05 April 2024 21:34

Combinatorial Optimisation

Topic ÷ Submodular function minimization

- Rishav Gupta

05 April 2024 21:35

Definitions :-

1.) Submodular function :-

A function $f: 2^E \to \mathbb{R}$ is said to be submodular iff $\forall X, Y \subseteq E$

$$f(X) + f(Y) \geq f(X \cap Y) + f(X \cup Y)$$

Another equivalent definition :-

A function $f: 2^E \to \mathbb{R}$ is said to be submodular if $\forall X, Y \subseteq E$ such that $Y \supseteq X$, $e \in Y^c$ we have

$$f(X \cup \{e\}) + f(Y) \geq f(X) + f(Y \cup \{e\})$$

05 April 2024 21:44

Associated Polyhedrons :

For a submodular function f, we associate following three polyhedrons

* $P_f = \{ x \in \mathbb{R}_+^E \mid \forall\, U \subseteq E,\ x(U) \leq f(U) \}$
* $EP_f = \{ x \in \mathbb{R}^E \mid \forall\, U \subseteq E,\ x(U) \leq f(U) \}$
* $B_f = \{ x \in \mathbb{R}^E \mid \forall\, U \subseteq E,\ x(U) \leq f(U),\ x(E) = f(E) \}$

where $x(U) = \sum_{u \in U} x_u$

Facts :

* P_f is said to be the polymatroid associated to f.
* EP_f is said to be the extended polymatroid associated to f.
* B_f is said to be the base polytope

05 April 2024 21:57

Greedy Algorithm for optimising over polymatroid (extended) :-

* Let f be a submodular function, $f: 2^E \to \mathbb{R}$
* Consider $w: E \to \mathbb{R}_{\geq 0}$ be a non negative weight function.
* Now we want to maximize $w \cdot x$, $x \in EP_f$.
* (Note, w is taken to be non-negative because if w was not non-negative then the above optimum would be unbounded.)
* Similarly, WLOG we can take $f(\phi) = 0$ because if $f(\phi) < 0$, EP_f is ϕ $\because x(\phi) = 0 < f(\phi)$ and if $EP_f \neq \phi$, i.e $f(\phi) \geq 0$ setting $f(\phi) = 0$ would not violate submodularity condition as now the only cases to check is when X, Y are disjoint or one of them is ϕ.

05 April 2024 22:21

Greedy Algorithm (Contd.)

1.) Order $E = \{s_1, s_2 \cdots s_n\}$ such that
$$w(s_1) \geq w(s_2) \cdots \geq w(s_n)$$

2.) Let $A_i = \{s_1, s_2 \cdots s_i\}$ for $1 \leq i \leq n$, $A_0 = \phi$

3.) Let $x'(s_i) = f(A_i) - f(A_{i-1})$, for $1 \leq i \leq n$

4.) Output x'

Note: The same algorithm we saw for case of matroids just replace f by rank of the matroid.

* For correctness of above algorithm we need to show feasiblity and optimality of x'.

Proof idea:

* We will use submodularity for primal feasiblity.
* Then find a feasible dual solution with same cost.

* Then find a feasible dual [illegible]
cost.

05 April 2024 22:34

Proof of feasiblity :

We need to show $\forall U \subseteq E$, $x'(U) \leq f(U)$

for $U = \phi$ it is obvious as $x'(\phi) = 0$

Let's say $U = \{ s_{i_1}, s_{i_2} \text{----} s_{i_k} \}$

Then $x'(U) = \sum_{j=1}^{k} f(A_{i_j}) - f(A_{i_j - 1})$

$$\sum_{j=1}^{k} f(A_{i_j}) - f(A_{i_{j-1}})$$

Let $B_j = \{ s_{i_1}, s_{i_2} \text{---} s_{i_j} \}$

Now $B_{j-1} \subseteq A_{i_j - 1}$

$$\Rightarrow \left. \begin{array}{l} f(B_{j-1} \cup \{ s_{i_j} \}) - f(B_{j-1}) \geq \\ f(A_{i_j - 1} \cup \{ s_{i_j} \}) - f(A_{i_j - 1}) \end{array} \right\} \text{(By submodularity)}$$

