Submodular Functions, Optimization, and Applications to Machine Learning

— Spring Quarter, Lecture 12 —

http://j.ee.washington.edu/~bilmes/classes/ee596b_spring_2014/

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

May 12th, 2014

\[
f(A) + f(B) \geq f(A \cup B) + f(A \cap B)
\]

f(A) + 2f(C) + f(B) \geq f(A) + f(C) + f(B)

f(A \cap B) \leq f(A) + f(B) - f(A \cup B)

Clockwise from top left:
Read chapters 1 and 2, and sections 3.1-3.2 from Fujishige’s book.


Read Tom McCormick’s overview paper on SFM http://people.commerce.ubc.ca/faculty/mccormick/sfmchap8a.pdf

Read chapters 1 - 3 from Fujishige book.

Announcements, Assignments, and Reminders

- **Weekly Office Hours:** Wednesdays, 5:00-5:50, or by skype or google hangout (email me).
Class Road Map - IT-I

- L1 (3/31): Motivation, Applications, & Basic Definitions
- L2: (4/2): Applications, Basic Definitions, Properties
- L3: More examples and properties (e.g., closure properties), and examples, spanning trees
- L4: proofs of equivalent definitions, independence, start matroids
- L5: matroids, basic definitions and examples
- L6: More on matroids, System of Distinct Reps, Transversals, Transversal Matroid, Matroid and representation
- L7: Dual Matroids, other matroid properties, Combinatorial Geometries
- L8: Combinatorial Geometries, matroids and greedy, Polyhedra, Matroid Polytopes,
- L9: From Matroid Polytopes to Polymatroids.
- L10: Polymatroids and Submodularity
- L11: More properties of polymatroids, SFM special cases
- L12: polymatroid properties, extreme points polymatroids,
- L13: sat, dep, supp, exchange capacity, examples
- L14: Lattice theory: partially ordered sets; lattices; distributive, modular, submodular, and boolean lattices; ideals and join irreducibles.
- L15: Supp, Base polytope, polymatroids and entropic Venn diagrams, exchange capacity,
- L16: proof that minimum norm point yields min of submodular function, and the lattice of minimizers of a submodular function, Lovasz extension
- L17: Lovasz extension, Choquet Integration, more properties/examples of Lovasz extension, convex minimization and SFM.
- L18: Lovasz extension examples and structured convex norms, The Min-Norm Point Algorithm detailed.
- L19: symmetric submodular function minimization, maximizing monotone submodular function w. card constraints.
- L20: maximizing monotone submodular function w. other constraints, non-monotone maximization.

Finals Week: June 9th-13th, 2014.
A polymatroid function’s polyhedron is a polymatroid.

**Theorem 12.2.4**

Let $f$ be a polymatroid function defined on subsets of $E$. For any $x \in \mathbb{R}_E^+$, and any $P_f^+$-basis $y^x \in \mathbb{R}_E^+$ of $x$, the component sum of $y^x$ is

$$y^x(E) = \text{rank}(x) = \max \left( y(E) : y \leq x, y \in P_f^+ \right) = \min \left( x(A) + f(E \setminus A) : A \subseteq E \right)$$  \hfill (12.34)

As a consequence, $P_f^+$ is a polymatroid, since r.h.s. is constant w.r.t. $y^x$.

By taking $B = \text{supp}(x)$ (so elements $E \setminus B$ are zero in $x$), and for $b \in B$, $x(b)$ is big enough, the r.h.s. min has solution $A^* = E \setminus B$. We recover submodular function from the polymatroid polyhedron via the following:

$$f(B) = \max \left\{ y(B) : y \in P_f^+ \right\}$$  \hfill (12.35)

In fact, we will ultimately see a number of important consequences of this theorem (other than just that $P_f^+$ is a polymatroid).
Tight sets $\mathcal{D}(y)$ are closed, and max tight set $\text{sat}(y)$

Recall the definition of the set of tight sets at $y \in P_f^+$:

$$\mathcal{D}(y) \triangleq \{ A : A \subseteq E, y(A) = f(A) \}$$  \hspace{1cm} (12.18)

**Theorem 12.2.1**

For any $y \in P_f^+$, with $f$ a polymatroid function, then $\mathcal{D}(y)$ is closed under union and intersection.

**Proof.**

We have already proven this as part of Theorem 9.4.5

Also recall the definition of $\text{sat}(y)$, the maximal set of tight elements relative to $y \in \mathbb{R}_E^+$.  

$$\text{sat}(y) \overset{\text{def}}{=} \bigcup \{ T : T \in \mathcal{D}(y) \}$$  \hspace{1cm} (12.19)
Bipartite Matching

- Given a matching $A \subseteq E$ (which might be empty), we can increase the matching if we can find an augmenting path $S$.
- The updated matching becomes $A' = A \setminus S \cup S \setminus A = A \oplus S$, where $\oplus$ is the symmetric difference operator.
- The algorithm becomes:

**Algorithm 8.1: Alternating Path Bipartite Matching**

1. Let $A$ be an arbitrary (including empty) matching in $G = (V, F, E)$;
2. while There exists an augmenting path $S$ in $G$ do
   3. $A \leftarrow A \oplus S$;

- This can easily be made to run in $O(m^2n)$, where $|V| = m$, $|F| = n$, $m \leq n$, but it can be made to run much faster as well (see Schrijver-2003).
Matroid Intersection

- Let $M_1 = (V, \mathcal{I}_1)$ and $M_2 = (V, \mathcal{I}_2)$ be two matroids. Consider their common independent sets $\mathcal{I}_1 \cap \mathcal{I}_2$.
- While $(V, \mathcal{I}_1 \cap \mathcal{I}_2)$ is typically not a matroid (Exercise: show graphical example.), we might be interested in finding the maximum size common independent set. That is, find $\max |X|$ such that both $X \in \mathcal{I}_1$ and $X \in \mathcal{I}_2$.

**Theorem 12.2.5**

Let $M_1$ and $M_2$ be given as above, with rank functions $r_1$ and $r_2$. Then the size of the maximum size set in $\mathcal{I}_1 \cap \mathcal{I}_2$ is given by

\[
(r_1 * r_2)(V) \triangleq \min_{X \subseteq V} \left( r_1(X) + r_2(V \setminus X) \right) \tag{12.7}
\]

This is an instance of the convolution of two submodular functions, $f_1$ and $f_2$ that, evaluated at $Y \subseteq V$, is written as:

\[
(f_1 * f_2)(Y) = \min_{X \subseteq Y} \left( f_1(X) + f_2(Y \setminus X) \right) \tag{12.8}
\]
Lemma 12.2.3

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in $M$.

Proof.

- Suppose, to the contrary, that there are two distinct circuits $C_1, C_2$ such that $C_1 \cup C_2 \subseteq I \cup \{e\}$.
- Then $e \in C_1 \cap C_2$, and by (C2), there is a circuit $C_3$ of $M$ s.t. $C_3 \subseteq (C_1 \cup C_2) \setminus\{e\} \subseteq I$.
- This contradicts the independence of $I$.

In general, let $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (commonly called the fundamental circuit in $M$ w.r.t. $I$ and $e$).
Matroid Intersection Algorithm Idea

- Hopefully (eventually) we'll find an odd length sequence $S = (v_1, v_2, \ldots, v_s)$ such that we will be independent in both $M_1$ and $M_2$ and thus be one greater in size than $I$.

- We will have $v_i \notin I$ for $i$ odd (will be shown in blue), and will have $v_i \in I$ for $i$ even (will be shown in green), while $v \in I \setminus S$ will be shown in red.

- We then replace $I$ with $I \ominus S$ (quite analogous to the bipartite matching case), and start again.
Identifying Augmenting Sequences

Theorem 12.2.6

Let $I_p$ and $I_{p+1}$ be intersections of $M_1$ and $M_2$ with $p$ and $p + 1$ elements respectively. Then there exists an augmenting sequence $S \subseteq I_p \ominus I_{p+1}$ w.r.t. $I_p$.

Theorem 12.2.7

An intersection is of maximum cardinality iff it admits no augmenting sequence.

Theorem 12.2.8

For any intersection $I$, there exists a maximum cardinality intersection $I^*$ such that $\text{span}_1(I) \subseteq \text{span}_1(I^*)$ and $\text{span}_2(I) \subseteq \text{span}_2(I^*)$.

All this can be made to run in poly time.
Suppose \( M_i = (E, \mathcal{I}_i) \) is a matroid and that we have \( k \) of them on the same ground set \( E \).
Matroid Partition Problem

- Suppose $M_i = (E, \mathcal{I}_i)$ is a matroid and that we have $k$ of them on the same ground set $E$.
- We wish to, if possible, partition $E$ into $k$ blocks, $I_i, i \in \{1, 2, \ldots, k\}$ where $I_i \in \mathcal{I}_i$. 

Suppose $M_i = (E, \mathcal{I}_i)$ is a matroid and that we have $k$ of them on the same ground set $E$.

We wish to, if possible, partition $E$ into $k$ blocks, $I_i, i \in \{1, 2, \ldots, k\}$ where $I_i \in \mathcal{I}_i$.

Moreover, we want partition to be lexicographically maximum, that is $|I_1|$ is maximum, $|I_2|$ is maximum given $|I_1|$, and so on.
**Theorem 12.3.1**

Let $M_i$ be a collection of $k$ matroids as described. Then, a set $S \subseteq E$ can be partitioned into $k$ subsets $I_i, i = 1 \ldots k$ where $I_i \in \mathcal{I}_i$ is independent in matroid $i$, if and only if, for all $A \subseteq S$

$$|A| \leq \sum_{i=1}^{k} r_i(A)$$

(12.1)

where $r_i$ is the rank function of $M_i$. 

