Course Notes for EE227C (Spring 2018): Convex Optimization and Approximation

Instructor: Moritz Hardt
Email: hardt+ee227c@berkeley.edu

Graduate Instructor: Max Simchowitz
Email: msimchow+ee227c@berkeley.edu

May 6, 2018

26 Submodular functions and Lovász extension

Definition 26.1 (Submodular functions)**.** Let $N = \{1, \dots, n\}$, and 2^N denote the power set of N. Then a function $f : 2^N \to \mathbb{R}$ is *submodular* if:

$$f(A \cup \{j\}) - f(A) \geqslant f(B \cup \{j\}) - f(B)$$

for all $A \subseteq B \subseteq N$ and for all $j \in N$.

This is equivalent to:

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

for all $\forall A, B \subseteq N$.

Some examples of submodular functions:

- Let A be a subset of nodes of a bipartite graph. Then $f(A) = |\text{neighborhood}(A)|$ is submodular.
- Let $X = \{X_1, \dots X_n\}$ be a set of random variables, and $A \subseteq X$. Then $f(A) = H(A)$, where H denotes entropy, is submodular.
- Let E, V be the edges and nodes of a graph, and $A \subseteq V$. Then the size of the graph cut, $f(A) = \big|\{(u,v) \in E : u \in A, v \in V \setminus A\}\big|$, is submodular.

There is an equivalence between submodular and boolean functions: a submodular function $f : 2^N \to \mathbb{R}$ can be expressed as a boolean function $f : \{0,1\}^n \to \mathbb{R}$.

Definition 26.2 (Lovász extension)**.** Let $f(x)$ be a submodular function. Then its Lovász extension, $\widehat{f}(z) : [0,1]^n \to \mathbb{R}$ is given by:

$$\widehat{f}(z) = \mathbb{E}_{\lambda \sim \mathrm{U}(0,1)} \left[f(\{i : z_i \geqslant \lambda\}) \right]$$

where $\mathrm{U}(0,1)$ denotes the uniform distribution on $[0,1]$.

Some observations about $\widehat{f}(z)$:

- For any $x \in \{0,1\}^n$, $\widehat{f}(x) = f(x)$. That is, f and $\widehat{f}$ agree on the domain of f.
- For any $z \in [0,1]^n$, there exists an $x \in \{0,1\}^n$, such that $f(x) \leqslant \widehat{f}(z)$. Also, $\min_{x \in \{0,1\}^n} f(x) = \min_{z \in [0,1]^n} \widehat{f}(z)$.

Theorem 26.3 (Submodularity and convexity)**.** *Consider a function $f(x) : 2^N \to \mathbb{R}$ and its Lovász extension $\widehat{f}(z) : 2^N \to \mathbb{R}$. Then $f(x)$ is submodular iff $\widehat{f}(z)$ is convex.*

Proof. We only show the forward direction (f is submodular $\Rightarrow \widehat{f}$ is convex), the reverse direction is left as an exercise.

Without loss of generality, assume z is order: $z_1 \geqslant z_2 \geqslant \ldots \geqslant z_n$.
Also, let $S_i = \{1, 2, \ldots, i\}$. We also take $f(\emptyset) = 0$.
Then:

$$\begin{aligned}
\widehat{f}(z) &= \sum_{i=1}^{n-1} P(z_{i+1} \leqslant \lambda \leqslant z_i) \cdot f(S_i) + z_n \cdot f(S_n) \\
&= \sum_{i=1}^{n-1} (z_{i+1} - z_i) \cdot f(S_i) + z_n \cdot f(S_n)
\end{aligned}$$

Now, consider the following lemma:

Lemma 26.4. *The Lovász extension $\widehat{f}(z)$ of submodular $f(x)$ can be expressed an LP:*

$$\widehat{f}(z) = \max_x x^T z : x(S) \leqslant f(S), x(N) = f(N), \forall S \subseteq N$$

where $x(S) = \sum_{i \in S} x_i$. Let F denote the feasible region.
Given Lemma 26.4, the rest of the proof follows easily:

$$\begin{aligned}
\widehat{f}(\lambda z + (1-\lambda) z') &= \max_{x \in F} x^T (\lambda z + (1-\lambda) z') \\
&\leqslant \lambda \max_{x \in F} x^T z + (1-\lambda) \max_{x \in F} x^T z' \\
&= \lambda \widehat{f}(z) + (1-\lambda) \widehat{f}(z')
\end{aligned}$$

