1 Outline

In this lecture, we cosider

- convexity of Lovász extension
- submodular function minimization

2 Convexity of Lovász Extension

In the previous lecture, we have introduced the Lovász extension $\hat{f}$ of a set function $f : \{0,1\}^{|E|} \to \mathbb{R}$, which is defined as

$$\hat{f}(x) = \sum_{i=0}^{|E|} \lambda_i f(\mathbf{1}_{S_i})$$

where $\emptyset = S_0 \subset S_1 \subset \cdots \subset S_{|E|} = E$ and $\lambda_i \geq 0$ with $\sum_{i=0}^{|E|} \lambda_i = 1$ such that

$$x = \sum_{i=0}^{|E|} \lambda_i \mathbf{1}_{S_i}.$$

Equivalently, we can express the Lovász extension as

$$\hat{f}(x) = \sum_{i=0}^{|E|} (x_{\pi(i)} - x_{\pi(i+1)}) f(\mathbf{1}_{S_i})$$

for $x \in [0,1]^{|E|}$, where $\pi : \{1,2,\ldots,|E|\} \to \{1,2,\ldots,|E|\}$ is a permutation such that

$$x_{\pi(1)} \geq x_{\pi(2)} \geq \cdots \geq x_{\pi(|E|)},$$

$S_0 = \emptyset$, and $S_i = \{\pi(1), \pi(2), \ldots, \pi(i)\}$ for $i = 1,2,\ldots,|E|$. In this section, we discuss the convexity of the Lovász extension.

It can be shown that if f is submodular, then the Lovász extension $\hat{f}$ is indeed a convex function. To elaborate on this, we take a convex lower envelope of f. To be specific, the lower envelope is the largest convex function that is lower than or equal to f on the vertices of the hypercube $[0,1]^{|E|}$:

$$f^-(x) = \min\left\{ \sum_{S\subseteq E} \lambda_S f(\mathbf{1}_S) : \sum_{S\subseteq E} \lambda_S \mathbf{1}_S = x,\ \sum_{S\subseteq E} \lambda_S = 1,\ \lambda_S \geq 0 \text{ for } S \subseteq E \right\}.$$

Let us first show that the lower envelope agrees with f on $\{0,1\}^{|E|}$.

Lemma 17.1. *For any $f : \{0,1\}^{|E|} \to \mathbb{R}$ and $S \subseteq E$, we have $f^-(\mathbf{1}_S) = f(\mathbf{1}_S)$.*

Proof. Let $\bar{S} \subseteq E$. Then

$$\sum_{S \subseteq E} \lambda_S \mathbf{1}_S = \mathbf{1}_{\bar{S}}, \ \sum_{S \subseteq E} \lambda_S = 1, \ \lambda_S \geq 0 \text{ for } S \subseteq E$$

holds if and only if $\lambda_{\bar{S}} = 1$ and $\lambda_T = 0$ for all $T \neq \bar{S}$. Therefore, we have $f^-(\mathbf{1}_{\bar{S}}) = f(\mathbf{1}_{\bar{S}})$, as required. □

Moreover, the lower envelope is a convex function.

Lemma 17.2. *For any $f : \{0,1\}^{|E|} \to \mathbb{R}$, the lower envelope f^- is convex.*

Proof. Let $x, y \in [0,1]^{|E|}$ and $\alpha \in [0,1]$. Take $\{\lambda_S\}_{S \subseteq E}$ and $\{\mu_S\}_{S \subseteq E}$ such that

$$f^-(x) = \sum_{S \subseteq E} \lambda_S f(\mathbf{1}_S), \quad \sum_{S \subseteq E} \lambda_S \mathbf{1}_S = x, \quad \sum_{S \subseteq E} \lambda_S = 1, \quad \lambda_S \geq 0 \text{ for } S \subseteq E,$$

and

$$f^-(y) = \sum_{S \subseteq E} \mu_S f(\mathbf{1}_S), \quad \sum_{S \subseteq E} \mu_S \mathbf{1}_S = y, \quad \sum_{S \subseteq E} \mu_S = 1, \quad \mu_S \geq 0 \text{ for } S \subseteq E.$$

Then, we have

$$\begin{aligned} f^-(\alpha x + (1-\alpha) y) &\leq \sum_{S \subseteq E} (\alpha \lambda_S + (1-\alpha)\mu_S) f(\mathbf{1}_S) \\ &= \alpha \sum_{S \subseteq E} \lambda_S f(\mathbf{1}_S) + (1-\alpha) \sum_{S \subseteq E} \mu_S f(\mathbf{1}_S) \\ &= \alpha f^-(x) + (1-\alpha) f^-(y), \end{aligned}$$

implying the convexity of f^-. □

To show that the Lovász extension $\hat{f}$ is a convex function when f is submodular, we prove that $\hat{f}$ coincides with the lower envelope f^- if and only if f is submodular.