**Prof. Jeff Bilmes**

EE596b/Spring 2014/Submodularity - Lecture 12 - May 12th, 2014
Theorem 12.3.1

Let $M_i$ be a collection of $k$ matroids as described. Then, a set $S \subseteq E$ can be partitioned into $k$ subsets $I_i, i = 1 \ldots k$ where $I_i \in \mathcal{I}_i$ is independent in matroid $i$, if and only if, for all $A \subseteq S$

$$|A| \leq \sum_{i=1}^{k} r_i(A)$$

(12.1)

where $r_i$ is the rank function of $M_i$.

Now, if all matroids are the same $M_i = M$ for all $i$, we get condition

$$|A| \leq kr(A) \ \forall A \subseteq E$$

(12.2)
Matroid Partition Problem

**Theorem 12.3.1**

Let \( M_i \) be a collection of \( k \) matroids as described. Then, a set \( S \subseteq E \) can be partitioned into \( k \) subsets \( I_i, i = 1 \ldots k \) where \( I_i \in \mathcal{I}_i \) is independent in matroid \( i \), if and only if, for all \( A \subseteq S \)

\[
|A| \leq \sum_{i=1}^{k} r_i(A) \quad (12.1)
\]

where \( r_i \) is the rank function of \( M_i \).

- Now, if all matroids are the same \( M_i = M \) for all \( i \), we get condition

\[
|A| \leq kr(A) \quad \forall A \subseteq E \quad (12.2)
\]

- But considering vector of all ones \( \mathbf{1} \in \mathbb{R}^E_+ \), this is the same as

\[
\frac{1}{k}|A| = \frac{1}{k}\mathbf{1}(A) \leq r(A) \quad \forall A \subseteq E \quad (12.3)
\]
Recall definition of matroid polytope

\[ P_r^+ = \{ y \in \mathbb{R}_+^E : y(A) \leq r(A) \text{ for all } A \subseteq E \} \]  

(12.4)
Matroid Partition Problem

Recall definition of matroid polytope

\[ P_r^+ = \{ y \in \mathbb{R}_+^E : y(A) \leq r(A) \text{ for all } A \subseteq E \} \]  \hspace{1cm} (12.4)

Then we see that this special case of the matroid partition problem is just testing if \( \frac{1}{k} \mathbf{1} \in P_r^+ \), a problem of testing the membership in matroid polyhedra.
Recall definition of matroid polytope

\[ P_r^+ = \{ y \in \mathbb{R}_+^E : y(A) \leq r(A) \text{ for all } A \subseteq E \} \]  

Then we see that this special case of the matroid partition problem is just testing if \( \frac{1}{k} \mathbf{1} \in P_r^+ \), a problem of testing the membership in matroid polyhedra.

This is therefore a special case of submodular function minimization.
Review

- The next two slides from respectively from Lecture 9 and Lecture 8.
A polymatroid is a compact set $P \subseteq \mathbb{R}^E_+$ satisfying

1. $0 \in P$
2. If $y \leq x \in P$ then $y \in P$ (called down monotone).
3. For every $x \in \mathbb{R}^E_+$, any maximal vector $y \in P$ with $y \leq x$ (i.e., any $P$-basis of $x$), has the same component sum $y(E)$.

- Vectors within $P$ (i.e., any $y \in P$) are called independent, and any vector outside of $P$ is called dependent.
- Since all $P$-bases of $x$ have the same component sum, if $\mathcal{B}_x$ is the set of $P$-bases of $x$, then $\text{rank}(x) = y(E)$ for any $y \in \mathcal{B}_x$. 
Theorem 12.4.6

Let $M = (V, \mathcal{I})$ be a matroid, with rank function $r$, then for any weight function $w \in \mathbb{R}_+^V$, there exists a chain of sets $U_1 \subset U_2 \subset \cdots \subset U_n \subseteq V$ such that

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i)$$

(12.19)

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i 1_{U_i}$$

(12.20)
Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.
Polymatroidal polyhedron and greedy

- Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.
- Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.
Polymatroidal polyhedron and greedy

- Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.
- Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.
- For a matroid, we saw that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \(w \in \mathbb{R}^E_+\), the greedy algorithm leads to a set \(I \in \mathcal{I}\) of maximum weight \(w(I)\).
Polymatroidal polyhedron and greedy

- Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.
- Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.
- For a matroid, we saw that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \(w \in \mathbb{R}^E_+\), the greedy algorithm leads to a set \(I \in \mathcal{I}\) of maximum weight \(w(I)\).
- Stated succinctly, considering \(\max \{w(I) : I \in \mathcal{I}\}\), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.
Polymatroidal polyhedron and greedy

- Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.

- Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.

- For a matroid, we saw that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \(w \in \mathbb{R}^E_+\), the greedy algorithm leads to a set \(I \in \mathcal{I}\) of maximum weight \(w(I)\).

- Stated succinctly, considering \(\max \{w(I) : I \in \mathcal{I}\}\), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.

- Can we also characterize a polymatroid in this way?
Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.

Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.

For a matroid, we saw that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \(w \in \mathbb{R}^E_+\), the greedy algorithm leads to a set \(I \in \mathcal{I}\) of maximum weight \(w(I)\).

Stated succinctly, considering \(\max\{w(I) : I \in \mathcal{I}\}\), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.

Can we also characterize a polymatroid in this way?

That is, if we consider \(\max\{wx : x \in P^+_f\}\), where \(P^+_f\) represents the “independent vectors”, is it the case that \(P^+_f\) is a polymatroid iff greedy works for this maximization?
Polymatroidal polyhedron and greedy

Let \((E, \mathcal{I})\) be a set system and \(w \in \mathbb{R}^E_+\) be a weight vector.

Recall greedy algorithm: Set \(A = \emptyset\), and repeatedly choose \(y \in E \setminus A\) such that \(A \cup \{y\} \in \mathcal{I}\) with \(w(y)\) as large as possible, stopping when no such \(y\) exists.

For a matroid, we saw that set system \((E, \mathcal{I})\) is a matroid iff for each weight function \(w \in \mathbb{R}^E_+\), the greedy algorithm leads to a set \(I \in \mathcal{I}\) of maximum weight \(w(I)\).

Stated succinctly, considering \(\max \{w(I) : I \in \mathcal{I}\}\), then \((E, \mathcal{I})\) is a matroid iff greedy works for this maximization.

Can we also characterize a polymatroid in this way?

That is, if we consider \(\max \\{wx : x \in P_f^+\}\), where \(P_f^+\) represents the “independent vectors”, is it the case that \(P_f^+\) is a polymatroid iff greedy works for this maximization?

Can we even relax things so that \(w \in \mathbb{R}^E\)?
Polymatroidal polyhedron and greedy

What is the greedy solution in this setting, when \( w \in \mathbb{R}^E \)?
Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when \( w \in \mathbb{R}^E \)?
- Sort elements of \( E \) w.r.t. \( w \) so that, w.l.o.g.
  \[
  E = (e_1, e_2, \ldots, e_m) \text{ with } w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m).
  \]
What is the greedy solution in this setting, when \( w \in \mathbb{R}^E \)?

Sort elements of \( E \) w.r.t. \( w \) so that, w.l.o.g.

\[
E = (e_1, e_2, \ldots, e_m) \text{ with } w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m).
\]

Let \( k + 1 \) be the first point (if any) at which we are non-positive, i.e., \( w(e_k) > 0 \) and \( 0 \geq w(e_{k+1}) \).

That is, we have

\[
w(e_1) \geq w(e_2) \geq \cdots \geq w(e_k) > 0 \geq w(e_{k+1}) \geq \cdots \geq w(e_m) \quad (12.5)
\]
Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g.
  \[ E = (e_1, e_2, \ldots, e_m) \text{ with } w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m). \]
- Let $k + 1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.
- Next define partial accumulated sets $E_i$, for $i = 0 \ldots m$, we have w.r.t. the above sorted order:

  \[ E_i \overset{\text{def}}{=} \{e_1, e_2, \ldots e_i\} \tag{12.6} \]

  (note $E_0 = \emptyset$, $f(E_0) = 0$, and $E$ and $E_i$ is always sorted w.r.t $w$).
What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?

Sort elements of $E$ w.r.t. $w$ so that, w.l.o.g.

$E = (e_1, e_2, \ldots, e_m)$ with $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$.

Let $k + 1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.

Next define partial accumulated sets $E_i$, for $i = 0 \ldots m$, we have w.r.t. the above sorted order:

$$E_i \overset{\text{def}}{=} \{e_1, e_2, \ldots e_i\}$$

(note $E_0 = \emptyset$, $f(E_0) = 0$, and $E$ and $E_i$ is always sorted w.r.t $w$).