$$\Rightarrow f(B_j) - f(B_{j-1}) \geq f(A_{i_j}) - f(A_{i_j - 1})$$

$$\Rightarrow \sum_{j=1}^{k} f(B_j) - f(B_{j-1}) \geq \sum_{j=1}^{k} f(A_{i_j}) - f(A_{i_j - 1})$$

$$\Rightarrow f(U) \geq x'(U)$$

$$\Rightarrow \quad f(U) \geq x'(U)$$

$\Rightarrow$ x' is feasible

05 April 2024 22:56

Proof of optimality :

Primal :

$$\max \sum_{e \in E} w_e x_e$$

$$\sum_{e \in U} x_e \leq f(U), \forall U \subseteq E$$

Dual :

$$\min \sum_{U \subseteq E} y_U f(U)$$

$$\sum_{e \in U} y_U = w_e, \forall e \in E$$

$$y_U \geq 0, \forall U \subseteq E$$

Define the dual solution :

$$y'(A_n) = y'(E) = w(S_n)$$

$$y'(A_i) = w(S_i) - w(S_{i+1}), \text{ for } 1 \leq i \leq n-1$$

$$y'(U) = 0, \text{ for other } U \subseteq E$$

Feasiblity of y' :

To show, $\sum_{e \in U} y'(U) = w(e)$

WLOG, $e = S_i$

$$\sum_{e \in U} y'(U) = \sum_{j \geq i} y'(A_j)$$

$$w(S_n) + \sum_{i \leq j \leq n-1} w(S_j) - w(S_{j+1})$$

$$= \omega(s_i)$$

$$= \omega(e)$$

Hence feasible for dual

Now it only remains to show y' has same cost as x' in primal

$$\sum f(U)\, y(U)$$
$$= f(A_n)\,\omega(S_n) + \sum_{1 \leq i \leq n-1} f(A_i)\big(\omega(S_i) - \omega(S_{i+1})\big)$$
$$= \sum_{1 \leq j \leq n} \omega(S_j)\big(f(A_j) - f(A_{j-1})\big)$$
$$= \text{Primal cost for } x'$$

$\Rightarrow$ By Duality theorem x' is optimal

$\Rightarrow$ Greedy works for polymatroids :)

05 April 2024 23:26

An observation :

Lemma : Let f be a submodular function on S with $f(\phi) = 0$, Then we have

$$f(U) = \max\{x(U) \mid x \in EP_f\}$$

Proof : Just in the previous greedy algorithm consider

$$w(e) = \begin{cases} 1 & , \text{ if } e \in U \\ 0 & , \text{ otherwise} \end{cases}$$

and previous greedy algorithm will give us $\sum w(e)\, x'(e) = f(U)$

So we can extract f from EP_f as above.

Hence there is a one-one correspondence b/w f and EP_f.

Note : [The previous algorithm also works

and P_f

Note: The previous algorithm also works with arbitary weight functions and P_f

i.e $\quad \max\, w \cdot x \,, \; x \in P_f$

Similarly we get $f(U) = \max\{x(U) \mid x \in P_f\}$

and f is a non negative monotone submodular function.

06 April 2024 10:45

Characterization of vertices of associated polyhedron:

Consider a strict linear order

$\omega_1 > \omega_2 \cdots > \omega_n > 0$

*We claim, for the above strict weights, we have a unique optimum at x' as devised by greedy algorithm.

