_f be its epigraph. Then

$$\mathrm{conv}(Q_f) = \left\{ (y, z) \in \mathbb{R} \times [0,1]^{|E|} : \; y \geq \pi^\top z \quad \textit{for all } \pi \in EP_f \right\}.$$

- Given $(y, z) \in \mathbb{R} \times [0,1]^{|E|}$, deciding whether $(y, z) \in \mathrm{conv}(Q_f)$ boils down to computing the maximum value of $z^\top \pi$ over all $\pi \in EP_f$.

Submodular function minimization

- Edmonds proved that there is a greedy algorithm for computing the maximum of a linear function over the extended polymatroid EP_f.

Theorem (Edmonds)

Let $z \in \mathbb{R}^{|E|}$. Then the linear program

$$\max \left\{ \sum_{e \in E} z_e \pi_e : \; \pi \in EP_f \right\} \tag{P}$$

can be solved in $O(|E| \log |E|)$ time by a greedy algorithm.

Submodular function minimization

- Recall that the equivalence of optimization and separation is based on the ellipsoid method.
- Grötschel, Lovász, and Schrijver showed that the algorithm can be turned into a strongly polynomial time algorithm.

Theorem (Grötschel, Lovász, and Schrijver)

Let $f : 2^E \to \mathbb{R}$ be submodular over the element set E. Then one can find $S \subseteq E$ minimizing f in strongly polynomial time.

- Later, Iwata, Fleischer, and Fujishige and Schrijver independently provided combinatorial algorithms for submodular function minimization.

Inventory planning

- We consider an inventory planning problem.
- A retail store prepares some inventory of items before the market opens.
- The retail store can observe the actual **demand** after the market opens.

Inventory planning

- y: the amount of items that the retail store prepares before the market opens.
- h: the unit cost of preparing items before the market opens.
- b: the stochastic **demand** for items.

Inventory planning

Assumption

- There are n possibilities, given by $b_1, \dots, b_n$, for the stochastic demand b.
- Historically, the demand is equal to value b_i with probability p_i, i.e.,

$$\mathbb{P}[b = b_i] = p_i.$$

- Here, $p_1, \dots, p_n \geq 0$ and $\sum_{i=1}^{n} p_i = 1$.
- We assume that the probability distribution is known to the decision-maker.

Each case of demand realization is called a **scenario**

Assumption

- There are n possibilities, given by $b_1, \ldots, b_n$, for the stochastic demand b.
- Historically, the demand is equal to value b_i with probability p_i, i.e.,

$$\mathbb{P}[b = b_i] = p_i.$$

- Here, $p_1, \ldots, p_n \geq 0$ and $\sum_{i=1}^n p_i = 1$.
- We assume that the probability distribution is known to the decision-maker.

Each case of demand realization is called a **scenario**

Chance-constrained programs

- The purchase decision is based on the distribution of the stochastic demand.
- The first attempt is to prepare again all possible scenarios.
- Basically, we target the largest possible demand by solving

$$
\begin{aligned}
\min \quad & hy \\
\text{s.t.} \quad & y \geq b_i, \quad i = 1, \dots, n, \\
& y \in \mathbb{R}_+.
\end{aligned}
$$

- However, targeting the largest possible demand may be a too conservative decision.
- Maybe the largest possible demand value occurs with probability less than 0.1% while we would face a moderate demand level with proability in most cases.

Chance-constrained programs

- Let us consider

$$\begin{aligned} \min \quad & hy \\ \text{s.t.} \quad & \mathbb{P}\left[y \geq b\right] \geq 0.95 \\ & y \in \mathbb{R}_+. \end{aligned}$$

- This optimization model is called a **chance-constrained program**
- Note that the constraint requires that we satisfy the stochastic demand with at least 95% chance.
- We might not satisfy the demand in some cases, but as long as the failure probability is at most 5%, we hare happy.

Chance-constrained programs

- In fact, the chance-constrained program can be reformulated as an integer program.
- Note that

$$\mathbb{P}[y \geq b] \geq 0.95$$

  is equivalent to

$$\mathbb{P}[y < b] \leq 0.05.$$

- Moreover,

$$\mathbb{P}[y < b] = \sum_{i=1}^{n} p_i \cdot 1[y < b_i]$$

  where

$$1[y < b_i] = \begin{cases} 1, & \text{if } y < b_i, \\ 0, & \text{otherwise.} \end{cases}$$

Chance-constrained programs

- Let
$$z_i = \begin{cases} 0, & \text{if the demand for scenario } i \text{ is satisfied,} \\ 1, & \text{otherwise.} \end{cases}$$
- Basically, we use the binary variable z_i to model the indicator function $1[y < b_i]$.
- Then the chance-constrained program can be reformulated as the following integer program.

$$\begin{aligned} \min \quad & hy \\ \text{s.t.} \quad & y + b_i z_i \geq b_i, \quad i = 1, \dots, n, \\ & \sum_{i=1}^{n} p_i z_i \leq 0.05, \\ & y \in \mathbb{R}_+, z \in \{0, 1\}^n. \end{aligned}$$

Binary mixing set

- The solution set of this model

$$\{(y,z)\in\mathbb{R}\times\{0,1\}^n:\ y+b_iz_i\geq b_i,\quad i=1,\ldots,n\}$$

  is called the **binary mixing set**.
- The convex hull of the mixing set is also well-understood.
- Let us define a function $f:\{0,1\}^n\to\mathbb{R}$ as

$$f(z)=\max\{b_i(1-z_i): i\in\{1,\ldots,n\}\}.$$

- Note that the binary mixing set can be equivalently written as

$$Q_f=\{(y,z)\in\mathbb{R}\times\{0,1\}^n:\ y\geq f(z)\}.$$

Binary mixing set

Lemma

The function $f : \{0,1\}^n \to \mathbb{R}$ *with* $f(z) = \max\{b_i(1 - z_i) : i \in \{1, \dots, n\}\}$ *is submodular.*

Binary mixing set

Based on Lemma 4, we may deduce the following approach to solve the chance-constrained program.

1. We solve the LP relaxation of the integer programming formulation.
2. If the current solution $(y, z) \notin \mathrm{conv}(Q_f)$, then we separate an inequality based on the greedy algorithm of Theorem 2.