The greedy solution is the vector $x \in \mathbb{R}^E_+$ with elements defined as:

$$x(e_1) \overset{\text{def}}{=} f(E_1) = f(e_1) = f(e_1|E_0) = f(e_1|\emptyset)$$

(12.7)

$$x(e_i) \overset{\text{def}}{=} f(E_i) - f(E_{i-1}) = f(e_i|E_{i-1}) \text{ for } i = 2 \ldots k$$

(12.8)

$$x(e_i) \overset{\text{def}}{=} 0 \text{ for } i = k + 1 \ldots m = |E|$$

(12.9)
Some Intuition: greedy and gain

Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$
Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i | E_{i-1}) \leq f(e_i | E')$ for any $E' \subseteq E_{i-1}$
- So $x(e_1) = f(e_1)$ and this corresponds to $w(e_1) \geq w(e_i)$ for all $i \neq 1$. 
Some Intuition: greedy and gain

- Note \( x(e_i) = f(e_i | E_{i-1}) \leq f(e_i | E') \) for any \( E' \subseteq E_{i-1} \)
- So \( x(e_1) = f(e_1) \) and this corresponds to \( w(e_1) \geq w(e_i) \) for all \( i \neq 1 \).
- Hence, for the largest value of \( w \) (namely \( w(e_1) \)), we use for \( x(e_1) \) the largest possible gain value of \( e_1 \) (namely \( f(e_1|\emptyset) \geq f(e_1|A) \) for any \( A \subseteq E \setminus \{e_1\} \)).
Some Intuition: greedy and gain

- Note \( x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E') \) for any \( E' \subseteq E_{i-1} \).

- So \( x(e_1) = f(e_1) \) and this corresponds to \( w(e_1) \geq w(e_i) \) for all \( i \neq 1 \).

- Hence, for the largest value of \( w \) (namely \( w(e_1) \)), we use for \( x(e_1) \) the largest possible gain value of \( e_1 \) (namely \( f(e_1|\emptyset) \geq f(e_1|A) \) for any \( A \subseteq E \setminus \{e_1\} \)).

- For the next largest value of \( w \) (namely \( w(e_2) \)), we use for \( x(e_2) \) the next largest gain value of \( e_2 \) (namely \( f(e_2|e_1) \)), while still ensuring (as we will soon see in Theorem 12.4.1) that the resulting \( x \in P_f \).
Some Intuition: greedy and gain

- Note \( x(e_i) = f(e_i | E_{i-1}) \leq f(e_i | E') \) for any \( E' \subseteq E_{i-1} \).
- So \( x(e_1) = f(e_1) \) and this corresponds to \( w(e_1) \geq w(e_i) \) for all \( i \neq 1 \).
- Hence, for the largest value of \( w \) (namely \( w(e_1) \)), we use for \( x(e_1) \) the largest possible gain value of \( e_1 \) (namely \( f(e_1 | \emptyset) \geq f(e_1 | A) \) for any \( A \subseteq E \setminus \{e_1\} \)).
- For the next largest value of \( w \) (namely \( w(e_2) \)), we use for \( x(e_2) \) the next largest gain value of \( e_2 \) (namely \( f(e_2 | e_1) \)), while still ensuring (as we will soon see in Theorem 12.4.1) that the resulting \( x \in P_f \).
- This process continues, using the next largest possible gain of \( e_i \) for \( x(e_i) \) while ensuring we do not leave the polytope, given the values we’ve already chosen for \( x(e_{i'}) \) for \( i' < i \).
Polymatroidal polyhedron and greedy

Theorem 12.4.1

The vector $x \in \mathbb{R}_+^E$ as previously defined using the greedy algorithm maximizes $wx$ over $P_f^+$, with $w \in \mathbb{R}_+^E$, if $f$ is submodular.

Proof.
Theorem 12.4.1

The vector $x \in \mathbb{R}_+^E$ as previously defined using the greedy algorithm maximizes $wx$ over $P_f^+$, with $w \in \mathbb{R}_+^E$, if $f$ is submodular.

Proof.

Consider the LP strong duality equation:

$$\max(wx : x \in P_f^+) = \min\left(\sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}_+^{2^E}, \sum_{A \subseteq E} y_A 1_A \geq w\right)$$

(12.10)
Theorem 12.4.1

The vector \( x \in \mathbb{R}^E_+ \) as previously defined using the greedy algorithm maximizes \( wx \) over \( P^+_f \), with \( w \in \mathbb{R}^E_+ \), if \( f \) is submodular.

Proof.

- Consider the LP strong duality equation:

\[
\max (wx : x \in P^+_f) = \min \left( \sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}^{2^E}_+, \sum_{A \subseteq E} y_A 1_A \geq w \right)
\]

(12.10)

- Define the following vector \( y \in \mathbb{R}^{2^E}_+ \) as

\[
y_{E_i} \leftarrow w(e_i) - w(e_{i+1}) \quad \text{for} \quad i = 1 \ldots (m - 1),
\]

(12.11)

\[
y_E \leftarrow w(e_m), \quad \text{and}
\]

(12.12)

\[
y_A \leftarrow 0 \quad \text{otherwise}
\]

(12.13)
Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \leq f(A), \forall A$).
Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P^+_f$ (that is $x(A) \leq f(A), \forall A$).
- Order $A = (a_1, a_2, \ldots, a_k)$ based on order $(e_1, e_2, \ldots, e_m)$.

<table>
<thead>
<tr>
<th></th>
<th>$a_1$</th>
<th>$a_2$</th>
<th>$a_3$</th>
<th>$a_4$</th>
<th>$a_5$</th>
<th>$\ldots$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_1$</td>
<td>$e_2$</td>
<td>$e_3$</td>
<td>$e_4$</td>
<td>$e_5$</td>
<td>$e_6$</td>
<td>$e_7$</td>
</tr>
</tbody>
</table>

...
Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy \( x \in P_f^+ \) (that is \( x(A) \leq f(A), \forall A \)).
- Order \( A = (a_1, a_2, \ldots, a_k) \) based on order \( (e_1, e_2, \ldots, e_m) \).

<table>
<thead>
<tr>
<th></th>
<th>( a_1 )</th>
<th>( a_2 )</th>
<th>( a_3 )</th>
<th>( a_4 )</th>
<th>( a_5 )</th>
<th>( \ldots )</th>
<th>( e_1 )</th>
<th>( e_2 )</th>
<th>( e_3 )</th>
<th>( e_4 )</th>
<th>( e_5 )</th>
<th>( e_6 )</th>
<th>( e_7 )</th>
<th>( e_8 )</th>
<th>( e_9 )</th>
<th>( e_{10} )</th>
<th>( e_{11} )</th>
<th>( \ldots )</th>
<th>( e_m )</th>
</tr>
</thead>
</table>
- Define \( e^{-1} : E \to \{1, \ldots, m\} \) so that \( e^{-1}(e_i) = i \).

This means that with \( A = \{a_1, a_2, \ldots, a_k\} \), and \( \forall j \leq k \):

\[
\{a_1, a_2, \ldots, a_j\} \subseteq \{e_1, e_2, \ldots, e_{e^{-1}(a_j)}\}
\]  
(12.14)

and

\[
\{a_1, a_2, \ldots, a_{j-1}\} \subseteq \{e_1, e_2, \ldots, e_{e^{-1}(a_j)-1}\}
\]  
(12.15)

Also recall matlab notation: \( a_{1:j} \equiv \{a_1, a_2, \ldots, a_j\} \).

E.g., with \( j = 4 \) we get \( e^{-1}(a_4) = 9 \), and

\[
\{a_1, a_2, a_3, a_4\} \subseteq \{e_1, e_2, \ldots, e_9\}
\]  
(12.16)
Proof.

- We first will see that greedy \( x \in P^+_f \) (that is \( x(A) \leq f(A), \forall A \)).

- Order \( A = (a_1, a_2, \ldots, a_k) \) based on order \( (e_1, e_2, \ldots, e_m) \).

- Define \( e^{-1} : E \to \{1, \ldots, m\} \) so that \( e^{-1}(e_i) = i \).

- Then, we have \( x \in P^+_f \) since for all \( A \):

\[
f(A) = \sum_{i=1}^{k} f(a_i|a_1:i-1) \geq \sum_{i=1}^{k} f(a_i|e_1:e^{-1}(a_i)-1) = \sum_{a \in A} f(a|e_1:e^{-1}(a)-1) = x(A)
\]
Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \leq f(A), \forall A$).
- Order $A = (a_1, a_2, \ldots, a_k)$ based on order $(e_1, e_2, \ldots, e_m)$.

|       | $a_1$ | $a_2$ | $a_3$ | $a_4$ | $a_5$ | $\ldots$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ | $e_8$ | $e_9$ | $e_{10}$ | $e_{11}$ | $\ldots$ | $e_m$ |
|-------|-------|-------|-------|-------|-------|----------|-------|-------|-------|-------|-------|-------|-------|-------|-----------|-----------|----------|-------|
- Define $e^{-1}: E \rightarrow \{1, \ldots, m\}$ so that $e^{-1}(e_i) = i$.
- Then, we have $x \in P_f^+$ since for all $A$:

\[
f(A) = \sum_{i=1}^{k} f(a_i | a_{1:i-1}) \leq \sum_{i=1}^{k} f(a_i | e_{1:e^{-1}(a_i)-1}) = \sum_{a \in A} f(a | e_{1:e^{-1}(a)-1}) = x(A)
\]
Next, \( y \) is also feasible for the dual constraints in Eq. 12.10 since:

\[
\sum_{A : e_i \in A} y_A = \sum_{j \geq i} y_{E_j} = m - 1 \sum_{i = 1}^f(E_i) (w(e_i) - w(e_i + 1)) + f(E) w(e_m) = \sum_{A \subseteq E} y_A f(A)
\]
Polymatroidal polyhedron and greedy

**Proof.**

- Next, $y$ is also feasible for the dual constraints in Eq. 12.10 since:
- Next, we check that $y$ is dual feasible. Clearly, $y \geq 0$, 

...
Proof.