Lemma 17.3. *Let $f : \{0,1\}^{|E|} \to \mathbb{R}$. Then, $\hat{f}(x) = f^-(x)$ for all $x \in [0,1]^{|E|}$ if and only if f is submodular.*

Proof. ($\Leftarrow$) Suppose that f is submodular. Recall that

$$f^-(x) = \min \left\{ \sum_{S \subseteq E} \lambda_S f(\mathbf{1}_S) : \sum_{S \subseteq E} \lambda_S \mathbf{1}_S = x, \ \sum_{S \subseteq E} \lambda_S = 1, \ \lambda_S \geq 0 \text{ for } S \subseteq E \right\}.$$

Among the coefficient vector λ that achieves the minimum, we choose the one maximizing

$$\sum_{S \subseteq E} \lambda_S |S|^2.$$

Suppose that there exist two sets $A, B \subseteq E$ such that $\lambda_A \geq \lambda_B > 0$, and neither $A \subseteq B$ nor $B \subseteq A$. Then we update λ by

$$\lambda_A \leftarrow \lambda_A - \lambda_B, \quad \lambda_B \leftarrow 0, \quad \lambda_{A \cup B} \leftarrow \lambda_{A \cup B} + \lambda_B, \quad \lambda_{A \cap B} \leftarrow \lambda_{A \cap B} + \lambda_B.$$

By the submodularity of f, this update does not increase the objective value. However, this update increases $\sum_{S\subseteq E} \lambda_S |S|^2$, because

$$|A\cup B|^2 + |A\cap B|^2 = |A|^2 + |B|^2 + 2|B\setminus A||A\setminus B| > |A|^2 + |B|^2.$$

Therefore, the support of λ forms a chain, which implies that $\hat{f}(x) = f^-(x)$.

($\Rightarrow$) Suppose that f is not submodular. Then, there exist $A, B \subseteq E$ such that

$$f(A) + f(B) < f(A\cup B) + f(A\cap B).$$

Take $x = \frac{1}{2}\mathbf{1}_A + \frac{1}{2}\mathbf{1}_B$. Then, we have

$$x = \frac{1}{2}\mathbf{1}_{A\cup B} + \frac{1}{2}\mathbf{1}_{A\cap B}.$$

Then the Lovász extension at x is given by

$$\hat{f}(x) = \frac{1}{2} f(\mathbf{1}_{A\cup B}) + \frac{1}{2} f(\mathbf{1}_{A\cap B}).$$

On the other hand, the lower envelope at x satisfies that

$$f^-(x) \le \frac{1}{2} f(\mathbf{1}_A) + \frac{1}{2} f(\mathbf{1}_B) < \frac{1}{2} f(\mathbf{1}_{A\cup B}) + \frac{1}{2} f(\mathbf{1}_{A\cap B}) = \hat{f}(x),$$

implying that $\hat{f}(x) \neq f^-(x)$. □

2.1 Submodular Function Minimization via Convex Optimization

We have just argued that the Lovász extension $\hat{f}$ of a submodular function f is a convex function. This fact allows us to reformulate the submodular function minimization problem

$$\text{minimize} \quad f(z) \quad \text{subject to} \quad z \in \{0,1\}^{|E|} \tag{17.1}$$

as a convex optimization problem, given by

$$\text{minimize} \quad \hat{f}(x) \quad \text{subject to} \quad x \in [0,1]^{|E|}. \tag{17.2}$$

This means that we can apply algorithms for convex optimization to solve (17.1). For example, we can use the projected subgradient method to solve (17.2). Note that the projection onto the hypercube $[0,1]^{|E|}$ is simply given by component-wise clipping. The remaining issue is how to compute a subgradient of the Lovász extension $\hat{f}$. What would be a subgradient of $\hat{x}$? To compute a subgradient of $\hat{f}$ at $x \in [0,1]^{|E|}$, we sort the components of x in non-increasing order with a permutation $\pi : \{1, 2, \ldots, |E|\} \to \{1, 2, \ldots, |E|\}$ such that

$$x_{\pi(1)} \ge x_{\pi(2)} \ge \cdots \ge x_{\pi(|E|)}.$$

Then, we define $S_0 = \emptyset$ and $S_i = \{\pi(1), \pi(2), \ldots, \pi(i)\}$ for $i = 1, 2, \ldots, |E|$ and set $x_{\pi(0)} = 1$ and $x_{\pi(|E|+1)} = 0$. With these notations, note that $\hat{f}$ can be expressed As

$$\hat{f}(x) = \sum_{i=0}^{|E|} (x_{\pi(i)} - x_{\pi(i+1)}) f(\mathbf{1}_{S_i}) = f(\emptyset) + \sum_{i=1}^{|E|} x_{\pi(i)} \left(f(\mathbf{1}_{S_i}) - f(\mathbf{1}_{S_{i-1}}) \right).$$

Based on this expression, the vector $g \in \mathbb{R}^{|E|}$ defined as

$$g_{\pi(i)} = f(\mathbf{1}_{S_i}) - f(\mathbf{1}_{S_{i-1}}) \quad \text{for } i = 1, 2, \dots, |E|$$

is a subgradient of $\hat{f}$ at x. Note that g can be computed by evaluating f at most $|E|$ times. Therefore, if we have an efficient oracle to evaluate f, then we can efficiently compute a subgradient of $\hat{f}$ and thus solve (17.2) using the projected subgradient method.

Suppose that we have obtained an optimal solution x^* of (17.2). How can we recover an optimal solution of the original problem (17.1)? Since (17.2) is equivalent to (17.1), it follows that $f(\mathbf{1}_S) \geq \hat{f}(x^*)$ for all $S \subseteq E$. On the other hand, we have

$$\hat{f}(x^*) = \sum_{i=0}^{|E|} (x^*_{\pi(i)} - x^*_{\pi(i+1)}) f(\mathbf{1}_{S_i^*}) \geq \min_{i=0,1,\dots,|E|} f(\mathbf{1}_{S_i^*})$$

where π and S_i^* are defined in the same way as before but with x^* in place of x. This implies that any of $f(\mathbf{1}_{S_i^*})$ for $i = 0, 1, \dots, |E|$ is an optimal value of (17.1).