Supplement to the Briefing on Min-Norm-Point Algorithm, Lovasz Extension

Abstract:

This note makes up the important contents on Lovasz extension that I did not cover in the last presentation: why the minimum value of Lovasz extension coincides with the minimum of a submodular function, and why the vector in the base polytope with minimal norm outputs the minimum of Lovasz extension (the proof of this statement implements minimax theorem).

In the previous briefing I did not answer one question:

- Why minimizing the norm of a vector on the base polytope also minimizes the result of Lovasz extension?

We can answer this question using minimax theorem.

Revisiting Definitions and Conclusions Introduced Before

1) A *submodular function* is a function *F* on a set *V* such that: if $A \subseteq B \subseteq V$ and $e \notin B$, then $F(A \cup \{e\}) - F(A) \geq F(B \cup \{e\}) - F(B)$. We often assume $F(\emptyset) = 0$. Equivalently, it is a function on *V* such that for two sets $S, T \subseteq V$:

$$F(S) + F(T) \geq F(S \cap T) + F(S \cup T)$$

2) Lovasz extension: given a vector $x \in [0,1]^n$ and a submodular function *F* on $V = \{1, 2, \cdots, n\}$ with $F(\emptyset) = 0$, assuming that entries of *x* are ordered as

$$x_{S_1} \leq x_{S_2} \leq \cdots x_{S_n}$$

then the Lovasz extension *f* of *x* is:

$$f(x) = x_{S_1} F(\{S_k\}_{k=1}^{n}) + \sum_{k=2}^{n} (x_{S_k} - x_{S_{k-1}}) F\left(\{S_j\}_{j=k}^{n}\right) \quad (1)$$

which is equivalent to:

$$\sum_{k=1}^{n} x_{S_k} \left[F\left(\{S_j\}_{j=k}^{n}\right) - F\left(\{S_j\}_{j=k}^{n} \setminus \{S_k\}\right) \right] \quad (2)$$

i.e. the sum of each x_{S_k} times the "marginal gain" of the function value of some *F(S)* when S_k is added to *S*, with the sequence of addition being S_n, S_{n-1}, ..., S_1.

Additionally, formula (2) is equivalent to:

$$\max_y\{\langle x,y\rangle | y(S) \le F(S) \forall S \subseteq V, y(V) = F(V)\} \quad (3)$$

where *y(S)* denotes the sum of the x^{th} entries of *y* for all *x* in *S*.

The set of vectors *y* that satisfy the restrictions in formula (3) is called the *base polytope* of *F*. A base polytope is convex.

Equating the Minimum of Lovasz Extension to That of *F*

First we may notice that the minimum value of the Lovasz extension always occurs at the boundary of the domain of *x*, i.e. $\{0,1\}^n$.

Proposition 1

Given a submodular function *F* on set *V={1, 2, ..., n}*, and name the Lovasz extension as *f*, there is an $x \in \{0,1\}^n$ such that $f(x) = \min_{z\in[0,1]^n} f(z)$.

Proof: We can show that for each *z* that has a non-zero-or-one entry, there is another vector *z'* on the boundary such that *f(z')≤f(z)*.

Arrange the entries of *z* in increasing order:

$$z_{S_1} \le z_{S_2} \le \cdots z_{S_n} \quad (4)$$

and assume $0 < z_{S_j} < 1$. As is shown, $f(z) = \langle z,y\rangle$, $y = [y_1,\cdots,y_n]^T$, for some *y* on the base polytope. Since $y_{S_k} = F(\{S_t\}_{t=k}^n) - F(\{S_t\}_{t=k}^n \setminus \{S_k\}) \forall k$, when entries of *z* change, as long as the order of entries remains as formula (4), *y* still makes $\langle y,z\rangle$ equal to *f(z)*. Therefore, we can make the following changes on z_{S_j}:

Step 1: If $z_{S_j} y_{S_j} \geq 0$, then decrease z_{S_j} until it equals to *0* or $z_{S_{j-1}}$; If $z_{S_j} y_{S_j} < 0$, then increase z_{S_j} until it reaches *1* or $z_{S_{j+1}}$.

Step 2: If now z_{S_j} equals to some $z_{S_{j-1}}$ (or $z_{S_{j+1}}$. For the simplicity of description we assume it is $z_{S_{j-1}}$) that is not *0* or *1*, then consider $z_{S_j} y_{S_j} + z_{S_{j-1}} y_{S_{j-1}}$. If it is not less than *0*, then decrease the value of z_{S_j} and $z_{S_{j-1}}$ until they reach *0* or some other entry (or entries). If it is less than *0* then increase their values. Repeat this process to move all the encountered entries to *0* or *1*.

Step 3: If there is still an entry that is not *0* or *1*, repeat steps 1 and 2 to this entry until we obtain a vector $z' \in \{0,1\}^n$.

