

EE595A – Submodular functions, their optimization and applications – Spring 2011

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Spring Quarter, 2011

http://ssli.ee.washington.edu/~bilmes/ee595a_spring_2011/

Lecture 18 - June 1st, 2011

Announcements

- Last lecture, and final presentations, will take place Thursday, June 9th, from 3-7:30pm. The lecture will be from 3:00-5:00pm, and the final presentations will be from 5:00-7:30pm. Please bring dinner.
- Today: short lecture (due to many deadlines this week).

Class Road Map

We need to find one makeup lecture this term.

- L1 (3/30):
- L2 (4/1):
- L3 (4/6):
- L4 (4/8):
- L5 (4/13):
- L6 (4/15):
- L7 (4/20):
- L8 (4/27):
- L9 (4/29):
- L10 (5/4):
- L11 (5/6): On SFM, polymatroid member & greedy, Lovász ext.
- L12 (5/11): Lovász ext. + polymatroid props.
- L13 (5/13): More polymatroids, start lattices
- L14 (5/18): lattices/submodular
- L15 (5/20): lattices, → SFM.
- L16 (5/25): → SFM
- L17 (5/27): dep/sat
- L18 (6/1): exchange capacities
- L19 (6/3):
- L20: (6/9): 3-7:30pm (EEB-303)?

dep and partial order

- We have

dep and partial order

- We have

Theorem 2.1

If $x \in P_f$ is an extreme point, then \preceq is a partial order on $\text{sat}(x)$ where for $a, e \in \text{sat}(x)$, the order \preceq is defined by: $a \preceq e$ iff $a \in \text{dep}(x, e)$.

dep and partial order

- We have

Theorem 2.1

If $x \in P_f$ is an extreme point, then \preceq is a partial order on $\text{sat}(x)$ where for $a, e \in \text{sat}(x)$, the order \preceq is defined by: $a \preceq e$ iff $a \in \text{dep}(x, e)$.

- In fact, we have a stronger result that extreme points are characterized by this construct:

dep and partial order

- We have

Theorem 2.1

If $x \in P_f$ is an extreme point, then \preceq is a partial order on $\text{sat}(x)$ where for $a, e \in \text{sat}(x)$, the order \preceq is defined by: $a \preceq e$ iff $a \in \text{dep}(x, e)$.

- In fact, we have a stronger result that extreme points are characterized by this construct:

Theorem 2.2

$x \in P_f$ is an extreme point, iff $\text{supp}(x) \subseteq \text{sat}(x)$ and $\text{dep}(x, a) \neq \text{dep}(x, b)$ for every pair of distinct points $a, b \in \text{sat}(x)$.

the partial order of extreme points

Theorem 2.3

Let x be an extreme point of P_f and \preceq be its partial order. Let $B \subseteq E$ be an ordered set. Then B generates x using the greedy algorithm iff we have $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$ and B is compatible with \preceq .

Corollary 2.4

If x is an extreme point of P_f and $B \subseteq E$ is given such that $\text{supp}(x) \subseteq B \subseteq \text{sat}(x)$, then x is generated using greedy by some ordering of B .

Extreme point testing and partial order generation

input : Vector $x \in \mathbb{R}^E$, polymatroid function f on E .

output: That x is not extreme point, or if it is, minimal tight sets $\text{dep}(x, e)$ for $e \in \text{sat}(x)$ thus defining \preceq . Moreover, $\text{dep}(x, e_j) = A_j$ for $1 \leq j \leq n$ where $n = |\text{sat}(x)|$.

$j \leftarrow 0$; $B \leftarrow \emptyset$;

while true do

$j \leftarrow j + 1$;

if $\exists e \in E \setminus B$ with $x(B + e) = f(B + e)$ **then**

$B \leftarrow B + e$; $e_j \leftarrow e$;

else

 STOP, if $\text{supp}(x) \subseteq B$ then x is extreme, otherwise not.

