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 19 - June 3st, 2011

Announcements

- Last lecture, and final presentations, will take place Thursday, June 9th, from 3-7:30pm in room EEB-303.
- The lecture will be from 3:00-5:00pm,
- The final presentations will be from 5:00-7:30pm. Feel free to bring dinner.
- Final project reports due next Wednesday, June 8th, at 11:45pm (on the web page, dropbox will be posted shortly).
- Final slides due by 1:00pm on Thursday, also via the dropbox.
- Final talks: Plan for 10 minutes of talking about your project (perhaps 8 minutes of lecture and 2 minutes of questions).

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): SFM algorithm
- L20: (6/9): 3-7:30pm (EEB-303)?

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

$$y = \sum_{i \in I} \lambda_i x^{(i)} \tag{1}$$

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

$$y = \sum_{i \in I} \lambda_i x^{(i)} \tag{1}$$

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
$$y = \sum_{i \in I} \lambda_i x^{(i)} \tag{1}$$
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

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{2}$$

Saturation Capacity

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

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{2}$$

- Identical to:

$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{3}$$

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

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{2}$$

- Identical to:

$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{3}$$

since $B \subseteq E$ such that $e \notin B$ have the same value $(x + \alpha \mathbf{1}_e)(B) = x(B)$.

- Again identical to:

$$\max \{\alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\}\} \tag{4}$$

or

$$\max \{\alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}\} \tag{5}$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding
$$\max\{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{2}$$
- Identical to:
$$\max\{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{3}$$
since $B \subseteq E$ such that $e \notin B$ have the same value $(x + \alpha \mathbf{1}_e)(B) = x(B)$.
- Again identical to:
$$\max\{\alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\}\} \tag{4}$$
or
$$\max\{\alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}\} \tag{5}$$
- This max is achieved when
$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min\{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{6}$$

Saturation Capacity

- For $x \in P_f$, and $e \in E$, consider finding
$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{2}$$
- Identical to:
$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{3}$$
since $B \subseteq E$ such that $e \notin B$ have the same value $(x + \alpha \mathbf{1}_e)(B) = x(B)$.
- Again identical to:
$$\max \{\alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\}\} \tag{4}$$
or
$$\max \{\alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}\} \tag{5}$$
- This max is achieved when
$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{6}$$
- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.

Saturation Capacity

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

$$\hat{c}(x;e) \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{e\}\} \tag{7}$$

$$= \max \{ \alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f \} \tag{8}$$

Saturation Capacity

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

$$\hat{c}(x;e) \stackrel{\text{def}}{=} \min\{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (7)$$

$$= \max\{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (8)$$

- 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 $x \in P_f$,

$$\hat{c}(x;e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \qquad (7)$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \qquad (8)$$

- 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 $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (7)$$
$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (8)$$

- 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 $x \in P_f$,
$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \quad (7)$$
$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \quad (8)$$
- 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 $x + \alpha \mathbf{1}_e \notin P_f$.

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 $x + \alpha \mathbf{1}_e \notin P_f$.
- Consider

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\} \tag{9}$$

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 $x + \alpha \mathbf{1}_e \notin P_f$.
- Consider

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\} \tag{9}$$

- Identical to:

$$\max \{\alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A\} \tag{10}$$

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 $x + \alpha \mathbf{1}_e \notin P_f$.
- Consider

$$\max\{\alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\} \tag{9}$$

- Identical to:

$$\max\{\alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A\} \tag{10}$$

- 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

$$\max\left\{\alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{11}$$

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 $x + \alpha \mathbf{1}_e \notin P_f$.
- Consider

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\} \tag{9}$$

- Identical to:

$$\max \{\alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A\} \tag{10}$$

- 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

$$\max \left\{\alpha : \alpha \in \mathbb{R}, (x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{11}$$

- Which is identical to:

$$\max \left\{\alpha : \alpha \in \mathbb{R}, \alpha(\mathbf{1}_e - \mathbf{1}_{e'}))(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{12}$$

Exchange Capacity

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

$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{13}$$

Exchange Capacity

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

$$\max \left\{ \alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \right\} \tag{13}$$

- Restating, we've got

$$\max \left\{ \alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A \right\} \tag{14}$$

Exchange Capacity

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

$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{13}$$

- Restating, we've got

$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{14}$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min\left\{f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{15}$$

Exchange Capacity

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

$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{13}$$

- Restating, we've got

$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{14}$$

- This max is achieved when

$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min\left\{f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{15}$$

- $\hat{c}(x; e, e')$ 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
$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{13}$$
- Restating, we've got
$$\max\left\{\alpha : \alpha \in \mathbb{R}, \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{14}$$
- This max is achieved when
$$\alpha = \hat{c}(x; e, e') \stackrel{\text{def}}{=} \min\left\{f(A) - x(A), \forall A \supseteq \{e\}, e' \notin A\right\} \tag{15}$$
- $\hat{c}(x; e, e')$ 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 $x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f$.