Now we have $\hat{x}(s_i) = f(A_i) - f(A_{i-1})$

Consider the following quantity

$$\sum_{e\in E} w(e)\hat{x}(e) - \sum_{e\in E} w(e)y(e)$$

$$= \sum_{e\in E} w(e)\left(\hat{x}(e) - y(e)\right)$$

$$= \sum_{i=1}^{n} \sum_{e\in A_i} \omega_i (\hat{x}(e) - y(e)) - \sum_{e\in A_{i-1}} \omega_i (\hat{x}(e) - y(e))$$

$$= \sum_{i=1}^{n} \left(\sum_{e\in A_i} \omega_i (\hat{x}(e) - y(e))\right) - \left(\sum_{e\in A_{i-1}} \omega_i (\hat{x}(e) - y(e))\right)$$

$$= \left(\sum^{n-1} (\omega_i - \omega_{i+1}) \sum_{e\in A_i} (\hat{x}(e) - y(e))\right) + \omega_n \sum_{e\in A_n} (\hat{x}(e) - y(e))$$

$$= \left(\sum_{i=1}^{n-1} (\omega_i - \omega_{i+1}) \sum_{e \in A_i} (\hat{x}(e) - y(e)) \right) + \omega_n \sum_{e \in A_n}$$

$$= \left(\sum_{i=1}^{n-1} (\omega_i - \omega_{i+1}) \cdot \left(\left(\sum_{e \in A_i} \hat{x}(e) \right) - \left(\sum_{e \in A_i} y(e) \right) \right) \right) +$$

$$\omega_n \left(\sum_{e \in A_n} \hat{x}(e) - \sum_{e \in A_n} y(e) \right)$$

$$(\because A_n = E)$$

we get

$$= \sum_{i=1}^{n-1} (\omega_i - \omega_{i+1}) \left(f(A_i) - y(A_i) \right) + \omega_n \left(f(E) - y(E) \right)$$

Now if weighing is strict and +ve

$\Rightarrow (\omega_i - \omega_{i+1}) > 0$

Now if $y \in EP_f \Rightarrow$ $y(A_i) \leq f(A_i) \ \forall i$

$y(E) \leq f(E)$

if $\forall i \in [1 \cdots n]$, $y(A_i) = f(A_i)$

$\Rightarrow y(s_i) = f(A_i) - f(A_{i-1})$

$\Rightarrow y = \hat{x}$

$\Rightarrow \exists$ some i s.that $y(A_i) < f(A_i)$

if $y \neq \hat{x}$ and $y \in P_f$

$\Rightarrow$ For above weights there exist Unique optimum at $\hat{x}$

$\Rightarrow$ $\hat{x}$ is a vertex of P_f

as well ...

$\Rightarrow \hat{x}$ is a vertex of P_f

as well as B_f.

$\therefore \hat{x} \in B_f$

Now we need to prove these are the only vertices

To prove this we will use the fact at a vertex n, linearly independent constraints are tight

if suppose for the above LP, a vertex y

$$\left.\begin{aligned} \sum_{e\in U_1} x_e &\leq f(U_1) \\ \sum_{e\in U_2} x_e &\leq f(U_2) \\ &\vdots \\ \sum_{e\in U_n} x_e &\leq f(U_n) \end{aligned}\right\} \text{These constraints are tight at } y$$

$$\Rightarrow \quad \sum_{e\in U_i} y_i = f(U_i), \ \forall i\in[1 \,..\, n]$$

Now consider the following

$$\sum_{i=1}^{n} \sum_{e\in U_i} y_i = \sum_{i=1}^{n} f(U_i)$$

Now for the objective function $\sum_{i=1}^{n} \sum_{e\in U_i} y_i$ will have a unique optimum and the weight defined by this function is non-negative so it must attain optimum at some vertex described by greedy.

$\Rightarrow$ We get a one to one correspondence between linear orders and vertices.

$\Rightarrow$ we get a one to one correspo[ndence]
between linear orders and vertices.