$A_j \leftarrow B$; $k \leftarrow j - 1$;

while $x(A_j - e_k) = f(A_j - e_k)$ and $k > 0$ **do**

$A_j = A_j - e_k$; $k \leftarrow k - 1$;

On Greedy, and linear programming max

Theorem 2.5

Let $y \in P_f$ be an extreme point, and let \preceq be the partial order of y . Let $c \in \mathbb{R}^E$. Then, y is the solution in:

$$c^T y = \max \{c^T x : x \in P_f\} \quad (1)$$

iff the following three conditions hold:

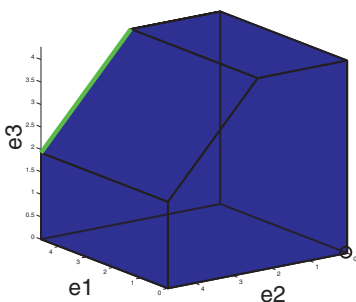
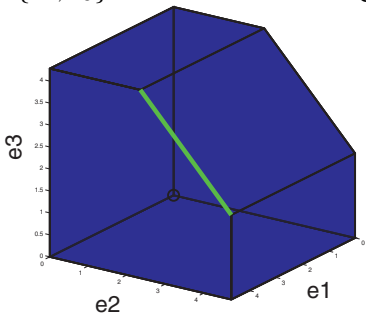
- (1) $c(e) \geq 0$ for every $e \in \text{supp}(y)$
- (2) $c(e) \leq 0$ for every $e \in E \setminus \text{sat}(y)$, and
- (3) For $d, e \in \text{sat}(y)$ and $d \preceq e$ imply that $c(d) \geq c(e)$.

Another revealing theorem

Theorem 2.6

Let f be a polymatroid function and suppose that E can be partitioned into (E_1, E_2, \dots, E_k) such that $f(A) = \sum_{i=1}^k f(A \cap E_i)$ for all $A \subseteq E$, and k is maximum. Then the base polytope $B_f = \{x \in P_f : x(E) = f(E)\}$ (the E -tight subset of P_f) has dimension $|E| - k$.

- Example f with independence between $A = \{e_2, e_3\}$ and $B = \{e_1\}$, i.e., $e_1 \perp\!\!\!\perp \{e_2, e_3\}$, with B_f marked in green.



Base polytope existence and location

- Given polymatroid function f , the base polytope $B_f = \{x \in \mathbb{R}_+^E : x(A) \leq f(A) \forall A \subseteq E, \text{ and } x(E) = f(E)\}$ always exists.

Base polytope existence and location

- Given polymatroid function f , the base polytope $B_f = \{x \in \mathbb{R}_+^E : x(A) \leq f(A) \forall A \subseteq E, \text{ and } x(E) = f(E)\}$ always exists.
- For any $A \subseteq E$, we have

$$B_f \cap \{x \in \mathbb{R}_+^E : x(A) = f(A)\} \neq \emptyset \quad (2)$$

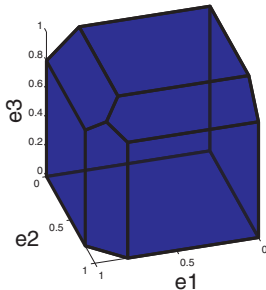
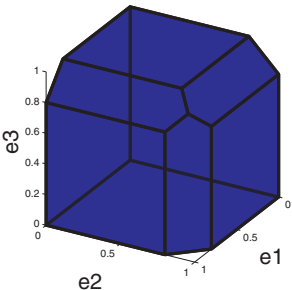
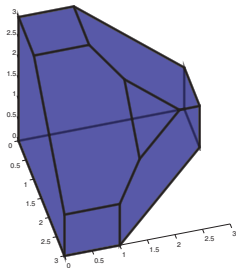
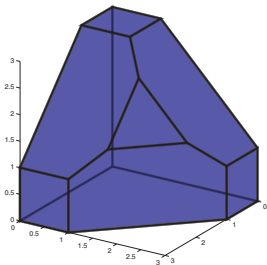
Base polytope existence and location