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

$$\text{dry}(x) = \left\{e' : x(A) < f(A), \forall A \not\ni e'\right\} \tag{16}$$

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

$$\text{dry}(x) = \left\{e' : x(A) < f(A), \forall A \not\ni e'\right\} \tag{16}$$

- 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

$$\text{dry}(x) = \left\{e' : x(A) < f(A), \forall A \not\ni e'\right\} \tag{16}$$

- 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

- 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

$$\text{dry}(x, e) = \left\{e' : x(A) < f(A), \forall A \not\ni e', e \in A\right\} \tag{17}$$

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

$$\text{dry}(x, e) = \left\{e' : x(A) < f(A), \forall A \not\ni e', e \in A\right\} \tag{17}$$

- 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

$$\text{dry}(x, e) = \left\{e' : x(A) < f(A), \forall A \not\ni e', e \in A\right\} \tag{17}$$

- 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 $\text{dep}(x, e)$:

$$\text{dep}(x, e) = \{e' : x(A) < f(A), \forall A \not\ni e', e \in A\} \tag{18}$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{19}$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{20}$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{21}$$

$$= \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \not\ni e', e \in A\} \tag{22}$$

dep revisited

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

$$\text{dep}(x,e) = \{e' : x(A) < f(A), \forall A \not\ni e', e \in A\} \tag{18}$$
$$= \{e' : \exists \alpha > 0,\ \text{s.t. } \alpha \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{19}$$
$$= \{e' : \exists \alpha > 0,\ \text{s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{20}$$
$$= \{e' : \exists \alpha > 0,\ \text{s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{21}$$
$$= \{e' : \exists \alpha > 0,\ \text{s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \not\ni e', e \in A\} \tag{22}$$

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

dep revisited

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

$$\text{dep}(x,e) = \{e' : x(A) < f(A), \forall A \not\ni e', e \in A\} \tag{18}$$
$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{19}$$
$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha \mathbf{1}_e(A) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{20}$$
$$= \{e' : \exists \alpha > 0, \text{ s.t. } \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A) - x(A), \forall A \not\ni e', e \in A\} \tag{21}$$
$$= \{e' : \exists \alpha > 0, \text{ s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A \not\ni e', e \in A\} \tag{22}$$

- Now, $1_e(A) - \mathbf{1}_{e'}(A) = 0$ 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 $A \not\ni e', e \in A$, that is we get

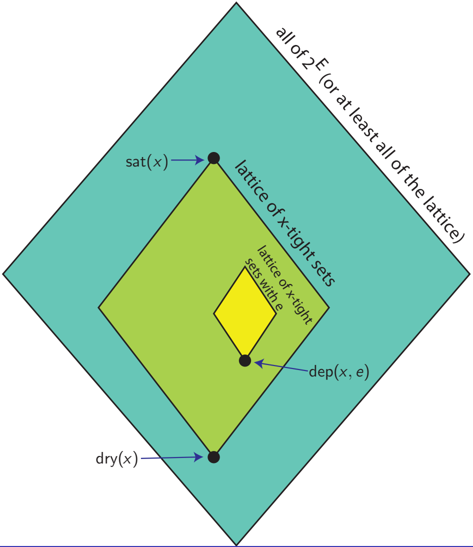
$$\text{dep}(x, e) = \left\{e' : \exists \alpha > 0,\ \text{s.t. } x(A) + \alpha(\mathbf{1}_e(A) - \mathbf{1}_{e'}(A)) \leq f(A), \forall A\right\} \quad (23)$$

- This is then identical to

$$\text{dep}(x, e) = \left\{e' : \exists \alpha > 0,\ \text{s.t. } x + \alpha(\mathbf{1}_e - \mathbf{1}_{e'}) \in P_f\right\} \quad (24)$$

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 $w = x - x(a)\mathbf{1}_a$.*

From Vertex to Vertex

- For (a), let x be generated by $E_i = (e_1, e_2, \ldots, e_{k-1}, b, a, e_{k+2}, \ldots, e_i)$ and consider generating w with an order with a and b swapped, i.e., $E_i' = (e_1, e_2, \ldots, e_{k-1}, a, b, e_{k+2}, \ldots, e_i)$

From Vertex to Vertex

- For (a), let x be generated by $E_i = (e_1, e_2, \ldots, e_{k-1}, b, a, e_{k+2}, \ldots, e_i)$ and consider generating w with an order with a and b swapped, i.e., $E_i' = (e_1, e_2, \ldots, e_{k-1}, a, b, e_{k+2}, \ldots, e_i)$
- Then

$$x(b) = f(e_1, \ldots, e_{k-1}, b) - f(e_1, \ldots, e_{k-1}) \tag{25}$$

$$x(a) = f(e_1, \ldots, e_{k-1}, b, a) - f(e_1, \ldots, e_{k-1}, b) \tag{26}$$

$$w(b) = f(e_1, \ldots, e_{k-1}, a, b) - f(e_1, \ldots, e_{k-1}, a) \tag{27}$$

$$w(a) = f(e_1, \ldots, e_{k-1}, a) - f(e_1, \ldots, e_{k-1}) \tag{28}$$

From Vertex to Vertex

- For (a), let x be generated by $E_i = (e_1, e_2, \ldots, e_{k-1}, b, a, e_{k+2}, \ldots, e_i)$ and consider generating w with an order with a and b swapped, i.e., $E'_i = (e_1, e_2, \ldots, e_{k-1}, a, b, e_{k+2}, \ldots, e_i)$
- Then

$$x(b) = f(e_1, \ldots, e_{k-1}, b) - f(e_1, \ldots, e_{k-1}) \tag{25}$$
$$x(a) = f(e_1, \ldots, e_{k-1}, b, a) - f(e_1, \ldots, e_{k-1}, b) \tag{26}$$
$$w(b) = f(e_1, \ldots, e_{k-1}, a, b) - f(e_1, \ldots, e_{k-1}, a) \tag{27}$$
$$w(a) = f(e_1, \ldots, e_{k-1}, a) - f(e_1, \ldots, e_{k-1}) \tag{28}$$

- First, we have $(w - x)(e) = 0$ for all $e \notin \{a, b\}$ since in each case the differences are the same.