- Next, $y$ is also feasible for the dual constraints in Eq. 12.10 since:
  - Next, we check that $y$ is dual feasible. Clearly, $y \geq 0$,
  - and also, considering $y$ component wise, for any $i$, we have that

$$\sum_{A: e_i \in A} y_A = \sum_{j \geq i} y_{E_j} = m - 1 \sum_{j=i}^{m-1} (w(e_j) - w(e_{j+1})) + w(e_m) = w(e_i).$$
Polymatroidal polyhedron and greedy

**Proof.**

- Next, \( y \) is also feasible for the dual constraints in Eq. 12.10 since:
- Next, we check that \( y \) is dual feasible. Clearly, \( y \geq 0 \),
- and also, considering \( y \) component wise, for any \( i \), we have that

\[
\sum_{A: e_i \in A} y_A = \sum_{j \geq i} y_{E_j} = \sum_{j = i}^{m-1} (w(e_j) - w(e_{j+1})) + w(e_m) = w(e_i).
\]

- Now optimality for \( x \) and \( y \) follows from strong duality, i.e.:

\[
wx = \sum_{e \in E} w(e)x(e) = \sum_{i=1}^{m} w(e_i)f(e_i | E_{i-1}) = \sum_{i=1}^{m} w(e_i)\left(f(E_i) - f(E_{i-1})\right)
\]

\[
= \sum_{i=1}^{m-1} f(E_i)\left(w(e_i) - w(e_{i+1})\right) + f(E)w(e_m) = \sum_{A \subseteq E} y_A f(A) \quad \ldots
\]
Polymatroidal polyhedron and greedy

Proof.

The equality in prev. Eq. follows via Abel summation:

\[ wx = \sum_{i=1}^{m} w_i x_i \]  \hspace{1cm} (12.17)

\[ = \sum_{i=1}^{m} w_i \left( f(E_i) - f(E_{i-1}) \right) \]  \hspace{1cm} (12.18)

\[ = \sum_{i=1}^{m} w_i f(E_i) - \sum_{i=1}^{m-1} w_{i+1} f(E_i) \]  \hspace{1cm} (12.19)

\[ = w_m f(E_m) + \sum_{i=1}^{m-1} \left( w_i - w_{i+1} \right) f(E_i) \]  \hspace{1cm} (12.20)
What about $w \in \mathbb{R}^E$

- When $w$ contains negative elements, we have $x(e_i) = 0$ for $i = k + 1, \ldots, m$, where $k$ is the last positive element of $w$ when it is sorted in decreasing order.
What about $w \in \mathbb{R}^E$?

- When $w$ contains negative elements, we have $x(e_i) = 0$ for $i = k + 1, \ldots, m$, where $k$ is the last positive element of $w$ when it is sorted in decreasing order.

- Exercise: show a modification of the previous proof that works for arbitrary $w \in \mathbb{R}^E$. 
Polymatroidal polyhedron and greedy

Theorem 12.4.1

Conversely, suppose \( P_f^+ \) is a polytope of form

\[
P_f^+ = \left\{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \right\},
\]

then the greedy solution to \( \max(wx : x \in P) \) is optimum only if \( f \) is submodular.

Proof.

- Order elements of \( E \) arbitrarily as \((e_1, e_2, \ldots, e_m)\) and define \( E_i = (e_1, e_2, \ldots, e_i) \). Also, choose \( A \) and \( B \) arbitrarily.
Conversely, suppose $P_f^+$ is a polytope of form

$$P_f^+ = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max(wx : x \in P)$$

is optimum only if $f$ is submodular.

Proof.

- Order elements of $E$ arbitrarily as $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$. Also, choose $A$ and $B$ arbitrarily.
- For $1 \leq p \leq q \leq m$, define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$
  and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p)$.
Polymatroidal polyhedron and greedy

Theorem 12.4.1

Conversely, suppose $P_f^+$ is a polytope of form

$$P_f^+ = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max (wx : x \in P)$$

is optimum only if $f$ is submodular.

Proof.

- Order elements of $E$ arbitrarily as $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$. Also, choose $A$ and $B$ arbitrarily.

- For $1 \leq p \leq q \leq m$, define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$
  and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p)$

- Note, then we have $A \cap B = \{e_1, \ldots, e_k\} = E_k$, and $A \cup B = E_q$. 
Theorem 12.4.1

Conversely, suppose $P_f^+$ is a polytope of form

$P_f^+ = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E \}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if $f$ is submodular.

Proof.

1. Order elements of $E$ arbitrarily as $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$. Also, choose $A$ and $B$ arbitrarily.
2. For $1 \leq p \leq q \leq m$, define $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p)$.
3. Note, then we have $A \cap B = \{e_1, \ldots, e_k\} = E_k$, and $A \cup B = E_q$.
4. Define $w \in \{0, 1\}^m$ as:

$$w \overset{\text{def}}{=} \sum_{i=1}^{q} 1_{e_i} = 1_{A \cup B} \quad (12.21)$$
Polymatroidal polyhedron and greedy

**Theorem 12.4.1**

Conversely, suppose $P_f^+$ is a polytope of form

$$P_f^+ = \{ x \in \mathbb{R}^E_+ : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to

$$\max(wx : x \in P)$$

is optimum only if $f$ is submodular.

**Proof.**

- Order elements of $E$ arbitrarily as $(e_1, e_2, \ldots, e_m)$ and define $E_i = (e_1, e_2, \ldots, e_i)$. Also, choose $A$ and $B$ arbitrarily.
- For $1 \leq p \leq q \leq m$, define $A = \{ e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p \} = E_p$ and $B = \{ e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q \} = E_k \cup (E_q \setminus E_p)$.
- Note, then we have $A \cap B = \{ e_1, \ldots, e_k \} = E_k$, and $A \cup B = E_q$.
- Define $w \in \{0, 1\}^m$ as:

$$w \overset{\text{def}}{=} \sum_{i=1}^{q} 1_{e_i} = 1_{A \cup B} \quad (12.21)$$

- Suppose optimum solution $x$ is given by the greedy procedure.
Polymatroidal polyhedron and greedy

Proof.

Then

\[ \sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k} (f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B) \]

(12.22)
Polymatroidal polyhedron and greedy

Proof.

Then

$$\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k} (f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B)$$

(12.22)

and

$$\sum_{i=1}^{p} x_i = f(E_1) + \sum_{i=2}^{p} (f(E_i) - f(E_{i-1})) = f(E_p) = f(A)$$

(12.23)
Then

\[
\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k} (f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B)
\]  
(12.22)

and

\[
\sum_{i=1}^{p} x_i = f(E_1) + \sum_{i=2}^{p} (f(E_i) - f(E_{i-1})) = f(E_p) = f(A)
\]  
(12.23)

and

\[
\sum_{i=1}^{q} x_i = f(E_1) + \sum_{i=2}^{q} (f(E_i) - f(E_{i-1})) = f(E_q) = f(A \cup B)
\]  
(12.24)
Polymatroidal polyhedron and greedy

Proof.

Thus, we have

\[ x(B) = \sum_{i \in 1, \ldots, k, p+1, \ldots, q} x_i = \sum_{i: e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \]

(12.25)
Polymatroidal polyhedron and greedy

Proof.

Thus, we have

\[ x(B) = \sum_{i=1,...,k,p+1,...,q} x_i = \sum_{i:e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \] (12.25)

But given that the greedy algorithm gives the optimal solution to \( \max(wx : x \in P_f^+) \), we have that \( x \in P_f^+ \) and thus \( x(B) \leq f(B) \).
Polymatroidal polyhedron and greedy

Proof.

• Thus, we have

\[
x(B) = \sum_{i \in 1, \ldots, k, p+1, \ldots, q} x_i = \sum_{i: e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A)
\]

(12.25)

• But given that the greedy algorithm gives the optimal solution to \(\max(wx : x \in P_f^+)\), we have that \(x \in P_f^+\) and thus \(x(B) \leq f(B)\).

• Thus,

\[
x(B) = f(A \cup B) + f(A \cap B) - f(A) = \sum_{i: e_i \in B} x_i \leq f(B)
\]

(12.26)

ensuring the submodularity of \(f\), since \(A\) and \(B\) are arbitrary.
Thus, restating the above results into a single complete theorem, we have a result very similar to what we saw for matroids (i.e., Theorem 8.6.1)

**Theorem 12.4.1**

If $f : 2^E \rightarrow \mathbb{R}_+$ is given, and $P$ is a polytope in $\mathbb{R}_+^E$ of the form

$$P = \{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E \},$$

then the greedy solution to the problem $\max(wx : x \in P)$ is $\forall w$ optimum iff $f$ is monotone non-decreasing submodular (i.e., iff $P$ is a polymatroid).
Multiple Polytopes associated with arbitrary $f$

- Given an arbitrary submodular function $f : 2^V \rightarrow \mathbb{R}$ (not necessarily a polymatroid function, so it need not be positive, monotone, etc.).
Given an arbitrary submodular function \( f : 2^V \rightarrow \mathbb{R} \) (not necessarily a polymatroid function, so it need not be positive, monotone, etc.). If \( f(\emptyset) \neq 0 \), we can set \( f'(A) = f(A) - f(\emptyset) \) without destroying submodularity. This also does not change any minima, so we assume all functions are normalized \( f(\emptyset) = 0 \).

Note that due to constraint \( x(\emptyset) \leq f(\emptyset) \), we must have \( f(\emptyset) \geq 0 \) since if not (i.e., if \( f(\emptyset) < 0 \)), then \( P_f^{+} \) doesn’t exist.