- Given polymatroid function f , the base polytope $B_f = \{x \in \mathbb{R}_+^E : x(A) \leq f(A) \forall A \subseteq E, \text{ and } x(E) = f(E)\}$ always exists.
- For any $A \subseteq E$, we have

$$B_f \cap \{x \in \mathbb{R}_+^E : x(A) = f(A)\} \neq \emptyset \quad (2)$$

- In words, B_f intersects all “multi-axis orthogonal” subsets of \mathbb{R}_+^E .

Not polymatroidal polytopes



→ SFM on arbitrary submodular g : transformation

- Given any arbitrary submodular function g with the goal of finding $A^* \in \operatorname{argmin}_{A \subseteq E} g(A)$
- We reduce this to:

$$A^* \in \operatorname{argmin}_{A \subseteq E'} \left(f(A) - m(A) \right) \quad (3)$$

where

- f is a polymatroid function on $2^{E'}$
- m is a modular function on $2^{E'}$ with $m \in \mathbb{R}_+^{E'}$.
- $E' \subseteq E$.
- In the sequel, we assume this form, with ground set E .
- Moreover, we may assume that P_f is a polymatroidal polytope, with $P_f \subset \mathbb{R}_+^E$.

A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

Theorem 2.7

Let f be a polymatroid function defined on subsets of E . For any $x \in \mathbb{R}_+^E$, then

$$\max (y(E) : y \leq x, y \in P_f) = \min (f(A) + x(E \setminus A) : A \subseteq E) \quad (4)$$

A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

Theorem 2.7

Let f be a polymatroid function defined on subsets of E . For any $x \in \mathbb{R}_+^E$, then

$$\max (y(E) : y \leq x, y \in P_f) = \min (f(A) + x(E \setminus A) : A \subseteq E) \quad (4)$$

- Thus, this can act as a certificate of optimality for any submodular function minimization problem on g even if g is not polymatroidal.

A characterization of the optimality of the SFM problem

- We proved in Lecture 7 (and again, in Lecture 11) the following theorem:

Theorem 2.7

Let f be a polymatroid function defined on subsets of E . For any $x \in \mathbb{R}_+^E$, then

$$\max (y(E) : y \leq x, y \in P_f) = \min (f(A) + x(E \setminus A) : A \subseteq E) \quad (4)$$

- Thus, this can act as a certificate of optimality for any submodular function minimization problem on g even if g is not polymatroidal.
- We need only find a feasible y on the max (left) side, and an A^* on the min (right) side that achieves equality, then A^* is a SFM solution in $A^* \in \operatorname{argmin}_{A \subseteq E} g(A)$ where x is the aforementioned modular function, and $f(A) = g(A) + m(A) - g(\emptyset)$.

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.
- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of P_f (each of which may be generated by greedy).

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.
- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of P_f (each of which may be generated by greedy).
- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.
- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of P_f (each of which may be generated by greedy).
- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.
- That is, let I be an index set, and $x^{(i)}$ be an extreme point of P_f for $i \in I$. We then keep y as

$$y = \sum_{i \in I} \lambda_i x^{(i)} \quad (5)$$

where λ_i are convex coefficients.

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.
- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of P_f (each of which may be generated by greedy).
- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.
- That is, let I be an index set, and $x^{(i)}$ be an extreme point of P_f for $i \in I$. We then keep y as

$$y = \sum_{i \in I} \lambda_i x^{(i)} \quad (5)$$

where λ_i are convex coefficients.

- At each step of the algorithm, we either find a larger y , or demonstrate y 's optimality by finding a minimizing A .