06 April 2024 12:57

Lovasz Extension Theorem ÷

Let p be a weight function such that

$$p_{i_1} \geq p_{i_2} \ldots \geq p_{i_n}$$

$A_j = \{ i_1, i_2 \ldots i_j \}$, then

$$\hat{f}(x) = (1 - p_{i_1}) f(\phi) + \sum_{j=1}^{n} (p_{i_j} - p_{i_{j+1}}) f(A_j)$$

From previous greedy algorithm

If f is submodular then $\hat{f}(x) = \max \{ x \cdot y \mid y \in EP_f \}$

Theorem ÷ A set function $f : 2^E \to \mathbb{R}$ with $f(\phi) = 0$ is submodular iff $\hat{f}$ is convex.

Proof ÷ If f is submodular. take $x, x' \in \mathbb{R}_n^+$, $t \in [0,1]$

$$z = tx + (1-t)x'$$

To show $\hat{f}$ is convex we need to show

$$\hat{f}(z) \leq \hat{f}(tx) + \hat{f}((1-t)x')$$

Let $y^* \in EP_f$ such that

$$\hat{f}(z) = z \cdot y^* = t x \cdot y^* + (1-t) x' \cdot y^*$$

We know by optimality

$$\hat{f}(tx) \geq t x \cdot y^*$$

$$\hat{f}((1-t)x') \geq (1-t) x' \cdot y^*$$

$\hat{f}((1-t)x') \geq (1-t)x' \cdot y$

$\Rightarrow \hat{f}(z) \leq \hat{f}(tx) + \hat{f}((1-t)x')$

$\Rightarrow \hat{f}$ is convex

Now if $\hat{f}$ is convex

Notation
$$\chi(S) = \begin{cases} \chi(S)(v) = 1 & \text{if } v \in S \\ = 0, & \text{otherwise} \end{cases}$$

$$\chi(S) : 2^E \to \mathbb{R}^n$$

Note
$$\hat{f}(\chi(A) + \chi(B)) = \hat{f}(\chi(A \cup B)) + \hat{f}(\chi(A \cap B))$$

as for $\chi(A) + \chi(B)$

the weights are $\begin{cases} 2, & \text{if } v \in A \cap B \\ 1, & \text{if } v \in A \setminus B \text{ or } B \setminus A \\ 0, & \text{otherwise} \end{cases}$

So the elements in $A \cap B$ will be ordered first

So we get that it is same as

$\hat{f}(\chi(A \cap B)) + \hat{f}(\chi(A \cup B))$

as for $A \cup B$ any permutation with correct ordering according to weights by $\chi_{A \cup B}$ will give optimum

So choose the one which ranks elements of $A \cap B$ first

$$\Rightarrow \hat{f}(\chi(A) + \chi(B)) = \hat{f}(\chi(A \cup B)) + \hat{f}(\chi(A \cap B))$$

$$\Rightarrow \hat{f}(\chi_{(A)} + \chi_{(B)}) = \hat{f}(\chi_{(A\cup B)}) + f(\chi(A\cap B))$$

$\because$ $\hat{f}$ is convex $\Rightarrow$ $\hat{f}\left(\frac{\chi_A + \chi_B}{2}\right) \leq \frac{\hat{f}(\chi_A) + \hat{f}(\chi_B)}{2}$

$\because$ scaling doesnt change ordering

$$\Rightarrow \hat{f}\left(\frac{\chi_A + \chi_B}{2}\right) = \frac{\hat{f}(\chi_A + \chi_B)}{2}$$

$$\Rightarrow \hat{f}(\chi_A + \chi_B) \leq \hat{f}(\chi_A) + \hat{f}(\chi_B)$$

$$\Rightarrow \hat{f}(\chi_{A\cup B}) + \hat{f}(\chi_{A\cap B}) \leq \hat{f}(\chi_A) + \hat{f}(\chi_B)$$