Another form of normalization can do is:

\[
f'(A) = \begin{cases} 
  f(A) & \text{if } A \neq \emptyset \\
  0 & \text{if } A = \emptyset 
\end{cases} 
\] (12.27)

This preserves submodularity due to \( f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \), and if \( A \cap B = \emptyset \) then r.h.s. only gets smaller when \( f(\emptyset) \geq 0 \).
Multiple Polytopes associated with arbitrary $f$

- Given an arbitrary submodular function $f : 2^V \rightarrow \mathbb{R}$ (not necessarily a polymatroid function, so it need not be positive, monotone, etc.).
- If $f(\emptyset) \neq 0$, we can set $f'(A) = f(A) - f(\emptyset)$ without destroying submodularity. This also does not change any minima, so we assume all functions are normalized $f(\emptyset) = 0$.
- We can define several polytopes:

\[
P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \} \tag{12.27}
\]

\[
P_f^+ = P_f \cap \{ x \in \mathbb{R}^E : x \geq 0 \} \tag{12.28}
\]

\[
B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \} \tag{12.29}
\]
Multiple Polytopes associated with arbitrary $f$

- Given an arbitrary submodular function $f : 2^V \rightarrow \mathbb{R}$ (not necessarily a polymatroid function, so it need not be positive, monotone, etc.).
- If $f(\emptyset) \neq 0$, we can set $f'(A) = f(A) - f(\emptyset)$ without destroying submodularity. This also does not change any minima, so we assume all functions are normalized $f(\emptyset) = 0$.
- We can define several polytopes:

$$P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \}$$  \hspace{1cm} (12.27)

$$P_f^+ = P_f \cap \{ x \in \mathbb{R}^E : x \geq 0 \}$$  \hspace{1cm} (12.28)

$$B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \}$$  \hspace{1cm} (12.29)

- $P_f$ is what is sometimes called the extended polytope (sometimes notated as $EP_f$).
Multiple Polytopes associated with arbitrary $f$

- Given an arbitrary submodular function $f : 2^V \rightarrow \mathbb{R}$ (not necessarily a polymatroid function, so it need not be positive, monotone, etc.).
- If $f(\emptyset) \neq 0$, we can set $f'(A) = f(A) - f(\emptyset)$ without destroying submodularity. This also does not change any minima, so we assume all functions are normalized $f(\emptyset) = 0$.
- We can define several polytopes:

  $$P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \} \quad (12.27)$$

  $$P^+_f = P_f \cap \{ x \in \mathbb{R}^E : x \geq 0 \} \quad (12.28)$$

  $$B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \} \quad (12.29)$$

- $P_f$ is what is sometimes called the extended polytope (sometimes notated as $EP_f$).
- $P^+_f$ is $P_f$ restricted to the positive orthant.
Multiple Polytopes associated with arbitrary $f$

- Given an arbitrary submodular function $f : 2^V \rightarrow \mathbb{R}$ (not necessarily a polymatroid function, so it need not be positive, monotone, etc.).
- If $f(\emptyset) \neq 0$, we can set $f'(A) = f(A) - f(\emptyset)$ without destroying submodularity. This also does not change any minima, so we assume all functions are normalized $f(\emptyset) = 0$.
- We can define several polytopes:

$$P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \}$$  \hspace{1cm} (12.27)

$$P_f^+ = P_f \cap \{ x \in \mathbb{R}^E : x \geq 0 \}$$ \hspace{1cm} (12.28)

$$B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \}$$ \hspace{1cm} (12.29)

- $P_f$ is what is sometimes called the extended polytope (sometimes notated as $EP_f$).
- $P_f^+$ is $P_f$ restricted to the positive orthant.
- $B_f$ is called the base polytope.
Multiple Polytopes associated with $f$

\[ P_f^+ = P_f \cap \{ x \in \mathbb{R}^E : x \geq 0 \} \]  \hspace{1cm} (12.30)

\[ P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \} \]  \hspace{1cm} (12.31)

\[ B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \} \]  \hspace{1cm} (12.32)
Base Polytope in 3D

\[ P_f = \{ x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E \} \]  

(12.33)

\[ B_f = P_f \cap \{ x \in \mathbb{R}^E : x(E) = f(E) \} \]  

(12.34)
A polymatroid function’s polyhedron is a polymatroid.

**Theorem 12.5.1**

Let \( f \) be a submodular function defined on subsets of \( E \). For any \( x \in \mathbb{R}^E \), we have:

\[
\text{rank}(x) = \max \{ y(E) : y \leq x, y \in P_f \} = \min \{ x(A) + f(E \setminus A) : A \subseteq E \} \tag{12.35}
\]

Essentially the same theorem as Theorem 9.4.5. Taking \( x = 0 \) we get:

**Corollary 12.5.2**

Let \( f \) be a submodular function defined on subsets of \( E \). \( x \in \mathbb{R}^E \), we have:

\[
\text{rank}(0) = \max \{ y(E) : y \leq 0, y \in P_f \} = \min \{ f(A) : A \subseteq E \} \tag{12.36}
\]
Proof of Theorem 12.5.1

Let \( y^* \) be the optimal solution of the l.h.s. and let \( A \subseteq E \) be any subset.
Proof of Theorem 12.5.1.

Let $y^*$ be the optimal solution of the l.h.s. and let $A \subseteq E$ be any subset.

Then $y^*(E) = y^*(A) + y^*(E \setminus A) \leq f(A) + x(E \setminus A)$ since if $y^* \in P_f$, $y^*(A) \leq f(A)$ and since $y^* \leq x$, $y^*(E \setminus A) \leq x(E \setminus A)$. This is a form of weak duality.
Proof of Theorem 12.5.1

Let $y^*$ be the optimal solution of the l.h.s. and let $A \subseteq E$ be any subset.

Then $y^*(E) = y^*(A) + y^*(E \setminus A) \leq f(A) + x(E \setminus A)$ since if $y^* \in P_f$, $y^*(A) \leq f(A)$ and since $y^* \leq x$, $y^*(E \setminus A) \leq x(E \setminus A)$. This is a form of weak duality.

Also, for any $e \in E$, if $y^*(e) < x(e)$ then there must be some reason for this other than the constraint $y^* \leq x$, namely it must be that $\exists T \in D(x)$ with $e \in T$ (i.e., $e$ is a member of at least one of the tight sets).
Proof of Theorem 12.5.1.

Let $y^*$ be the optimal solution of the l.h.s. and let $A \subseteq E$ be any subset.

Then $y^*(E) = y^*(A) + y^*(E \setminus A) \leq f(A) + x(E \setminus A)$ since if $y^* \in P_f$, $y^*(A) \leq f(A)$ and since $y^* \leq x$, $y^*(E \setminus A) \leq x(E \setminus A)$. This is a form of weak duality.

Also, for any $e \in E$, if $y^*(e) < x(e)$ then there must be some reason for this other than the constraint $y^* \leq x$, namely it must be that $\exists T \in D(x)$ with $e \in T$ (i.e., $e$ is a member of at least one of the tight sets).

Hence, for all $e \notin \text{sat}(y^*)$ we have $y^*(e) = x(e)$, and moreover $y^*(\text{sat}(y^*)) = f(\text{sat}(y^*))$ by definition.
Proof of Theorem 12.5.1.

Let $y^*$ be the optimal solution of the l.h.s. and let $A \subseteq E$ be any subset.

Then $y^*(E) = y^*(A) + y^*(E \setminus A) \leq f(A) + x(E \setminus A)$ since if $y^* \in P_f$, $y^*(A) \leq f(A)$ and since $y^* \leq x$, $y^*(E \setminus A) \leq x(E \setminus A)$. This is a form of weak duality.

Also, for any $e \in E$, if $y^*(e) < x(e)$ then there must be some reason for this other than the constraint $y^* \leq x$, namely it must be that $\exists T \in D(x)$ with $e \in T$ (i.e., $e$ is a member of at least one of the tight sets).

Hence, for all $e \notin \text{sat}(y^*)$ we have $y^*(e) = x(e)$, and moreover $y^*(\text{sat}(y^*)) = f(\text{sat}(y^*))$ by definition.

Thus we have that $y^*(\text{sat}(y^*)) + y^*(E \setminus \text{sat}(y^*)) = f(\text{sat}(y^*)) + x(E \setminus \text{sat}(y^*))$, strong duality, showing that the two sides are equal for $y^*$. 
In Theorem 12.4.1, we can relax $P_f^+$ to $P_f$. 
In Theorem 12.4.1, we can relax $P_f^+$ to $P_f$.

If $\exists e$ such that $w(e) < 0$ then $\max(wx : x \in P_f) = \infty$ since we can let $x_e \to \infty$, unless we ignore the negative elements or assume $w \geq 0$. 
In Theorem 12.4.1, we can relax $P_f^+$ to $P_f$.

If $\exists e$ such that $w(e) < 0$ then $\max(wx : x \in P_f) = \infty$ since we can let $x_e \to \infty$, unless we ignore the negative elements or assume $w \geq 0$.

The proof, moreover, showed also that $x \in P_f$, not just $P_f^+$. 

Greedy and $P_f$

- In Theorem 12.4.1, we can relax $P_f^+$ to $P_f$.
- If $\exists e$ such that $w(e) < 0$ then $\max(wx : x \in P_f) = \infty$ since we can let $x_e \to \infty$, unless we ignore the negative elements or assume $w \geq 0$.
- The proof, moreover, showed also that $x \in P_f$, not just $P_f^+$.
- Moreover, in polymatroidal case, since the greedy constructed $x$ has $x(E) = f(E)$, we have that the greedy $x \in B_f$. 
In Theorem 12.4.1, we can relax $P^+_f$ to $P_f$.