Now we just have to prove Lemma 26.4.
Consider the dual problem to the above LP:

$$\min_y \sum_{S\subseteq N} y_S \cdot f(S) : \sum_{S\subseteq N} y_S e_S = z, y_S \geqslant 0$$
$$\text{where } (e_S)_i = \begin{cases} i & \text{if } i \in S \\ 0 & \text{else} \end{cases}$$

Our goal is to find a primal-dual feasible solution x^*, y^* that satisfies:

$$\widehat{f}(z) = c^T x^* = \sum_{S\subseteq N} y^*_S \cdot f(S)$$

Consider the following x^*, y^*:

$$x^*_i = f(S_i) - f(S_{i-1})$$
$$y^*_S = \begin{cases} z_i - z_{i+1} & \text{if } S = S_i \\ z_n & \text{if } S = N \\ 0 & \text{else} \end{cases}$$

First, we show that x^* is primal feasible.
For the constraint $x(N) = f(N)$, we see that:

$$\begin{aligned} x^*(N) = \sum_{i=1}^n x^*_i &= \sum_{i=1}^n f(S_i) - f(S_{i-1}) \\ &= f(N) - f(\emptyset) = f(N) \end{aligned}$$

For the other constraint, $x(S) \leqslant f(S)$, we proceed by induction on $|S|$ (the size of S).
As a base case, we take $S = \emptyset$.
Suppose that the constraint is satisfied for all sets of size $i-1$.
Then consider an arbitrary S, whose largest element is i:

$$\begin{aligned} f(S) + f(S_{i-1}) &\geqslant f(S \cup S_{i-1}) + f(S \cup S_{i-1}) \\ &= f(S_i) + f(S \setminus \{i\}) \\ &\geqslant f(S_i) + x^*(S \setminus \{i\}) \\ \Rightarrow f(S) &\geqslant f(S_i) - f(S_{i-1}) + x^*(S \setminus \{i\}) \\ &= x^*_i + x^*(S \setminus \{i\}) \\ &= x^*(S) \end{aligned}$$

where we used fact that f is submodular and that i is the largest element in S.
It follows that $x^*(S) \leqslant f(S)$, for all $S \subseteq N$, and thus x^* is primal feasible.

Next, we show that y^* is dual feasible.

$$\sum_{S\subseteq N} y^*_S \cdot f(S) = \sum_{j=1}^{n-1}(z_j - z_{j+1})e_{S_j} + z_n e_n$$

$$\begin{aligned}\left(\sum_{S\subseteq N} y^*_S \cdot f(S)\right)_i &= \left(\sum_{j=1}^{n-1}(z_j - z_{j+1})e_{S_j} + z_n e_n\right)_i \qquad \text{(consider it elementwise)}\\ &= \sum_{j=i}^{n}(z_j - z_{j+1}) + z_n\\ &= z_i\end{aligned}$$

Since this equality holds elementwise for each i, we see that $\sum_{S\subseteq N} y^*_S \cdot f(S) = z$.

We also observe that $y^*_S \geqslant 0$ due to the nonincreasing ordering of $z_1, \ldots, z_n$.

Thus, it follows that y^*_S is dual feasible.

Finally, we show that x^*, y^* are optimal, achieving a dualty gap of zero:

$$\begin{aligned}z^T x^* &= \sum_{i=1}^{n} z_i\big(f(S_i) - f(S_{i-1})\big)\\ &= \sum_{i=1}^{n}(z_i - z_{i+1})f(S_i) + z_n f(S_n)\\ &= \widehat{f}(z)\\ \widehat{f}(z) &= \sum_{i=1}^{n}(z_i - z_{i+1})f(S_i) + z_n f(S_n)\\ &= \sum_{S\subseteq N} y^*_S f(S)\end{aligned}$$

This concludes the proof of Lemma 26.4, and that of Theorem 26.3.

References