$$\Rightarrow f(A\cup B) + f(A\cap B) \leq f(A) + f(B)$$

Hence proved

06 April 2024 14:51

Min-Max Theorem

For a submodular function with $f(\phi)=0$

$$\min_{X \subseteq E} f(X) = \max\{z(E) \mid z \in EP_f, z \leq 0\}$$

$$= \max\{x^-(E) \mid x \in B_f\}$$

where $x^-(S) = \sum_{\substack{v \in S \\ x(v) \leq 0}} x_v$

Proof÷ If $z \in EP_f$, $z \leq 0$ for any $S \subseteq E$

$$z(E) \leq z(S) \leq f(S)$$

$$z(E) \leq f(S)$$

Now consider $f^0 : 2^V \to \mathbb{R}$

$$f^0(X) = \min_{Y \subseteq X} f(X)$$

If f is submodular $\Rightarrow$ f^0 is submodular

and $EP_{f^0} \subseteq EP_f$

Now for any $z \in B_{f^0}$, construct z using greedy algo

we have $z \leq 0$

and

$$z(E) = \min_{X \subseteq E} f(X) = f^0(E)$$

$$z(E) = \min_{X \subseteq E} f(X) = f^{-}(E)$$

$\Rightarrow$ The $z(E) = \min_{X \subseteq E} f(X)$

$\Rightarrow$ Equality is attained

$\because z \in EP_f$

07 April 2024 12:14

Proof: that f° is submodular

To show if $X \subseteq Y$, $e \notin Y$

$$\therefore f^\circ(X+e) - f^\circ(X) \geq f^\circ(Y+e) - f^\circ(Y)$$

or

$$f^\circ(Y) + f^\circ(X+e) \geq f^\circ(X) + f^\circ(Y+e)$$

For X minimum is attained at A'
" Y " " " " B'
" $X+e$ " " " " C'
" $Y+e$ " " " " D'

Note: The sets A', B', C', D' have the following property that value of f at these sets is smaller than value of f at any subset of these sets by minimality.

i.e. $\forall\, T \subseteq S$, $f(S) \leq f(T)$

Now for any set with above property we have the following result using submodularity of f

Let P be a set with above property

$$f(P) + f(T) \geq f(T \cup P) + f(T \cap P)$$

$$\because f(P) \leq f(T \cap P)$$

$$\Rightarrow f(T) \geq f(T \cup P)$$

$$\Rightarrow \quad f(T) \geq f(T \cup P)$$

So this tells us that taking union with such sets reduces the submodular function value.

$\Rightarrow$ WLOG we can assume

$$B' \cup C' \subseteq D', \quad B' \cap C' \supseteq A'$$

$\because$ if either B' or C' isn't contained in D' then we can take union with B' and C' and without reducing cost of f stay inside yte.

Similarly if $A' \not\subseteq B'$ or $A' \not\subseteq C'$

$B' \longrightarrow B' \cup A'$

$C' \longrightarrow C' \cup A'$

will also be optimal sets

Now we have, $A' \subseteq B' \cap C' \subseteq X$

and $D' \supseteq B' \cup C'$

By minimality of A' in X $\Rightarrow f(A') \leq f(B' \cap C')$

$\because D'$ has above property $\Rightarrow f(D') \leq f(B' \cup C')$

$$\Rightarrow \quad f(A') + f(D') \leq f(B' \cap C') + f(B' \cup C')$$
$$\leq f(B') + f(C')$$

$\Rightarrow \quad f(A') + \ldots$

$\leq f(B') + f(C')$
(By submodularity)

Hence proved

06 April 2024 23:36

Application of submodular minimization ÷

1.) Cuts ÷ Cuts in a graph forms a submodular function. So mincut can be found using submodular minimization.

2.) Matroid intersection ÷ $M_1 = (S, I_1)$, $M_2 = (S, I_2)$

$$f(U) = r_{M_1}(U) + r_{M_2}(S \setminus U)$$

f is a submodular function.

Minimizing f would give us max independent set in $I_1 \cap I_2$.