If $\exists e$ such that $w(e) < 0$ then $\max(wx : x \in P_f) = \infty$ since we can let $x_e \to \infty$, unless we ignore the negative elements or assume $w \geq 0$.

The proof, moreover, showed also that $x \in P_f$, not just $P^+_f$.

Moreover, in polymatroidal case, since the greedy constructed $x$ has $x(E) = f(E)$, we have that the greedy $x \in B_f$.

In fact, we next will see that the greedy $x$ is a vertex of $B_f$. 
Polymatroid extreme points

- The greedy algorithm does more than solve $\max(wx : x \in P_f^+).$ We can use it to generate vertices of polymatroidal polytopes.
Polymatroid extreme points

- The greedy algorithm does more than solve $\max(wx : x \in P_f^+)$. We can use it to generate vertices of polymatroidal polytopes.
- Consider $P_f^+$ and also $C_f^+ \overset{\text{def}}{=} \{x : x \in \mathbb{R}_+^E, x(e) \leq f(e), \forall e \in E\}$
The greedy algorithm does more than solve $\max(wx : x \in P_f^+)$. We can use it to generate vertices of polymatroidal polytopes.

Consider $P_f^+$ and also $C_f^+ \overset{\text{def}}{=} \{x : x \in \mathbb{R}_E^+, x(e) \leq f(e), \forall e \in E\}$

Then ordering $A = (a_1, \ldots, a_{|A|})$ arbitrarily with $A_i = \{a_1, \ldots, a_i\}$, $f(A) = \sum_i f(a_i|A_{i-1}) \leq \sum_i f(a_i)$, and hence $P_f^+ \subseteq C_f^+$. 

**Polymatroid extreme points**
Polymatroid extreme points

- The greedy algorithm does more than solve $\max(wx : x \in P_f^+)$. We can use it to generate vertices of polymatroidal polytopes.
- Consider $P_f^+$ and also $C_f^+ \overset{\text{def}}{=} \{ x : x \in \mathbb{R}^E_+, x(e) \leq f(e), \forall e \in E \}$
- Then ordering $A = (a_1, \ldots, a_{|A|})$ arbitrarily with $A_i = \{a_1, \ldots, a_i\}$, $f(A) = \sum_i f(a_i|A_{i-1}) \leq \sum_i f(a_i)$, and hence $P_f^+ \subseteq C_f^+$. 
Polymatroid extreme points

Since \( w \in \mathbb{R}^E_{+} \) is arbitrary, it may be that any \( e \in E \) is max (i.e., is such that \( w(e) > w(e') \) for \( e' \in E \setminus \{e\} \)).
Polymatroid extreme points

- Since \( w \in \mathbb{R}_+^E \) is arbitrary, it may be that any \( e \in E \) is max (i.e., is such that \( w(e) > w(e') \) for \( e' \in E \setminus \{e\} \)).

- Thus, intuitively, any first vertex of the polytope away from the origin might be obtained by advancing along the corresponding axis.
Polymatroid extreme points

- Since $w \in \mathbb{R}^E_+$ is arbitrary, it may be that any $e \in E$ is max (i.e., is such that $w(e) > w(e')$ for $e' \in E \setminus \{e\}$).

- Thus, intuitively, any first vertex of the polytope away from the origin might be obtained by advancing along the corresponding axis.

- Recall, base polytope defined as the extreme face of $P_f$. I.e.,

$$B_f = P_f \cap \{x \in \mathbb{R}^E_+: x(E) = f(E)\} \quad (12.37)$$
Polymatroid extreme points

- Since $w \in \mathbb{R}^E_+$ is arbitrary, it may be that any $e \in E$ is max (i.e., is such that $w(e) > w(e')$ for $e' \in E \setminus \{e\}$).

- Thus, intuitively, any first vertex of the polytope away from the origin might be obtained by advancing along the corresponding axis.

- Recall, base polytope defined as the extreme face of $P_f$. I.e.,
  \begin{equation}
  B_f = P_f \cap \{x \in \mathbb{R}^E_+ : x(E) = f(E)\} \quad (12.37)
  \end{equation}

  Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we'll reach a vertex in $B_f$, and if we advance only in some dimensions, we'll reach a vertex in $P_f \setminus B_f$. We formalize this next:
Polymatroid extreme points

Since \( w \in \mathbb{R}_+^E \) is arbitrary, it may be that any \( e \in E \) is max (i.e., is such that \( w(e) > w(e') \) for \( e' \in E \setminus \{e\} \)).

Thus, intuitively, any first vertex of the polytope away from the origin might be obtained by advancing along the corresponding axis.

Recall, base polytope defined as the extreme face of \( P_f \). I.e.,

\[
B_f = P_f \cap \{x \in \mathbb{R}_+^E : x(E) = f(E)\} \quad (12.37)
\]

Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we’ll reach a vertex in \( B_f \), and if we advance only in some dimensions, we’ll reach a vertex in \( P_f \setminus B_f \).
Polymatroid extreme points

- Since $w \in \mathbb{R}_+^E$ is arbitrary, it may be that any $e \in E$ is max (i.e., is such that $w(e) > w(e')$ for $e' \in E \setminus \{e\}$).

- Thus, intuitively, any first vertex of the polytope away from the origin might be obtained by advancing along the corresponding axis.

- Recall, base polytope defined as the extreme face of $P_f$. I.e.,

$$B_f = P_f \cap \{x \in \mathbb{R}_+^E : x(E) = f(E)\} \quad (12.37)$$

- Also, intuitively, we can continue advancing along the skeletal edges of the polytope to reach any other vertex, given the appropriate ordering. If we advance in all dimensions, we’ll reach a vertex in $B_f$, and if we advance only in some dimensions, we’ll reach a vertex in $P_f \setminus B_f$.

- We formalize this next:
Polymatroid extreme points

- Given any arbitrary order of \( E = (e_1, e_2, \ldots, e_m) \), define \( E_i = (e_1, e_2, \ldots, e_i) \).
Polymatroid extreme points

- Given any arbitrary order of $E = (e_1, e_2, \ldots, e_m)$, define $E_i = (e_1, e_2, \ldots, e_i)$.
- As before, a vector $x$ is generated by $E_i$ using the greedy procedure as follows

\[
\begin{align*}
x(e_1) &= f(E_1) = f(e_1) \\
x(e_j) &= f(E_j) - f(E_{j-1}) = f(e_j | E_{j-1}) \text{ for } 2 \leq j \leq i \hspace{1cm} (12.38) \\
x(e) &= 0 \text{ for } e \in E \setminus E_i \hspace{1cm} (12.40)
\end{align*}
\]
Polymatroid extreme points

- Given any arbitrary order of $E = (e_1, e_2, \ldots, e_m)$, define $E_i = (e_1, e_2, \ldots, e_i)$.

- As before, a vector $x$ is generated by $E_i$ using the greedy procedure as follows

  $x(e_1) = f(E_1) = f(e_1)$ (12.38)

  $x(e_j) = f(E_j) - f(E_{j-1}) = f(e_j | E_{j-1})$ for $2 \leq j \leq i$ (12.39)

  $x(e) = 0$ for $e \in E \setminus E_i$ (12.40)

- An extreme point of $P_f$ is a point that is not a convex combination of two other distinct points in $P_f$. Equivalently, an extreme point corresponds to setting certain inequalities in the specification of $P_f$ to be equalities, so that there is a unique single point solution.
Polymatroid extreme points

**Theorem 12.6.1**

For a given ordering \( E = (e_1, \ldots, e_m) \) of \( E \) and a given \( E_i = (e_1, \ldots, e_i) \) and \( x \) generated by \( E_i \) using the greedy procedure \( (x(e_i) = f(e_i|E_{i-1})) \), then \( x \) is an extreme point of \( P_f \).
Theorem 12.6.1

For a given ordering $E = (e_1, \ldots, e_m)$ of $E$ and a given $E_i = (e_1, \ldots, e_i)$ and $x$ generated by $E_i$ using the greedy procedure ($x(e_i) = f(e_i|E_{i-1})$), then $x$ is an extreme point of $P_f$

Proof.

- We already saw that $x \in P_f$ (Theorem 12.4.1).
Polymatroid extreme points

**Theorem 12.6.1**

For a given ordering $E = (e_1, \ldots, e_m)$ of $E$ and a given $E_i = (e_1, \ldots, e_i)$ and $x$ generated by $E_i$ using the greedy procedure ($x(e_i) = f(e_i | E_{i-1})$), then $x$ is an extreme point of $P_f$

**Proof.**

- We already saw that $x \in P_f$ (Theorem 12.4.1).
- To show that $x$ is an extreme point of $P_f$, note that it is the unique solution of the following system of equations

\[
\begin{align*}
x(E_j) &= f(E_j) \text{ for } 1 \leq j \leq i \leq m \\
x(e) &= 0 \text{ for } e \in E \setminus E_i
\end{align*}
\] (12.41)

There are $i \leq m$ equations and $i \leq m$ unknowns, and simple Gaussian elimination gives us back the $x$ constructed via the Greedy algorithm!!
Polymatroid extreme points

As an example, we have $x(E_1) = x(e_1) = f(e_1)$
Polymatroid extreme points

- As an example, we have \( x(E_1) = x(e_1) = f(e_1) \)
- \( x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2) \) so
- \( x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1) = f(e_2 | e_1) \).
Polymatroid extreme points

- As an example, we have $x(E_1) = x(e_1) = f(e_1)$

- $x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so
  \[ x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1) = f(e_2 | e_1). \]

- $x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so
  \[ x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2) = f(e_3 | e_1, e_2). \]
Polymatroid extreme points

As an example, we have $x(E_1) = x(e_1) = f(e_1)$

$x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so

$x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1) = f(e_2 | e_1)$.