Maximizing y

- The nature of SFM will be very similar to the Edmonds's matroid partition problem (recall, asking if E can be partitioned into $\{I_i\}$ each independent in a matroid M_i) and the core algorithm is very similar.
- Now, from convex polytope theory, any $x \in P_f$ can be represented as a convex combination of at most $|E| + 1$ extreme points of P_f (each of which may be generated by greedy).
- We keep a feasible solution to the max version of the problem as a convex combination of such extreme points.
- That is, let I be an index set, and $x^{(i)}$ be an extreme point of P_f for $i \in I$. We then keep y as

$$y = \sum_{i \in I} \lambda_i x^{(i)} \quad (5)$$

where λ_i are convex coefficients.

- At each step of the algorithm, we either find a larger y , or demonstrate y 's optimality by finding a minimizing A .
- Start with $y = 0$, $I = \{1\}$, $\lambda_1 = 1$, and $v^{(1)} = 0$.

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \quad (6)$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \quad (6)$$

- Identical to:

$$\max \{ \alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\} \} \quad (7)$$

since $B \subseteq E$ such that $e \notin B$ have the same value

$$(x + \alpha \mathbf{1}_e)(B) = x(B).$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \quad (6)$$

- Identical to:

$$\max \{ \alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\} \} \quad (7)$$

since $B \subseteq E$ such that $e \notin B$ have the same value

$$(x + \alpha \mathbf{1}_e)(B) = x(B).$$

- Again identical to:

$$\max \{ \alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\} \} \quad (8)$$

or

$$\max \{ \alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\} \} \quad (9)$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \quad (6)$$

- Identical to:

$$\max \{ \alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\} \} \quad (7)$$

since $B \subseteq E$ such that $e \notin B$ have the same value

$$(x + \alpha \mathbf{1}_e)(B) = x(B).$$

- Again identical to:

$$\max \{ \alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\} \} \quad (8)$$

or

$$\max \{ \alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\} \} \quad (9)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\} \} \quad (10)$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \quad (6)$$

- Identical to:

$$\max \{ \alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\} \} \quad (7)$$

since $B \subseteq E$ such that $e \notin B$ have the same value

$$(x + \alpha \mathbf{1}_e)(B) = x(B).$$

- Again identical to:

$$\max \{ \alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\} \} \quad (8)$$

or

$$\max \{ \alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\} \} \quad (9)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\} \} \quad (10)$$

- $\hat{c}(x; e)$ is known as the **saturation capacity** associated with $x \in P_f$ and e .

Saturation Capacity

- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (11)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (12)$$

Saturation Capacity

- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (11)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (12)$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.

Saturation Capacity

- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (11)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (12)$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, for $e \in \text{sat}(x)$, we have that $\hat{c}(x; e) = 0$.

Saturation Capacity

- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (11)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (12)$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, for $e \in \text{sat}(x)$, we have that $\hat{c}(x; e) = 0$.
- Note that any α with $0 \leq \alpha \leq \hat{c}(x; e)$ we have $x + \alpha \mathbf{1}_e \in P_f$.

Saturation Capacity

- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (11)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (12)$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, for $e \in \text{sat}(x)$, we have that $\hat{c}(x; e) = 0$.
- Note that any α with $0 \leq \alpha \leq \hat{c}(x; e)$ we have $x + \alpha \mathbf{1}_e \in P_f$.
- We also see that computing $\hat{c}(x; e)$ is a form of submodular function minimization.

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.
- Thus, for any $\alpha > 0$, we have $x + \alpha \mathbf{1}_a \notin P_f$ for $a = e$ and $a = e'$.