From Vertex to Vertex

- For a we have:

$$(w - x)(a) = f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1}) \tag{29}$$

$$- f(e_1, \dots, e_{k-1}, a, b) + f(e_1, \dots, e_{k-1}, b) \tag{30}$$

From Vertex to Vertex

- For a we have:

$$
\begin{aligned}
(w-x)(a) &= f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) && (29)\\
&\quad - f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) && (30)
\end{aligned}
$$

- For b, we have

$$
\begin{aligned}
(w-x)(b) &= f(e_1,\dots,e_{k-1},a,b) - f(e_1,\dots,e_{k-1},a) && (31)\\
&\quad - f(e_1,\dots,e_{k-1},b) + f(e_1,\dots,e_{k-1}) && (32)
\end{aligned}
$$

From Vertex to Vertex

- For a we have:

$$(w-x)(a) = f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) \tag{29}$$
$$- f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) \tag{30}$$

- For b, we have

$$(w-x)(b) = f(e_1,\dots,e_{k-1},a,b) - f(e_1,\dots,e_{k-1},a) \tag{31}$$
$$- f(e_1,\dots,e_{k-1},b) + f(e_1,\dots,e_{k-1}) \tag{32}$$

- so we immediately see that

$$(w-x)(a) = -(w-x)(b) \tag{33}$$

From Vertex to Vertex

- For a we have:

$$\begin{aligned}(w-x)(a) &= f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) &\quad (29)\\ &\quad - f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) &\quad (30)\end{aligned}$$

- For b, we have

$$\begin{aligned}(w-x)(b) &= f(e_1,\dots,e_{k-1},a,b) - f(e_1,\dots,e_{k-1},a) &\quad (31)\\ &\quad - f(e_1,\dots,e_{k-1},b) + f(e_1,\dots,e_{k-1}) &\quad (32)\end{aligned}$$

- so we immediately see that

$$(w-x)(a) = -(w-x)(b) \tag{33}$$

- So with $\alpha = (w-x)(a)$ we have

$$w = x + \alpha(\mathbf{1}_a - \mathbf{1}_b) \tag{34}$$

From Vertex to Vertex

- Moreover, we see that since

$$\begin{aligned}(w-x)(a) = f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) && (35)\\ - f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) && (36)\\ && (37)\end{aligned}$$

From Vertex to Vertex

- Moreover, we see that since

$$(w-x)(a) = f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) \tag{35}$$
$$- f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) \tag{36}$$
$$\tag{37}$$

- and since by submodularity we have

$$f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) \geq f(e_1,\dots,e_{k-1},a,b) - f(e_1,\dots,e_{k-1},b) \tag{38}$$

From Vertex to Vertex

- Moreover, we see that since

$$(w - x)(a) = f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1}) \quad (35)$$
$$- f(e_1, \dots, e_{k-1}, a, b) + f(e_1, \dots, e_{k-1}, b) \quad (36)$$
$$(37)$$

- and since by submodularity we have

$$f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1})$$
$$\geq f(e_1, \dots, e_{k-1}, a, b) - f(e_1, \dots, e_{k-1}, b) \quad (38)$$

- We have that

$$(w - x)(a) \geq 0 \quad (39)$$

From Vertex to Vertex

- Moreover, we see that since

$$(w - x)(a) = f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1}) \quad (35)$$
$$- f(e_1, \dots, e_{k-1}, a, b) + f(e_1, \dots, e_{k-1}, b) \quad (36)$$
$$(37)$$

- and since by submodularity we have

$$f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1})$$
$$\geq f(e_1, \dots, e_{k-1}, a, b) - f(e_1, \dots, e_{k-1}, b) \quad (38)$$

- We have that

$$(w - x)(a) \geq 0 \quad (39)$$

- Now both x and w are extreme points (generated by greedy algorithm), so any point $\lambda x + (1 - \lambda)w$ is in the base of $\text{supp}(x)$

From Vertex to Vertex

- Moreover, we see that since

$$(w-x)(a) = f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1}) \tag{35}$$
$$- f(e_1,\dots,e_{k-1},a,b) + f(e_1,\dots,e_{k-1},b) \tag{36}$$
$$\tag{37}$$

- and since by submodularity we have

$$f(e_1,\dots,e_{k-1},a) - f(e_1,\dots,e_{k-1})$$
$$\geq f(e_1,\dots,e_{k-1},a,b) - f(e_1,\dots,e_{k-1},b) \tag{38}$$

- We have that

$$(w-x)(a) \geq 0 \tag{39}$$

- Now both x and w are extreme points (generated by greedy algorithm), so any point $\lambda x + (1-\lambda)w$ is in the base of $\text{supp}(x)$
- If all E is used, then each of x and w are in B_f and the convex combination lies on one of the skeletal edges of B_f.