$x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so

$x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2) = f(e_3 | e_1, e_2)$

And so on . . . , but we see that this is just Gaussian elimination.
Polymatroid extreme points

- As an example, we have $x(E_1) = x(e_1) = f(e_1)$
- $x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so
  
  $x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1) = f(e_2|e_1)$.

- $x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so
  
  $x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2) = f(e_3|e_1, e_2)$

- And so on . . . , but we see that this is just Gaussian elimination.

- Also, since $x \in P_f$, for each $i$, we see that,

  $$x(E_j) = f(E_j) \text{ for } 1 \leq j \leq i$$  \hspace{1cm} (12.43)

  $$x(A) \leq f(A), \forall A \subseteq E$$  \hspace{1cm} (12.44)
Polymatroid extreme points

As an example, we have $x(E_1) = x(e_1) = f(e_1)$

$x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ so

$x(e_2) = f(e_1, e_2) - x(e_1) = f(e_1, e_2) - f(e_1) = f(e_2 | e_1)$.

$x(E_3) = x(e_1) + x(e_2) + x(e_3) = f(e_1, e_2, e_3)$ so

$x(e_3) = f(e_1, e_2, e_3) - x(e_2) - x(e_1) = f(e_1, e_2, e_3) - f(e_1, e_2) = f(e_3 | e_1, e_2)$

And so on . . . , but we see that this is just Gaussian elimination.

Also, since $x \in P_f$, for each $i$, we see that,

$$x(E_j) = f(E_j) \text{ for } 1 \leq j \leq i \quad (12.43)$$

$$x(A) \leq f(A), \forall A \subseteq E \quad (12.44)$$

Thus, the greedy procedure provides a modular function lower bound on $f$ that is tight on all points $E_i$ in the order. This can be useful in its own right.
Polymatroid extreme points
some examples
Polymatroid extreme points

Moreover, we have (and will ultimately prove)

**Corollary 12.6.2**

*If $x$ is an extreme point of $P_f$ and $B \subseteq E$ is given such that $\text{supp}(x) = \{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)) = \text{sat}(x)$, then $x$ is generated using greedy by some ordering of $B$.***
Moreover, we have (and will ultimately prove)

**Corollary 12.6.2**

If \( x \) is an extreme point of \( P_f \) and \( B \subseteq E \) is given such that
\[
\text{supp}(x) = \{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup \{ A : x(A) = f(A) \} = \text{sat}(x),
\]
then \( x \) is generated using greedy by some ordering of \( B \).

Note, \( \text{sat}(x) = \text{cl}(x) = \bigcup \{ A : x(A) = f(A) \} \) is also called the closure of \( x \) (recall that sets \( A \) such that \( x(A) = f(A) \) are called tight, and such sets are closed under union and intersection, as seen in Lecture 8, Theorem ??)
Polymatroid extreme points

Moreover, we have (and will ultimately prove)

**Corollary 12.6.2**

*If* $x$ *is an extreme point of* $P_f$ *and* $B \subseteq E$ *is given such that* $\text{supp}(x) = \{e \in E : x(e) \neq 0\} \subseteq B \subseteq \bigcup (A : x(A) = f(A)) = \text{sat}(x)$, *then* $x$ *is generated using greedy by some ordering of* $B$.

- Note, $\text{sat}(x) = \text{cl}(x) = \bigcup (A : x(A) = f(A))$ *is also called the closure of* $x$ *(recall that sets* $A$ *(such that* $x(A) = f(A)$ *are called tight, and such sets are closed under union and intersection, as seen in Lecture 8, Theorem ??)*

- Thus, $\text{cl}(x)$ *is a tight set.*
Moreover, we have (and will ultimately prove)

**Corollary 12.6.2**

*If* $x$ *is an extreme point of* $P_f$ *and* $B \subseteq E$ *is given such that*

$$\text{supp}(x) = \{e \in E : x(e) \neq 0\} \subseteq B \subseteq \bigcup(A : x(A) = f(A)) = \text{sat}(x),$$

*then* $x$ *is generated using greedy by some ordering of* $B$.

- Note, $\text{sat}(x) = \text{cl}(x) = \bigcup(A : x(A) = f(A))$ *is also called the closure of* $x$ *(recall that sets* $A$ *such that* $x(A) = f(A)$ *are called tight, and such sets are closed under union and intersection, as seen in Lecture 8, Theorem ??)*

- Thus, $\text{cl}(x)$ *is a tight set.*

- Also, $\text{supp}(x) = \{e \in E : x(e) \neq 0\}$ *is called the support of* $x$. 
Polymatroid extreme points

Moreover, we have (and will ultimately prove)

Corollary 12.6.2

If $x$ is an extreme point of $P_f$ and $B \subseteq E$ is given such that
\[ \text{supp}(x) = \{ e \in E : x(e) \neq 0 \} \subseteq B \subseteq \bigcup (A : x(A) = f(A)) = \text{sat}(x), \]
then $x$ is generated using greedy by some ordering of $B$.

Note, $\text{sat}(x) = \text{cl}(x) = \bigcup (A : x(A) = f(A))$ is also called the closure of $x$ (recall that sets $A$ such that $x(A) = f(A)$ are called tight, and such sets are closed under union and intersection, as seen in Lecture 8, Theorem ??)

Thus, $\text{cl}(x)$ is a tight set.

Also, $\text{supp}(x) = \{ e \in E : x(e) \neq 0 \}$ is called the support of $x$.

For arbitrary $x$, $\text{supp}(x)$ is not necessarily tight, but for an extreme point, $\text{supp}(x)$ is.
Recall $f(e|A) = f(A + e) - f(A)$

Notice how submodularity, $f(e|B) \leq f(e|A)$ for $A \subseteq B$, defines the shape of the polytope.

In fact, we have strictness here $f(e|B) < f(e|A)$ for $A \subset B$.

Also, consider how the greedy algorithm proceeds along the edges of the polytope.
Polymatroid with labeled edge lengths

- Recall \( f(e|A) = \)
  \( f(A + e) - f(A) \)

- Notice how
  submodularity,
  \( f(e|B) \leq f(e|A) \)
  for
  \( A \subseteq B \), defines the
  shape of the polytope.

- In fact, we have
  strictness here
  \( f(e|B) < f(e|A) \)
  for
  \( A \subset B \).

- Also, consider how the
  greedy algorithm
  proceeds along the
  edges of the polytope.
Intuition: why greedy works with polymatroids

- Given $w$, the goal is to find $x = (x(e_1), x(e_2))$ that maximizes $x^T w = x(e_1)w(e_1) + x(e_2)w(e_2)$.
- If $w(e_2) > w(e_1)$ the upper extreme point indicated maximizes $x^T w$ over $x \in P^+_f$.
- If $w(e_2) < w(e_1)$ the lower extreme point indicated maximizes $x^T w$ over $x \in P^+_f$. 

Maximal point in $P^+_f$ for $w$ in this region.
A polymatroid function’s polyhedron is a polymatroid.

**Theorem 12.7.4**

Let \( f \) be a polymatroid function defined on subsets of \( E \). For any \( x \in \mathbb{R}^E_+ \), and any \( P_f^+ \)-basis \( y^x \in \mathbb{R}^E_+ \) of \( x \), the component sum of \( y^x \) is

\[
y^x(E) = \text{rank}(x) = \max \left( y(E) : y \leq x, y \in P_f^+ \right)
\]

\[
= \min \left( x(A) + f(E \setminus A) : A \subseteq E \right) \quad (12.34)
\]

As a consequence, \( P_f^+ \) is a polymatroid, since r.h.s. is constant w.r.t. \( y^x \).

By taking \( B = \text{supp}(x) \) (so elements \( E \setminus B \) are zero in \( x \)), and for \( b \in B \), \( x(b) \) is big enough, the r.h.s. min has solution \( A^* = E \setminus B \). We recover submodular function from the polymatroid polyhedron via the following:

\[
f(B) = \max \left\{ y(B) : y \in P_f^+ \right\} \quad (12.35)
\]

In fact, we will ultimately see a number of important consequences of this theorem (other than just that \( P_f^+ \) is a polymatroid).
Considering Theorem 9.4.5, the matroid case is now a special case, where we have that:

**Corollary 12.7.2**

*We have that:*

\[
\max \{ y(E) : y \in P_{ind. set}(M), y \leq x \} = \min \{ r_M(A) + x(E \setminus A) : A \subseteq E \}
\]

(12.2)

where \( r_M \) is the matroid rank function of some matroid.
Most violated inequality problem in matroid polytope case

Consider

\[ P_r^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \} \]  \hspace{1cm} (12.45)
Consider

\[ P_r^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \} \] (12.45)

Suppose we have any \( x \in \mathbb{R}_+^E \) such that \( x \notin P_r^+ \).
Consider

\[ P_r^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \} \]  \hspace{1cm} (12.45)

Suppose we have any \( x \in \mathbb{R}^E_+ \) such that \( x \not\in P_r^+ \).