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.
- Thus, for any $\alpha > 0$, we have $x + \alpha \mathbf{1}_a \notin P_f$ for $a = e$ and $a = e'$.
- Consider

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f \} \quad (13)$$

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.
- Thus, for any $\alpha > 0$, we have $x + \alpha \mathbf{1}_a \notin P_f$ for $a = e$ and $a = e'$.
- Consider

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f \} \quad (13)$$

- Identical to:

$$\max \{ \alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \} \quad (14)$$

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.
- Thus, for any $\alpha > 0$, we have $x + \alpha \mathbf{1}_a \notin P_f$ for $a = e$ and $a = e'$.
- Consider

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f \} \quad (13)$$

- Identical to:

$$\max \{ \alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \} \quad (14)$$

- Note that if both $e, e' \in A$, then $\alpha(\mathbf{1}_e - \mathbf{1}_{e'})(A) = 0$ for any α , so to make this meaningful, we take $A : e' \notin A \supseteq \{e\}$, thus identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (15)$$

Exchange Capacity

- Now consider $x \in P_f$, $e \in \text{sat}(x)$ and $e' \in \text{dep}(x, e) \setminus \{e\}$
- recall that $\text{dep}(x, e) \setminus \{e\}$ is tight for $e \in \text{sat}(x)$, so $x(e') > 0$ for $e' \in \text{dep}(x, e) \setminus \{e\}$.
- Thus, for any $\alpha > 0$, we have $x + \alpha \mathbf{1}_a \notin P_f$ for $a = e$ and $a = e'$.
- Consider

$$\max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f \} \quad (13)$$

- Identical to:

$$\max \{ \alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \} \quad (14)$$

- Note that if both $e, e' \in A$, then $\alpha(\mathbf{1}_e - \mathbf{1}_{e'})(A) = 0$ for any α , so to make this meaningful, we take $A : e' \notin A \supseteq \{e\}$, thus identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (15)$$

- Which is identical to:

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha(\mathbf{1}_e - \mathbf{1}_{e'})(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (16)$$

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (19)$$

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (19)$$

- $\hat{c}(x; e, e')$ is known as the **exchange capacity** associated with $x \in P_f$ and e .

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (19)$$

- $\hat{c}(x; e, e')$ is known as the **exchange capacity** associated with $x \in P_f$ and e .
- For any α with $0 \leq \alpha \leq \hat{c}(x; e, e')$, we have that $x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f$.

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (19)$$

- $\hat{c}(x; e, e')$ is known as the **exchange capacity** associated with $x \in P_f$ and e .
- For any α with $0 \leq \alpha \leq \hat{c}(x; e, e')$, we have that $x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f$.
- As we will see, if e and e' are successive in an order that generates extreme point x , then we get a “neighbor” extreme point via $x' = x + \hat{c}(x; e, e')(\mathbf{1}_e - \mathbf{1}_{e'})$.

Exchange Capacity

- In such case, we get $\mathbf{1}_{e'}(A) = 0$, thus above identical to

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (17)$$

- Restating, we've got

$$\max \{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (18)$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \} \quad (19)$$

- $\hat{c}(x; e, e')$ is known as the **exchange capacity** associated with $x \in P_f$ and e .
- For any α with $0 \leq \alpha \leq \hat{c}(x; e, e')$, we have that $x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f$.
- As we will see, if e and e' are successive in an order that generates extreme point x , then we get a “neighbor” extreme point via $x' = x + \hat{c}(x; e, e')(\mathbf{1}_e - \mathbf{1}_{e'})$.
- Note that Eq. 19 is a form of SFM.

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets
 $\mathcal{D}(x) = \{A : x(A) = f(A)\}$

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets
 $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.
- Consider the “0” element of $\mathcal{D}(x)$, i.e., $\text{dry}(x) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x)\}$

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.
- Consider the “0” element of $\mathcal{D}(x)$, i.e., $\text{dry}(x) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x)\}$
- We can see $\text{dry}(x)$ as the elements that are necessary for tightness.