From Vertex to Vertex and exchange capacity

- Therefore, defining

$$\alpha = (w - x)(a) = f(e_1, \ldots, e_{k-1}, a) - f(e_1, \ldots, e_{k-1}) \tag{40}$$
$$- f(e_1, \ldots, e_{k-1}, a, b) + f(e_1, \ldots, e_{k-1}, b) \tag{41}$$
$$\tag{42}$$

we see that $\hat{c}(x; a, b) \geq \alpha$

From Vertex to Vertex and exchange capacity

- Therefore, defining

$$\alpha = (w - x)(a) = f(e_1, \ldots, e_{k-1}, a) - f(e_1, \ldots, e_{k-1}) \tag{40}$$
$$- f(e_1, \ldots, e_{k-1}, a, b) + f(e_1, \ldots, e_{k-1}, b) \tag{41}$$
$$\tag{42}$$

  we see that $\hat{c}(x; a, b) \geq \alpha$

- On the other hand, moving from x to w jumps between two extreme points, and if we moved any less than $\hat{c}(x; a, b)$ this wouldn't be the case.

From Vertex to Vertex and exchange capacity

- Therefore, defining

$$\alpha = (w - x)(a) = f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1}) \tag{40}$$
$$- f(e_1, \dots, e_{k-1}, a, b) + f(e_1, \dots, e_{k-1}, b) \tag{41}$$
$$\tag{42}$$

  we see that $\hat{c}(x; a, b) \geq \alpha$

- On the other hand, moving from x to w jumps between two extreme points, and if we moved any less than $\hat{c}(x; a, b)$ this wouldn't be the case.
- Thus,

$$\alpha = \hat{c}(x; a, b) \stackrel{\text{def}}{=} \min \{ f(A) - x(A), \forall A \supseteq \{a\}, b \notin A \} \tag{43}$$

From Vertex to Vertex and exchange capacity

- Therefore, defining

$$\alpha = (w - x)(a) = f(e_1, \dots, e_{k-1}, a) - f(e_1, \dots, e_{k-1}) \quad (40)$$
$$- f(e_1, \dots, e_{k-1}, a, b) + f(e_1, \dots, e_{k-1}, b) \quad (41)$$
$$(42)$$

  we see that $\hat{c}(x; a, b) \geq \alpha$

- On the other hand, moving from x to w jumps between two extreme points, and if we moved any less than $\hat{c}(x; a, b)$ this wouldn't be the case.
- Thus,

$$\alpha = \hat{c}(x; a, b) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{a\}, b \notin A\} \quad (43)$$

- Interestingly, this particular submodular minimization problem is easy.

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 $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e) \tag{44}$$

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 $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e)\mathbf{1}_e \tag{44}$$

- 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 $e \notin \text{sat}(x)$

$$x \leftarrow x + \hat{c}(x; e) I_e \tag{44}$$

- 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 $e \notin \text{sat}(x)$
$$x \leftarrow x + \hat{c}(x; e) \tag{44}$$
- 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 $c(x; e)$. 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 $x \in B_f$.

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \text{supp}(x)$, we have $x(e) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$ $= \rho_e(\text{dep}(x,e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \text{supp}(x)$, we have $x(e) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \dots, e_i)$ be that order, letting $e = e_k$.

e_k

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \text{supp}(x)$, we have $x(e) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \ldots, e_i)$ be that order, letting $e = e_k$.
- Then $\text{dep}(x, e) = \{e_1, \ldots, e_k\}$

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \text{supp}(x)$, we have $x(e) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \dots, e_i)$ be that order, letting $e = e_k$.
- Then $\text{dep}(x, e) = \{e_1, \dots, e_k\}$
- Since this generates x, we have $x(e) = f(x_1, \dots, x_k) - f(x_1, \dots, x_{k-1}) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \operatorname{supp}(x)$, we have $x(e) = f(\operatorname{dep}(x,e)) - f(\operatorname{dep}(x,e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \dots, e_i)$ be that order, letting $e = e_k$.
- Then $\operatorname{dep}(x,e) = \{e_1, \dots, e_k\}$
- Since this generates x, we have $x(e) = f(x_1, \dots, x_k) - f(x_1, \dots, x_{k-1}) = f(\operatorname{dep}(x,e)) - f(\operatorname{dep}(x,e) \setminus e)$
- Alt proof: We saw earlier that both $\operatorname{dep}(x,e)$ and $\operatorname{dep}(x,e) \setminus e$ are tight, since e is maximal in $\operatorname{dep}(x,e)$ using order $\preceq$ on x.

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$, then for any $e \in \text{supp}(x)$, we have $x(e) = f(\text{dep}(x,e)) - f(\text{dep}(x,e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \ldots, e_i)$ be that order, letting $e = e_k$.
- Then $\text{dep}(x,e) = \{e_1, \ldots, e_k\}$
- Since this generates x, we have $x(e) = f(x_1, \ldots, x_k) - f(x_1, \ldots, x_{k-1}) = f(\text{dep}(x,e)) - f(\text{dep}(x,e) \setminus e)$
- Alt proof: We saw earlier that both $\text{dep}(x,e)$ and $\text{dep}(x,e) \setminus e$ are tight, since e is maximal in $\text{dep}(x,e)$ using order $\preceq$ on x.
- Thus $f(\text{dep}(x,e)) = x(dep(x,e))$ and $f(\text{dep}(x,e) \setminus e) = x(\text{dep}(x,e) \setminus e)$

$x(\text{dep}(x,e) \setminus e) = x(\text{dep}(x,e)) - x(e) = f(\text{dep}(x,e)) - x(e)$

$f(\text{dep}(x,e) \setminus e)$

More on dep

Lemma 3.3

Given any extreme point $x \in P_f$*, then for any* $e \in \text{supp}(x)$*, we have* $x(e) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$

Proof.