Throughout the process when *z* is changed to *z'*, the order of its entries is not changed, and thus $f(z') = \langle z', y \rangle \leq \langle z, y \rangle = f(z)$. $\square$

Furthermore, from expression (1) we can see that for each $z' = [z_1^{'}, \cdots, z_n^{'}]^T$ with *j* being the smallest number such that $z_{S_j} = 1$, we have $f(z') = F(\{S_k\}_{k=j}^n)$. Hence:

1. Since each *f(z')* equals to some function value of *F*,

$$\min_{z \in [0,1]^n} f(z) = \min_{z' \in \{0,1\}^n} f(z') \geq \min_{S \subseteq V} F(S)$$

2. For the set *S'={s_1, s_2, ..., s_k}* that minimizes *F*, let *x* be an n-entry vector with the *s_1-th*, *s_2-th*, ..., *s_k-th* entries being *1* and other entries *0*, we have:

$$\min_{S \subseteq V} F(S) = f(x) \geq \min_{z' \in \{0,1\}^n} f(z)$$

Thus we conclude that **the minimum of Lovasz extension is the minimum of *F*.** This conclusion enables us to transform the search of the minimum of a discrete set function to the search of the minimum of a continuous function.

The minimum-norm-point provides the minimum in Lovasz Extension

This part refers to *"Lexicographically Optimal Base of a Polymatroid with Respect to a Weight Vector"* by Satoru Fujishige, published as Vol. 5, No. 2 of the May, 1980 release of *"Mathematics of Operations Research"*.

As described in the last presentation, using the minimax theorem to formula (3), if we name the base polytope of *F* as *B*, then

$$\min_{x\in[0,1]^n} \max_{y\in B} \langle x, y\rangle = \max_{y\in B} \min_{x\in[0,1]^n} \langle x, y\rangle = \max_{y\in B} y\big(S^-(y)\big)$$

where we can define the negative entries of *y* as *sink* (or *sink entries*), denoted as $S^-(y)$, and other entries as *sourse* (or source entries), denoted as $S^+(y)$.

As previously stated, we have the following claim: the minimum-norm vector *y'* in *B* satisfies that $y\big(S^-(y)\big) = \min_{S\subseteq V} F(S)$. To prove this we can prove the contrapositive statement: for a vector *y* in the base polytope, if the sum of its sink entries can be increased, then its norm is not minimal. To show this, we can use the following lemmas:

Lemma 1:

For a vector *y* in *B*, if $y\big(S^-(y)\big)$ is not maximal, then *y* contains both sink entries and source entries.

Proof: Suppose *y* contains only sink entries, then for any other vector *z* in *B*,

$$z(V) = y(V)$$

$$\Rightarrow z\big(S^+(z)\big) + z\big(S^-(z)\big) = y(V) = y\big(S^-(y)\big)$$

$$\Rightarrow z\big(S^-(z)\big) \le y\big(S^-(y)\big)$$

Then $y\big(S^-(y)\big)$ should be maximal, contradicting to the assumption.

Otherwise, when y contains only source entries, $y(S^-(y))$ is 0, which is the maximum value achievable by a sum of sink entries. □

Lemma 2

For a vector y in the base polytope B, if its sum of sink entries is not maximal, and another vector y' satisfies that $y'(S^-(y'))$ attains maximum, then there exists some k in $S^-(y)$ such that $y_k < y_k'$.

Proof: Otherwise the sink of y' contains that of y, and the sum of sink entries of y' is less than or equal to that of y, but the assumption is that $y(S^-(y)) < y'(S^-(y'))$. □

Since a base polytope is convex, in the process where y changes to y' linearly, it remains in the base polytope. By Lemma 2, during this change there are some negative entries increase. The way to ultimately associate minimal norm to maximal sum of sink entries is to illustrate that there exists a sink entry a and a source entry b such that to counter the increase in y_a, it suffices to decrease y_b by the same amount.

We shall introduce some definitions in proving this. For y in base polytope, define the *saturated set* as:

$$sat(y) = \{u | \exists S \text{ so that } u \in S \subset V, y(S) = F(S)\}$$

where $\subset$ signifies "is a proper subset of".

By the definition of submodular functions, if $y(S)=F(S)$ and $y(T)=F(T)$, then:

$$y(S \cap T) = F(S \cap T), y(S \cup T) = F(S \cup T)$$

and hence if u is in $sat(y)$, there is a minimal set that contains u such that the sum of entries of locations in this set equals to the value of submodular function. Hence we can define the *dependent set* of u in y as:

$$dep(y,u) = \begin{cases} \bigcap_{u \in S \subset V, y(S)=F(S)} S, & \text{if } u \in sat(y) \\ \emptyset & \text{otherwise} \end{cases}$$

This definition shows that if u is in *sat(y)* and y_u increases, then some entries in *dep(y,u)* must in total decrease by the same amount to keep y in the base polytope.

Now we can see that for a sink entry y_k that increases when y changes linearly to y', if k is not in *sat(y)*, then we can take any source entry a (which exists by Lemma 1), increase y_k by a small value d, and decrease y_a by d. Then we get a vector in the base polytope where y_k and y_a are closer to their average, and has a smaller norm than y.

When all sink entries that increase during the linear change are in the saturated set, then we claim that among them there is a k where *dep(y,k)* contains at least one positive source entry a, since otherwise the transformation of y into y' involves only increase in some sink entries and decrease in other sink entries or zeros, which, we may easily discover does not increase $y\left(S^{-}(y)\right)$. As there exists a sink entry k and a source entry a in *dep(y,k)*, we can perform the perturbance as described in the previous paragraph to obtain a vector in the base polytope that has lower norm.

Thus we can conclude that **when the sum of the sink vector can be increased, its norm can be decreased, i.e. the min-norm-point inside the base polytope maximizes the Lovasz Extension.**