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.
- Consider the “0” element of $\mathcal{D}(x)$, i.e., $\text{dry}(x) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x)\}$
- We can see $\text{dry}(x)$ as the elements that are necessary for tightness.
- That is, we can view $\text{dry}(x)$ as

$$\text{dry}(x) = \{e' : x(A) < f(A), \forall A \not\supseteq e'\} \quad (20)$$

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.
- Consider the “0” element of $\mathcal{D}(x)$, i.e., $\text{dry}(x) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x)\}$
- We can see $\text{dry}(x)$ as the elements that are necessary for tightness.
- That is, we can view $\text{dry}(x)$ as

$$\text{dry}(x) = \{e' : x(A) < f(A), \forall A \not\supseteq e'\} \quad (20)$$

- Perhaps, then, a better name for dry is $\text{nsat}(x)$, for the necessary for tightness.

dep revisited

- Given $x \in P_f$, recall distributive lattice of tight sets
 $\mathcal{D}(x) = \{A : x(A) = f(A)\}$
- We had that $\text{sat}(x) = \bigcup \{A : A \in \mathcal{D}(x)\}$ is the “1” element of this lattice.
- Consider the “0” element of $\mathcal{D}(x)$, i.e., $\text{dry}(x) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x)\}$
- We can see $\text{dry}(x)$ as the elements that are necessary for tightness.
- That is, we can view $\text{dry}(x)$ as

$$\text{dry}(x) = \{e' : x(A) < f(A), \forall A \not\supseteq e'\} \quad (20)$$

- Perhaps, then, a better name for dry is $\text{nsat}(x)$, for the necessary for tightness.
- This can be read as, for any $e' \in \text{dry}(x)$, any set that does not contain e' is not tight for x .

dep revisited

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.
- Analogously, we can define the “0” element of this lattice as $\text{dry}(x, e) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x, e)\}$.

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.
- Analogously, we can define the “0” element of this lattice as $\text{dry}(x, e) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x, e)\}$.
- We can see $\text{dry}(x, e)$ as the elements that are necessary for e -containing tightness, with $e \in \text{sat}(x)$.

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.
- Analogously, we can define the “0” element of this lattice as $\text{dry}(x, e) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x, e)\}$.
- We can see $\text{dry}(x, e)$ as the elements that are necessary for e -containing tightness, with $e \in \text{sat}(x)$.
- That is, we can view $\text{dry}(x, e)$ as

$$\text{dry}(x, e) = \{e' : x(A) < f(A), \forall A \not\supseteq e', e \in A\} \quad (21)$$

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.
- Analogously, we can define the “0” element of this lattice as $\text{dry}(x, e) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x, e)\}$.
- We can see $\text{dry}(x, e)$ as the elements that are necessary for e -containing tightness, with $e \in \text{sat}(x)$.
- That is, we can view $\text{dry}(x, e)$ as

$$\text{dry}(x, e) = \{e' : x(A) < f(A), \forall A \not\ni e', e \in A\} \quad (21)$$
- This can be read as, for any $e' \in \text{dry}(x, e)$, any e -containing set that does not contain e' is not tight for x .

dep revisited

- Now, given $x \in P_f$, and $e \in \text{sat}(x)$, recall distributive lattice of e -containing tight sets $\mathcal{D}(x, e) = \{A : e \in A, x(A) = f(A)\}$
- We can define the “1” element of this lattice as $\text{sat}(x, e) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x, e)\}$.
- Analogously, we can define the “0” element of this lattice as $\text{dry}(x, e) \stackrel{\text{def}}{=} \bigcap \{A : A \in \mathcal{D}(x, e)\}$.
- We can see $\text{dry}(x, e)$ as the elements that are necessary for e -containing tightness, with $e \in \text{sat}(x)$.
- That is, we can view $\text{dry}(x, e)$ as

$$\text{dry}(x, e) = \{e' : x(A) < f(A), \forall A \not\ni e', e \in A\} \quad (21)$$
- This can be read as, for any $e' \in \text{dry}(x, e)$, any e -containing set that does not contain e' is not tight for x .
- Notice also that $\text{dry}(x, e) = \text{dep}(x, e)$.