- Consider all orderings that generate x and choose the one where e occurs earliest.
- Let $E = (e_1, \dots, e_i)$ be that order, letting $e = e_k$.
- Then $\text{dep}(x, e) = \{e_1, \dots, e_k\}$
- Since this generates x, we have $x(e) = f(x_1, \dots, x_k) - f(x_1, \dots, x_{k-1}) = f(\text{dep}(x, e)) - f(\text{dep}(x, e) \setminus e)$
- Alt proof: We saw earlier that both $\text{dep}(x, e)$ and $\text{dep}(x, e) \setminus e$ are tight, since e is maximal in $\text{dep}(x, e)$ using order $\preceq$ on x.
- Thus $f(\text{dep}(x, e)) = x(dep(x, e))$ and $f(\text{dep}(x, e) \setminus e) = x(\text{dep}(x, e) \setminus e)$
- This gives the same result.

$\square$

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.
- At each iteration, we have a $y = \sum_{j \in I} \lambda_j x^j$ where x^j is an extremal point, and $\{\lambda_j\}_j$ is convex combination.

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.
- At each iteration, we have a $y = \sum_{j \in I} \lambda_j x^j$ where x^j is an extremal point, and $\{\lambda_j\}_j$ is convex combination.
- At each step we find a better y or show that it is maximizing by finding a minimizing A (using Edmond's theorem as certificate of optimality).

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.
- At each iteration, we have a $y = \sum_{j \in I} \lambda_j x^j$ where x^j is an extremal point, and $\{\lambda_j\}_j$ is convex combination.
- At each step we find a better y or show that it is maximizing by finding a minimizing A (using Edmond's theorem as certificate of optimality).
- Initially we take $y = 0$ with $I = \{1\}$, $\lambda_1 = 1$ and $x^1 = 0$.

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.
- At each iteration, we have a $y = \sum_{j \in I} \lambda_j x^j$ where x^j is an extremal point, and $\{\lambda_j\}_j$ is convex combination.
- At each step we find a better y or show that it is maximizing by finding a minimizing A (using Edmond's theorem as certificate of optimality).
- Initially we take $y = 0$ with $I = \{1\}$, $\lambda_1 = 1$ and $x^1 = 0$.
- For distinct $a, b \in E$, define

$$F(a, b) = \left\{ j \in I : a \text{ covers } b \text{ in partial order } \preceq \text{ of } x^j \right\} \tag{45}$$

The directed graph

- Let s, t be distinct elements not in E and define a directed graph similar to the one we used to solve matroid partition.
- Edges may go from s only to nodes in E, they may go between nodes in E, and edges may to to t only from nodes in E.
- At each iteration, we have a $y = \sum_{j \in I} \lambda_j x^j$ where x^j is an extremal point, and $\{\lambda_j\}_j$ is convex combination.
- At each step we find a better y or show that it is maximizing by finding a minimizing A (using Edmond's theorem as certificate of optimality).
- Initially we take $y = 0$ with $I = \{1\}$, $\lambda_1 = 1$ and $x^1 = 0$.
- For distinct $a, b \in E$, define

$$F(a,b) = \left\{ j \in I : a \text{ covers } b \text{ in partial order } \preceq \text{ of } x^j \right\} \tag{45}$$

- This will be used to produce edges from a to b in order to "augment" y via an exchange like operation.

The directed graph

- Similarly, for $a \in E$, define

$$F(a, \cancel{s}\,t) = \{j \in I : a \notin \mathrm{sat}(x^j)\} \tag{46}$$

The directed graph

- Similarly, for $a \in E$, define
$$F(a, s) = \{j \in I : a \notin \text{sat}(x^j)\} \tag{46}$$
- Used to produce edges between a and s to augment y via an "augment" operation (add a new dimension).

The directed graph

- Similarly, for $a \in E$, define
$$F(a,s) = \{j \in I : a \notin \text{sat}(x^j)\} \tag{46}$$
- Used to produce edges between a and s to augment y via an "augment" operation (add a new dimension).
- Recall the key relationship
$$\max\left(y(E) : y \leq x, y \in P_f\right) = \min\left(f(A) + x(E \setminus A) : A \subseteq E\right) \tag{47}$$
where x is the modular function so constructed making f totally normalized.

The directed graph

- Similarly, for $a \in E$, define
$$F(a, s) = \{j \in I : a \notin \text{sat}(x^j)\} \tag{46}$$
- Used to produce edges between a and s to augment y via an "augment" operation (add a new dimension).
- Recall the key relationship
$$\max\left(y(E) : y \leq x, y \in P_f\right) = \min\left(f(A) + x(E \setminus A) : A \subseteq E\right) \tag{47}$$
where x is the modular function so constructed making f totally normalized.
- We want each step to increase y while still being feasible to the l.h.s.

The directed graph

- We now define the edges and capacities of the graph.

The directed graph

- We now define the edges and capacities of the graph.
- For each $a \in E$ with $y(a) < x(a)$, we add an edge (s, a) with capacity $u(s, a) = x(a) - y(a)$.

The directed graph

- We now define the edges and capacities of the graph.
- For each $a \in E$ with $y(a) < x(a)$, we add an edge (s, a) with capacity $u(s, a) = x(a) - y(a)$.
- Compare before (matroid case): we added such an edge if a is not yet in one of the independent sets we are creating.

