Submodular functions 2: Lovász extension, greedy algorithm

Yaqiao Li

March 15, 2018

1 Lovász extension and convexity

Three ways to define the Lovász extension of a submodular function $f : 2^{[n]} \to \mathbb{R}$:

(i). $\widehat{f}(z) = \mathbb{E}_{\lambda \sim [0,1]} f(\{i : z_i \geq \lambda\})$.

(ii). analytical expansion of the above: order the z_i increasingly, suppose $z_n \leq z_{n-1} \leq \cdots \leq z_2 \leq z_1$, and let $S_i := \{1, 2, \ldots, i\}$, then

$$\widehat{f}(z) := \sum_{i=1}^{n} (z_i - z_{i+1}) f(S_i) \tag{1}$$

where we let $z_{n+1} = 0$. Observe that this immediately implies

$$\widehat{f}(cz) = c\widehat{f}(z), \quad \forall c \in \mathbb{R}. \tag{2}$$

Exchanging the sum, one has equivalently,

$$\widehat{f}(z) := \sum_{i=1}^{n} z_i (f(S_i) - f(S_{i-1})) \tag{3}$$

where $f(\emptyset) = 0$. Note that the above two analytical expression depends critically on the ordering of coordinates of z, so they do not imply that $\widehat{f}$ is a linear function.

(iii). $\widehat{f}(z)$ is the optimal value of the following LP: $\max z^T x$, s.t. $x \in B(f)$, where

$$B(f) := \{x \in \mathbb{R}^n : x(S) \leq f(S) \ \forall\ S; \text{ and } x([n]) = f([n])\}.$$

$B(f)$ is called the base polyhedron of the submodular function f: it is nonempty. The nonemptyness of $B(f)$ has explicit implication in allocating cost for a grand coalition in cooperative games.

Note that the definition (iii) immediately implies $\widehat{f}(z)$ is convex.

Theorem 1 (Lovász)**.** *f is submodular if and only if $\widehat{f}$ is convex.*

Sketch. $\Longleftarrow$: use property (2).

$\Longleftarrow$: This is the same as to show the definition in (ii) and (iii) are equivalent. By the primal-dual theory, one can look at the dual problem of the LP given in (iii). Then choose the feasible solution for primal according to (3), and choose the feasible solution for dual by (1), show the values obtained from them in the primal and dual LP are equal, hence they are both optimal. □

2 Maximizing a monotone submodular function with a size constraint

Given a monotone submodular $f : 2^V \to \mathbb{R}^+$.
Problem: $\max_{|S|=k} f(S) = ?$

Let $f_S(x) := f(S \cup \{x\}) - f(S)$ denote the marginal increase. It's easy to show f_S is also submodular and subadditive.
Greedy algorithm: start from the emtpy set, each step chooses an element that has the maximal marignal increase, until k elements has been chosen.

Theorem 2. *Greedy algorithm gives* $(1 - 1/e)$*-approximation.*

Sketch. Let A be the set of size k that is optimal, let $S_1, S_2, \ldots, S_k$ be the set that is successively chosen in the greedy algorithm. To show $f(S_k) \geq (1 - 1/e)f(A)$.

So in step i we have set S_i and the greedy algorithm chooses x, so $S_{i+1} = S \cup \{x\}$. By "greedy", one has

$$f(S_{i+1}) - f(S_i) = f_{S_i}(x) \geq f_{S_i}(y), \quad \forall\ y \in V.$$

Hence

$$k(f(S_{i+1}) - f(S_i)) \geq \sum_{y \in A} f_{S_i}(y) \geq f_{S_i}(A) = f(S_i \cup A) - f(S_i) \geq f(A) - f(S_i),$$

where the second inequality is by the fact that f_{S_i} is sub-additive. This shows,

$$f(S_{i+1}) \geq (1 - \frac{1}{k})f(S_i) + \frac{1}{k}f(A).$$

Use an induction to show $f(S_j) \geq (1 - (1 - \frac{1}{k})^j)f(A)$ to conclude the proof. □