dep revisited

- Now, we have the following equalities for $\text{dep}(x, e)$:

$$\text{dep}(x, e) = \{e' : x(A) < f(A), \forall A \not\cong e', e \in A\} \quad (22)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \leq f(A) - x(A), \forall A \not\cong e', e \in A\} \quad (23)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \not\cong e', e \in A\} \quad (24)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \not\cong e', e \in A\} \quad (25)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \not\cong e', e \in A\} \quad (26)$$

dep revisited

- Now, we have the following equalities for $\text{dep}(x, e)$:

$$\text{dep}(x, e) = \{e' : x(A) < f(A), \forall A \ni e', e \in A\} \quad (22)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \leq f(A) - x(A), \forall A \ni e', e \in A\} \quad (23)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \ni e', e \in A\} \quad (24)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \ni e', e \in A\} \quad (25)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \ni e', e \in A\} \quad (26)$$

- Now, $\mathbf{1}_e(A) - \mathbf{1}_{e'}(A) = 0$ if either $\{e, e'\} \subseteq A$, or $\{e, e'\} \cap A = \emptyset$.

dep revisited

- Now, we have the following equalities for $\text{dep}(x, e)$:

$$\text{dep}(x, e) = \{e' : x(A) < f(A), \forall A \not\subseteq e', e \in A\} \quad (22)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \leq f(A) - x(A), \forall A \not\subseteq e', e \in A\} \quad (23)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \not\subseteq e', e \in A\} \quad (24)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \not\subseteq e', e \in A\} \quad (25)$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \not\subseteq e', e \in A\} \quad (26)$$

- Now, $\mathbf{1}_e(A) - \mathbf{1}_{e'}(A) = 0$ if either $\{e, e'\} \subseteq A$, or $\{e, e'\} \cap A = \emptyset$.
- Also, if $e' \in A$ but $e \notin A$, then $x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) = x(A) - \alpha \leq f(A)$ since $x \in P_f$.

dep revisited

- thus, we get the same in the above if we remove the constraint

$A \not\supseteq e', e \in A$, that is we get

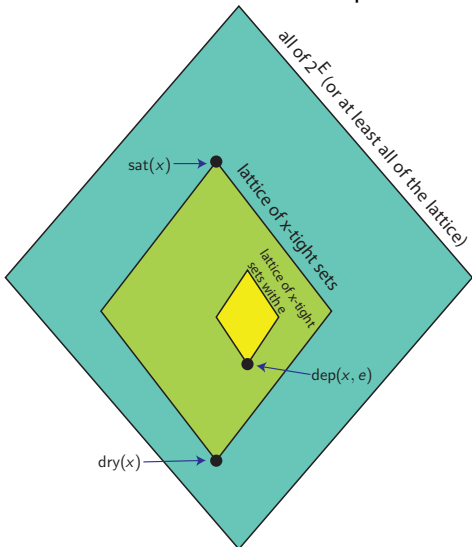
$$\text{dep}(x, e) = \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A\} \quad (27)$$

- This is then identical to

$$\text{dep}(x, e) = \{e' : \exists \alpha > 0, \text{ s.t. } x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\} \quad (28)$$

dep and sat

The following picture summarizes the relationships.



From vertex to vertex

- We will need to move from one extreme point to another (adjacent) extreme point, and will use an augmenting path like approach to do so.
- How do we characterize such adjacent extreme points?

From vertex to vertex

Theorem 3.1

Let x be an extreme point of P_f , and let \preceq be its partial order. Then, each of the following three operations will yield a new extreme point w :

- (a) Let $a, b \in E$ and a cover b relative to \preceq , so $b \sqsubset a$. Let $w = x + \alpha \mathbf{1}_a - \alpha \mathbf{1}_b$ with $\alpha = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.

From vertex to vertex

Theorem 3.1

Let x be an extreme point of P_f , and let \preceq be its partial order. Then, each of the following three operations will yield a new extreme point w :

- (a) Let $a, b \in E$ and a cover b relative to \preceq , so $b \sqsubset a$. Let $w = x + \alpha \mathbf{1}_a - \alpha \mathbf{1}_b$ with $\alpha = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.
- (b) Let $a \in E \setminus \text{sat}(x)$, and let $w = x + \alpha \mathbf{1}_a$ where $\alpha = f(\text{sat}(x) + a) - f(\text{sat}(x))$.