The directed graph

- We now define the edges and capacities of the graph.
- For each $a \in E$ with $y(a) < x(a)$, we add an edge (s, a) with capacity $u(s, a) = x(a) - y(a)$.
- Compare before (matroid case): we added such an edge if a is not yet in one of the independent sets we are creating.
- For each pair $a, b \in E + t$, that has $F(a, b) \neq \emptyset$, we create an edge (a, b) in the graph with capacity

$$u(a, b) = \sum_{j \in F(a,b)} \lambda_j \underline{\hat{c}(x^j; a, b)} \tag{48}$$

The directed graph

- We now define the edges and capacities of the graph.
- For each $a \in E$ with $y(a) < x(a)$, we add an edge (s, a) with capacity $u(s, a) = x(a) - y(a)$.
- Compare before (matroid case): we added such an edge if a is not yet in one of the independent sets we are creating.
- For each pair $a, b \in E + t$, that has $F(a, b) \neq \emptyset$, we create an edge (a, b) in the graph with capacity

$$u(a, b) = \sum_{j \in F(a,b)} \lambda_j \hat{c}(x^j; a, b) \tag{48}$$

- For $a, b \in E$, $\hat{c}(x^j; a, b)$ is the exchange capacity for a, b under extreme vector x^j.

The directed graph

- We now define the edges and capacities of the graph.
- For each $a \in E$ with $y(a) < x(a)$, we add an edge (s, a) with capacity $u(s, a) = x(a) - y(a)$.
- Compare before (matroid case): we added such an edge if a is not yet in one of the independent sets we are creating.
- For each pair $a, b \in E + t$, that has $F(a, b) \neq \emptyset$, we create an edge (a, b) in the graph with capacity

$$u(a, b) = \sum_{j \in F(a,b)} \lambda_j \hat{c}(x^j; a, b) \tag{48}$$

- For $a, b \in E$, $\hat{c}(x^j; a, b)$ is the exchange capacity for a, b under extreme vector x^j.
- For $a \in E$ and $b = t$, $\hat{c}(x^j; a, b) = \hat{c}(x^j; a)$ is the saturation capacity associated with x^j and a.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \ldots, a_{m+1} = t$.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \ldots, a_{m+1} = t$.
- For each k with $1 \leq k \leq m$, choose $j(k) \in F(a_k, a_{k+1})$.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \ldots, a_{m+1} = t$.
- For each k with $1 \leq k \leq m$, choose $j(k) \in F(a_k, a_{k+1})$.
- For $j \in I$, then let $n(j) = |\{k : j = j(k)\}|$.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \dots, a_{m+1} = t$.
- For each k with $1 \leq k \leq m$, choose $j(k) \in F(a_k, a_{k+1})$.
- For $j \in I$, then let $n(j) = |\{k : j = j(k)\}|$.
- Define $\theta = \min\{\lambda_j / n(j) : n(j) > 0\}$.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \ldots, a_{m+1} = t$.
- For each k with $1 \leq k \leq m$, choose $j(k) \in F(a_k, a_{k+1})$.
- For $j \in I$, then let $n(j) = |\{k : j = j(k)\}|$.
- Define $\theta = \min \{\lambda_j / n(j) : n(j) > 0\}$.
- Define $\epsilon' = \min \left\{\hat{c}(x^{j(k)}; a_k, a_{k+1}) : 1 \leq k \leq m\right\}$.

SFM

Theorem 3.4

If there is an s, t dipath in the graph, then there exists a $y' \in P_f$ with $y \leq y' \leq x$ with $y'(E) > y(E)$. If there is no such dipath, then $y(A) = f(A)$ and $y(E \setminus A) = x(E \setminus A)$ for some $A \subseteq E$.

- We prove this over the next few slides.
- First, suppose that there exists an s, t dipath and let the vertex set be $s = a_0, a_1, \ldots, a_{m+1} = t$.
- For each k with $1 \leq k \leq m$, choose $j(k) \in F(a_k, a_{k+1})$.
- For $j \in I$, then let $n(j) = |\{k : j = j(k)\}|$.
- Define $\theta = \min \{\lambda_j / n(j) : n(j) > 0\}$.
- Define $\epsilon' = \min \left\{\hat{c}(x^{j(k)}; a_k, a_{k+1}) : 1 \leq k \leq m\right\}$.
- Define $\epsilon = \min \{\epsilon', x(a_1) - y(a_1)\}$

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \geq 0$.

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \geq 0$.
- For each $1 \leq k < m$, let $z^k = x^{j(k)} + \epsilon(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \geq 0$.
- For each $1 \leq k < m$, let $z^k = x^{j(k)} + \epsilon(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$
- Also let $z^m = x^{j(m)} + \epsilon \mathbf{1}_{a_m}$

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \geq 0$.
- For each $1 \leq k < m$, let $z^k = x^{j(k)} + \epsilon(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$
- Also let $z^m = x^{j(m)} + \epsilon\mathbf{1}_{a_m}$
- Also, for each $j \in I$, let $\lambda'_j = \lambda_j - \theta n(j)$.

