Lecture #13

Last Time: Ellipsoid Algorithm
separation $\Rightarrow$ optimization

Today: Submodular Functions

Let N be a set with $|N| = n$

blackbox

A function $f: 2^N \to \mathbb{R}$ is submodular if

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

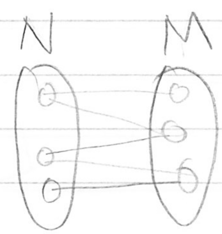
for all $A \subseteq B$, and j. "diminishing returns"

Alternatively $f(A \cap B) + f(A \cup B) \leq f(A) + f(B)$ $\forall A, B$

Examples:

(1) coverage function

e.g. bipartite graph



where $f(A) = |\{u \in M \text{ s.t. } u \text{ is adjacent to node in } A\}|$

(2) entropy where N is a collection of random variables $X_1, X_2 \ldots X_N$ and

$$f(A) = h(\{X_i\}_{i \in A})$$

(3) graph cut function for some graph $G = (\overset{\text{vertices}}{N}, E)$

e.g. $f(A) = |E(A, N/A)|$

Many questions we can ask about them:

maximize; minimize; max/min subject to constraints

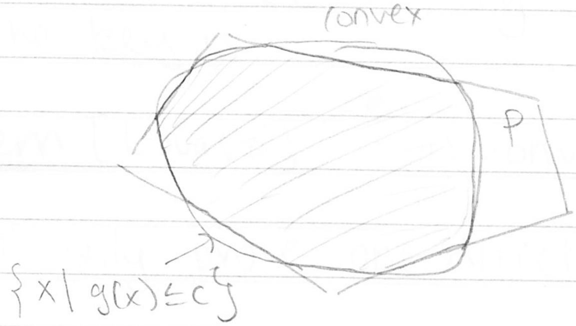
We will see how to efficiently minimize them, by developing analogy with convexity

A function $g: \mathbb{R}^n \to \mathbb{R}$ is <u>convex</u> if

$$\lambda \cdot g(x) + (1-\lambda) g(y) \geq g(\lambda x + (1-\lambda) y)$$

Such functions are easy to minimize over convex sets, picturally consider:

$$\{x \mid \underbrace{x \in P}_{\text{convex}}, \text{ and } g(x) \leq c\}$$



* actually use <u>sub-gradient</u> *

sketch:

We can use the Ellipsoid algorithm and given a query x, output separating hyperplane for P <u>OR</u> $-\nabla g(x)$ if value not small enough

<u>Main Idea:</u> Submodular functions are discrete analogues of convex functions

A submodular function $f : \{0,1\}^n \to \mathbb{R}$, let's extend it to all of $[0,1]^n$

The Lovasz extension $\hat{f} : [0,1]^n \to \mathbb{R}$ is

$$f(z) = \mathop{\mathbb{E}}_{\lambda \sim_u [0,1]} \left[f(\{ i \mid z_i \geq \lambda \}) \right]$$

will assume throughout wlog, to simplify notation

In particular, if $z_1 \geq z_2 \ldots \geq z_n$ and $S_i = \{1, 2, \ldots i\}$ then

$$f(z) = \sum_{i=1}^{n-1} (z_i - z_{i+1}) f(S_i) + z_n f(S_n)$$

let's assume $f(\phi) = 0$ $\leftarrow$ $\cancel{+ (1 - z_1) f(\phi)}$

This is well defined for any function $f : \{0,1\}^n \to \mathbb{R}$, but the key is

Theorem [Lovasz]: $\hat{f}$ is convex iff f is submodular

we will only prove one direction $(\Leftarrow)$

Proof: of $(\Leftarrow)$: Consider the following LPs

(P) $\max \; z^T x$ $\left(x(S) = \sum_{i \in S} x_i \right)$

$$\text{s.t. } x(S) \leq f(S) \quad \forall S \subsetneq N$$
$$X(N) = f(N)$$

The dual is the following

(D) $\min \sum_{S \subseteq N} y_S f(S)$

indicator for set S

s.t. $\sum_{S \subseteq N} y_S e_S = z$

$y_S \geq 0 \quad \forall S \subsetneq N$ (because we have equality constraint in (P) for $S = N$)

Let's guess the optimal solutions

$$x_i^* = f(S_i) - f(S_{i-1})$$

and let

$$y_S^* = \begin{cases} z_i - z_{i+1} & \text{if } S = S_i \\ z_n & \text{if } S = N \\ 0 & \text{else} \end{cases}$$

Claim #1: y^* is feasible

Proof: $z_1 \geq z_2 \dots \geq z_n$ so $y_S^* \geq 0$, moreover

$$\left(\sum_{S \subseteq N} y_S^* e_S \right)_i = y_{S_i}^* + y_{S_{i+1}}^* \dots y_{S_n}^*$$

$$= z_i \quad \boxtimes$$

Claim #2: x^* is feasible

we will come back to this

Now $z^T x^* = \sum_{i \in N} z_i (f(S_i) - f(S_{i-1}))$

$$= \sum_{i=1}^{n-1} (z_i - z_{i+1}) f(S_i) + z_n f(S_n) = \hat{f}(z)$$

$$= \sum_{S \subseteq N} y_S^* f(S) =$$

(make lemma, prove theorem first)

Now $\hat{f}(z)$ = optimal value of (P)

$$= \max_{x \text{ feasible}} z^T x$$

Hence $\hat{f}(z+z') = \max_{x \text{ feasible}} (z+z')^T x$

$$\leq \underbrace{\max_{x \text{ feasible}} z^T x}_{\hat{f}(z)} + \underbrace{\max_{x \text{ feasible}} z'^T x}_{\hat{f}(z')}$$

Thus $\hat{f}$ is convex! ☒

All that remains is to prove claim #2

Proof: Clearly $x^*(N) = f(N)$. Prove by induction on $|S|$. If $S = \emptyset$, then $f(\emptyset) = 0$ and $x^*(\emptyset) = 0$.

Let i be largest element in S, then

$$f(S) + \underbrace{f(S_{i-1})}_{\{1 \ldots i-1\}} \geq \underbrace{f(S \cup S_{i-1})}_{S_i} + \underbrace{f(S \cap S_{i-1})}_{S \setminus \{i\}}$$

Rearranging we get:

$$f(S) \geq \underbrace{f(S_i) - f(S_{i-1})}_{x_i^*} + \underset{\substack{\geq \text{ by induction} \\ x^*(S \setminus \{i\})}}{f(S \setminus \{i\})}$$

$$\geq x^*(S) \quad \boxed{}$$