From vertex to vertex

Theorem 3.1

Let x be an extreme point of P_f , and let \preceq be its partial order. Then, each of the following three operations will yield a new extreme point w :

- (a) Let $a, b \in E$ and a cover b relative to \preceq , so $b \sqsubset a$. Let $w = x + \alpha \mathbf{1}_a - \alpha \mathbf{1}_b$ with $\alpha = f(\text{dep}(x, a) - b) - x(\text{dep}(x, a) - b)$.
- (b) Let $a \in E \setminus \text{sat}(x)$, and let $w = x + \alpha \mathbf{1}_a$ where $\alpha = f(\text{sat}(x) + a) - f(\text{sat}(x))$.
- (c) Let $a \in \text{supp}(x)$ be maximal (w.r.t. \preceq), and let $w = x - x(a) \mathbf{1}_a$.

B_f dominates

Lemma 3.2

Let $x \in P_f$ and let $T = \text{sat}(x)$. Then there exists $y \in B_f$ such that $y \geq x$ with $y(e) = x(e)$ for $e \in T$.

Proof.

- Consider a form of the greedy procedure, where we update x

B_f dominates

Lemma 3.2

Let $x \in P_f$ and let $T = \text{sat}(x)$. Then there exists $y \in B_f$ such that $y \geq x$ with $y(e) = x(e)$ for $e \in T$.

Proof.

- Consider a form of the greedy procedure, where we update x
- Iterate the following procedure, for any $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e)\mathbf{1}_e \quad (29)$$

B_f dominates

Lemma 3.2

Let $x \in P_f$ and let $T = \text{sat}(x)$. Then there exists $y \in B_f$ such that $y \geq x$ with $y(e) = x(e)$ for $e \in T$.

Proof.

- Consider a form of the greedy procedure, where we update x
- Iterate the following procedure, for any $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e)\mathbf{1}_e \quad (29)$$

- Thus, after x update, e , we still have $x \in P_f$.

B_f dominates

Lemma 3.2

Let $x \in P_f$ and let $T = \text{sat}(x)$. Then there exists $y \in B_f$ such that $y \geq x$ with $y(e) = x(e)$ for $e \in T$.

Proof.

- Consider a form of the greedy procedure, where we update x
- Iterate the following procedure, for any $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e)\mathbf{1}_e \quad (29)$$

- Thus, after x update, e , we still have $x \in P_f$.
- Moreover, at each update there is a set S_e that achieves the min in the min form of $\hat{c}(x; e)$. This set S_e is tight for the new x and remains tight for all subsequent iterations.

B_f dominates

Lemma 3.2

Let $x \in P_f$ and let $T = \text{sat}(x)$. Then there exists $y \in B_f$ such that $y \geq x$ with $y(e) = x(e)$ for $e \in T$.

Proof.

- Consider a form of the greedy procedure, where we update x
- Iterate the following procedure, for any $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e)\mathbf{1}_e \quad (29)$$

- Thus, after x update, e , we still have $x \in P_f$.
- Moreover, at each update there is a set S_e that achieves the min in the min form of $\hat{c}(x; e)$. This set S_e is tight for the new x and remains tight for all subsequent iterations.
- Eventually we stop, and since $E = T \cup \bigcup_{e \notin T} S_e$ is the union of tight sets (for x), we see that the resulting x has $x \in B_f$.

Scratch Paper

Scratch Paper

Scratch Paper

Sources for Today's Lecture

- Bixby, Cunningham, Topkis, "The Partial Order of a Polymatroid Extreme Point", 1985.
- J. Edmonds, "Submodular Functions, Matroids, and Certain Polyhedra", 1970.
- Lovász, "Submodular Functions and Convexity", 1983.