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \cancel{\geq} > 0$.
- For each $1 \leq k < m$, let $z^k = x^{j(k)} + \epsilon(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$
- Also let $z^m = x^{j(m)} + \epsilon \mathbf{1}_{a_m}$
- Also, for each $j \in I$, let $\lambda'_j = \lambda_j - \theta n(j)$.
- we then generate y' as

$$y' = \sum_{j \in I} \lambda'_j x^j + \sum_{k: 1 \leq k \leq m} \theta z^k \tag{49}$$

$$= \sum_{j \in I} \lambda_j x^j + \theta \epsilon \mathbf{1}_{a_1} \tag{50}$$

SFM

- Now, since $\hat{c} >= 0$ and $x(a_1) > y(a_1)$ by graph construction, we have $\epsilon \geq 0$.
- For each $1 \leq k < m$, let $z^k = x^{j(k)} + \epsilon(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$
- Also let $z^m = x^{j(m)} + \epsilon \mathbf{1}_{a_m}$
- Also, for each $j \in I$, let $\lambda'_j = \lambda_j - \theta n(j)$.
- we then generate y' as

$$y' = \sum_{j \in I} \lambda'_j x^j + \sum_{k:1 \leq k \leq m} \theta z^k \quad (49)$$

$$= \sum_{j \in I} \lambda_j x^j + \theta \epsilon \mathbf{1}_{a_1} \quad (50)$$

- thus, we have $y' \in P_f$ with $y' \leq x$, and $y'(E) > y(E)$.

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \leq k < m$

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \leq k < m$
- And let $w^m = x^{j(k)} + \hat{c}(x^{j(m)}; a_m, a_{m+1})\mathbf{1}_{a_m}$.

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \leq k < m$
- And let $w^m = x^{j(k)} + \hat{c}(x^{j(m)}; a_m, a_{m+1})\mathbf{1}_{a_m}$.
- Then for $1 \leq k \leq m$, we have that w^k is also an extreme point (by previous theorem).

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \leq k < m$
- And let $w^m = x^{j(k)} + \hat{c}(x^{j(m)}; a_m, a_{m+1})\mathbf{1}_{a_m}$.
- Then for $1 \leq k \leq m$, we have that w^k is also an extreme point (by previous theorem).
- And so $z^k = (\epsilon/\delta_k)w^k + (1 - \epsilon/\delta_k)x^{j(k)}$.

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \le k < m$
- And let $w^m = x^{j(k)} + \hat{c}(x^{j(m)}; a_m, a_{m+1})\mathbf{1}_{a_m}$.
- Then for $1 \le k \le m$, we have that w^k is also an extreme point (by previous theorem).
- And so $z^k = (\epsilon/\delta_k)w^k + (1 - \epsilon/\delta_k)x^{j(k)}$.
- Therefore, y' is a convex combination of at most $|I| + m$ extreme points as given in Equation 49.

SFM

- Also, we have an expression for y' as a convex combination of extreme points of P_f.
- Let $\delta_k = \hat{c}(x^{j(k)}; a_k, a_{k+1})$ and let $w^k = x^{j(k)} + \delta_k(\mathbf{1}_{a_k} - \mathbf{1}_{a_{k+1}})$ for $1 \leq k < m$

  $\epsilon($ $)$
- And let $w^m = x^{j(k)} + \hat{c}(x^{j(m)}; a_m, a_{m+1})\mathbf{1}_{a_m}$.
- Then for $1 \leq k \leq m$, we have that w^k is also an extreme point (by previous theorem).
- And so $z^k = (\epsilon/\delta_k)w^k + (1 - \epsilon/\delta_k)x^{j(k)}$.
- Therefore, y' is a convex combination of at most $|I| + m$ extreme points as given in Equation 49.
- This can then be reduced in $O(n^3)$ time to at most $n+1$ extreme points using Carathéodory's theorem and associated linear algebra routines.

SFM

- Next, suppose that no s, t path exists.

SFM

- Next, suppose that no s, t path exists.
- Then there is a set $A \subseteq E$ such that no edge (a, b) of the graph has $a \in A + s$ and $b \notin A + s$ (i.e., a cut).

SFM

- Next, suppose that no s, t path exists.
- Then there is a set $A \subseteq E$ such that no edge (a, b) of the graph has $a \in A + s$ and $b \notin A + s$ (i.e., a cut).
- In such case, we (by construction) have that $y(k) = x(k)$ for $k \in E \setminus A$.

SFM

- Next, suppose that no s, t path exists.
- Then there is a set $A \subseteq E$ such that no edge (a, b) of the graph has $a \in A + s$ and $b \notin A + s$ (i.e., a cut).
- In such case, we (by construction) have that $y(k) = x(k)$ for $k \in E \setminus A$.
- Also, if it were the case A is x^j tight for $j \in I$, then we'd have

$$y(A) = \sum_{j \in I} \lambda_j x^j(A) \tag{51}$$

$x^j(A) = f(A)$

$$= \sum_{j \in I} \lambda_j f(A) \tag{52}$$

$$= f(A) \tag{53}$$

SFM

- Next, suppose that no s, t path exists.
- Then there is a set $A \subseteq E$ such that no edge (a, b) of the graph has $a \in A + s$ and $b \notin A + s$ (i.e., a cut).
- In such case, we (by construction) have that $y(k) = x(k)$ for $k \in E \setminus A$.
- Also, if it were the case A is x^j tight for $j \in I$, then we'd have

$$y(A) = \sum_{j \in I} \lambda_j x^j(A) \tag{51}$$

$$= \sum_{j \in I} \lambda_j f(A) \tag{52}$$

$$= f(A) \tag{53}$$

- Next we show that such an A is tight.

SFM

- Next we show that such an A is tight.

SFM

- Next we show that such an A is tight.
- If $a \in A$ and $j \in I$, then $a \in \text{sat}(x^j)$ since (a, t) is not an edge.

SFM

- Next we show that such an A is tight.
- If $a \in A$ and $j \in I$, then $a \in \text{sat}(x^j)$ since (a, t) is not an edge.
- Moreover, $\{b : b \preceq_j a\} \subseteq A$ since otherwise there is a $e \in A$ and $d \in E \setminus A$ such that e covers d in $\preceq_j$.

order defined by x^j

SFM

- Next we show that such an A is tight.
- If $a \in A$ and $j \in I$, then $a \in \text{sat}(x^j)$ since (a, t) is not an edge.
- Moreover, $\{b : b \preceq_j a\} \subseteq A$ since otherwise there is a $e \in A$ and $d \in E \setminus A$ such that e covers d in $\preceq_j$.
- But $\{b : b \preceq_j a\}$ is x^j-tight.

SFM

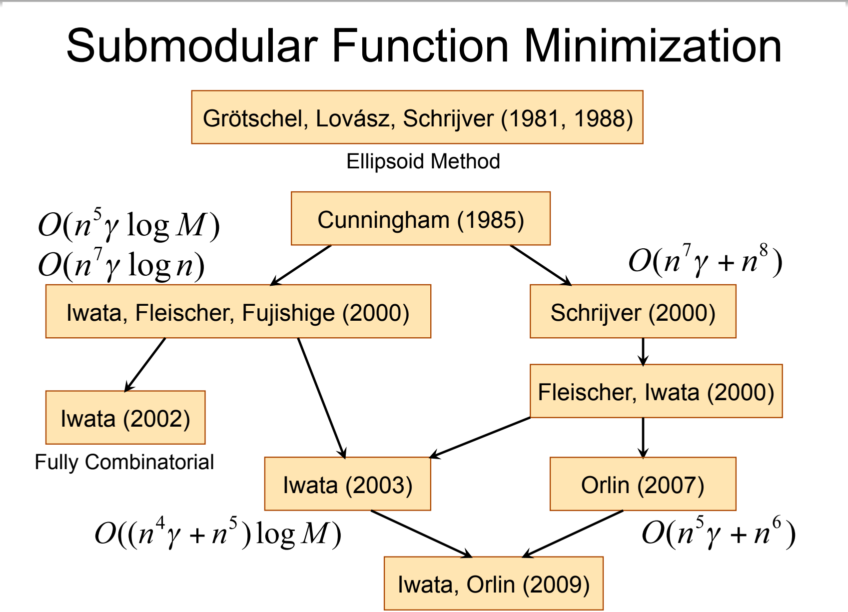
- Next we show that such an A is tight.
- If $a \in A$ and $j \in I$, then $a \in \text{sat}(x^j)$ since (a, t) is not an edge.
- Moreover, $\{b : b \preceq_j a\} \subseteq A$ since otherwise there is a $e \in A$ and $d \in E \setminus A$ such that e covers d in $\preceq_j$.
- But $\{b : b \preceq_j a\}$ is x^j-tight.
- Therefore, A is the union of x^j-tight sets and so itself is x^j-tight, as required.

SFM

Theorem 3.5

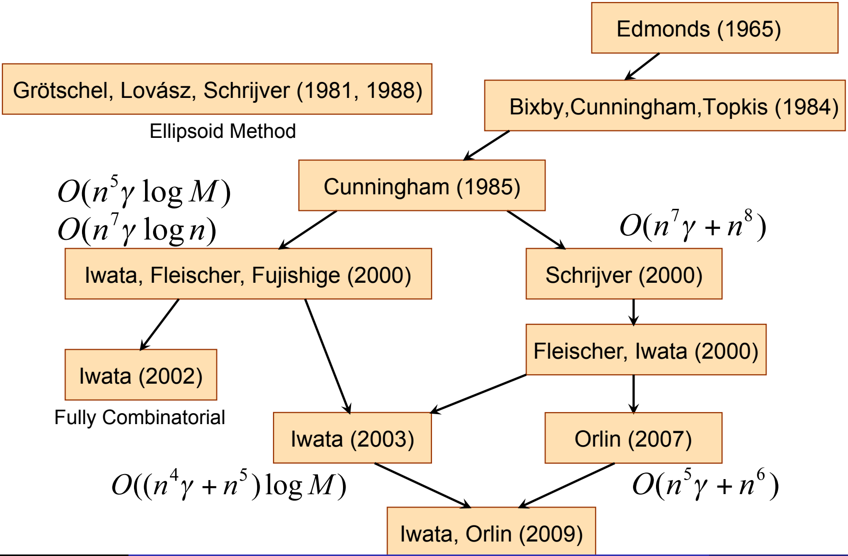
If f is integer valued, then so is x, and in such case it is possible to choose the steps δ in such a way that the running time of the algorithm is $O(Mn^6 \log(Mn)\beta)$ where M is an integer bound on the max value of f, $n = |E|$, and β is the cost of evaluating a submodular function query.

SFM Summary (from S. Iwata's slides on SFM)



SFM Summary (from S. Iwata's slides on SFM)

Submodular Function Minimization



Announcements

- Last lecture, and final presentations, will take place Thursday, June 9th, from 3-7:30pm in room EEB-303.
- The lecture will be from 3:00-5:00pm,
- The final presentations will be from 5:00-7:30pm. Feel free to bring dinner.
- Final project reports due next Wednesday, June 8th, at 11:45pm (on the web page, dropbox will be posted shortly).
- Final slides due by 1:00pm on Thursday, also via the dropbox.
- Final talks: Plan for 10 minutes of talking about your project (perhaps 8 minutes of lecture and 2 minutes of questions).

Scratch Paper

Scratch Paper

Scratch Paper

Sources for Today's Lecture

- Cunningham, "On Submodular Function Minimization", 1985.
- 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.