Hence, there must be a set of \( \mathcal{W} \subseteq 2^V \), each member of which corresponds to a violated inequality, i.e., equations of the form \( x(A) > r_M(A) \) for \( A \in \mathcal{W} \).
Most violated inequality problem in matroid polytope case

Consider

\[ P^+_r = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \} \] (12.45)

Suppose we have any \( x \in \mathbb{R}^E_+ \) such that \( x \not\in P^+_r \).

Hence, there must be a set of \( \mathcal{W} \subseteq 2^V \), each member of which corresponds to a violated inequality, i.e., equations of the form \( x(A) > r_M(A) \) for \( A \in \mathcal{W} \).

The most violated inequality when \( x \) is considered w.r.t. \( P^+_r \) corresponds to the set \( A \) that maximizes \( x(A) - r_M(A) \), i.e., the most violated inequality is valuated as:

\[ \max \{ x(A) - r_M(A) : A \in \mathcal{W} \} = \max \{ x(A) - r_M(A) : A \subseteq E \} \] (12.46)
Most violated inequality problem in matroid polytope case

Consider

$$P_r^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E \}$$  \hspace{1cm} (12.45)

Suppose we have any $x \in \mathbb{R}^E_+$ such that $x \not\in P_r^+$.

Hence, there must be a set of $\mathcal{W} \subseteq 2^V$, each member of which corresponds to a violated inequality, i.e., equations of the form $x(A) > r_M(A)$ for $A \in \mathcal{W}$.

The most violated inequality when $x$ is considered w.r.t. $P_r^+$ corresponds to the set $A$ that maximizes $x(A) - r_M(A)$, i.e., the most violated inequality is valuated as:

$$\max \{ x(A) - r_M(A) : A \in \mathcal{W} \} = \max \{ x(A) - r_M(A) : A \subseteq E \}$$  \hspace{1cm} (12.46)

Since $x$ is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in::

$$\min \{ r_M(A) + x(E \setminus A) : A \subseteq E \}$$  \hspace{1cm} (12.47)
Consider

\[ P_f^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E \} \]  \hspace{1cm} (12.48)
Consider

\[ P_f^+ = \{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E \} \]  

(12.48)

Suppose we have any \( x \in \mathbb{R}_+^E \) such that \( x \not\in P_f^+ \).
Most violated inequality/polymatroid membership/SFM

- Consider

\[ P_f^+ = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E \right\} \]  \hspace{1cm} (12.48)

- Suppose we have any \( x \in \mathbb{R}^E_+ \) such that \( x \not\in P_f^+ \).

- Hence, there must be a set of \( \mathcal{W} \subseteq 2^V \), each member of which corresponds to a violated inequality, i.e., equations of the form \( x(A) > r_M(A) \) for \( A \in \mathcal{W} \).

![Diagrams](http://example.com/diagrams.png)

Left: \( \mathcal{W} = \{\{1\}\} \)

Center: \( \mathcal{W} = \{\{2\}\} \)

Right: \( \mathcal{W} = \{\{1, 2\}\} \)
The most violated inequality when \( x \) is considered w.r.t. \( P_f^+ \) corresponds to the set \( A \) that maximizes \( x(A) - f(A) \), i.e., the most violated inequality is valuated as:

\[
\max \{ x(A) - f(A) : A \in \mathcal{W} \} = \max \{ x(A) - f(A) : A \subseteq E \} \quad (12.49)
\]
The most violated inequality when $x$ is considered w.r.t. $P_f^+$ corresponds to the set $A$ that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max \{ x(A) - f(A) : A \in \mathcal{W} \} = \max \{ x(A) - f(A) : A \subseteq E \} \quad (12.49)$$

Since $x$ is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min \{ f(A) + x(E \setminus A) : A \subseteq E \} \quad (12.50)$$
The most violated inequality when $x$ is considered w.r.t. $P_f^+$ corresponds to the set $A$ that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valued as:

$$\max \{x(A) - f(A) : A \in \mathcal{W}\} = \max \{x(A) - f(A) : A \subseteq E\} \quad (12.49)$$

Since $x$ is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in:

$$\min \{f(A) + x(E \setminus A) : A \subseteq E\} \quad (12.50)$$

More importantly, $\min \{f(A) + x(E \setminus A) : A \subseteq E\}$ is a form of submodular function minimization, namely $\min \{f(A) - x(A) : A \subseteq E\}$ for a submodular $f$ and $x \in \mathbb{R}_+^E$, consisting of a difference of polymatroid and modular function (so $f - x$ is no longer necessarily monotone, nor positive).
The most violated inequality when $x$ is considered w.r.t. $P_f^+$ corresponds to the set $A$ that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max \{ x(A) - f(A) : A \in \mathcal{W} \} = \max \{ x(A) - f(A) : A \subseteq E \} \quad (12.49)$$

Since $x$ is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min \{ f(A) + x(E \setminus A) : A \subseteq E \} \quad (12.50)$$

More importantly, $\min \{ f(A) + x(E \setminus A) : A \subseteq E \}$ is a form of submodular function minimization, namely

$$\min \{ f(A) - x(A) : A \subseteq E \}$$

for a submodular $f$ and $x \in \mathbb{R}^E_+$, consisting of a difference of polymatroid and modular function (so $f - x$ is no longer necessarily monotone, nor positive).

We will ultimately answer how general this form of SFM is.
Definition 12.8.1 (closed/flat/subspace)

A subset $A \subseteq E$ is closed (equivalently, a flat or a subspace) of matroid $M$ if for all $x \in E \setminus A$, $r(A \cup \{x\}) = r(A) + 1$.

A hyperplane is a flat of rank $r(M) - 1$.

Definition 12.8.2 (closure)

Given $A \subseteq E$, the closure (or span) of $A$, is defined by

$\text{span}(A) = \{b \in E : r(A \cup \{b\}) = r(A)\}$.

Therefore, a closed set $A$ has $\text{span}(A) = A$.

Definition 12.8.3 (circuit)

A subset $A \subseteq E$ is circuit or a cycle if it is an inclusionwise-minimal dependent set (i.e., if $r(A) < |A|$ and for any $a \in A$, $r(A \setminus \{a\}) = |A| - 1$).
Several circuit definitions for matroids.

**Theorem 12.8.1 (Matroid by circuits)**

Let $E$ be a set and $C$ be a collection of nonempty subsets of $E$, such that no two sets in $C$ are contained in each other. Then the following are equivalent.

1. $C$ is the collection of circuits of a matroid;
2. if $C, C' \in C$, and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $C$;
3. if $C, C' \in C$, and $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $C$ containing $y$;

Again, think about this for a moment in terms of linear spaces and matrices, and spanning trees.
Lemma 12.8.3

Let \( I \in \mathcal{I}(M) \), and \( e \in E \), then \( I \cup \{e\} \) contains at most one circuit in \( M \).

Proof.

- Suppose, to the contrary, that there are two distinct circuits \( C_1, C_2 \) such that \( C_1 \cup C_2 \subseteq I \cup \{e\} \).
- Then \( e \in C_1 \cap C_2 \), and by (C2), there is a circuit \( C_3 \) of \( M \) s.t. \( C_3 \subseteq (C_1 \cup C_2) \setminus \{e\} \subseteq I \)
- This contradicts the independence of \( I \).

In general, let \( C(I, e) \) be the unique circuit associated with \( I \cup \{e\} \) (commonly called the fundamental circuit in \( M \) w.r.t. \( I \) and \( e \)).
Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in $M$ w.r.t. $I$ and $e$, if it exists).
Matroids: The Fundamental Circuit

- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in $M$ w.r.t. $I$ and $e$, if it exists).
- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).
Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in $M$ w.r.t. $I$ and $e$, if it exists).

If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).

If $e \in I$, then $I + e = I$ doesn’t create a circuit. In such cases, $C(I, e)$ is not really defined.
Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in $M$ w.r.t. $I$ and $e$, if it exists).

- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).
- If $e \in I$, then $I + e = I$ doesn’t create a circuit. In such cases, $C(I, e)$ is not really defined.
- In such cases, we define $C(I, e) = \{e\}$, and we will soon see why we do this.
Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in $M$ w.r.t. $I$ and $e$, if it exists).

If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).

If $e \in I$, then $I + e = I$ doesn’t create a circuit. In such cases, $C(I, e)$ is not really defined.

In such cases, we define $C(I, e) = \{e\}$, and we will soon see why we do this.

If $e \notin \text{span}(I)$, then $C(I, e) = \emptyset$, since no circuit is created in this case.
Lemma 12.8.1

Let $\mathcal{B}(C)$ be the set of bases of $C$. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free set $C \subseteq E$, we have that:

$$
\bigcup_{B \in \mathcal{B}(C)} B = C.
$$

(12.51)
Lemma 12.8.1

Let $\mathcal{B}(C)$ be the set of bases of $C$. Then, given matroid $\mathcal{M} = (E, I)$, and any loop-free set $C \subseteq E$, we have that:

$$\bigcup_{B \in \mathcal{B}(C)} B = C. \quad (12.51)$$

Proof.

- Define $C' \triangleq \bigcup_{B \in \mathcal{B}(C)}$, and suppose $\exists c \in C$ such that $c \notin C'$. 
Lemma 12.8.1

Let $\mathcal{B}(C)$ be the set of bases of $C$. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free set $C \subseteq E$, we have that:

$$ \bigcup_{B \in \mathcal{B}(C)} B = C. \quad (12.51) $$

Proof.

- Define $C' \triangleq \bigcup_{B \in \mathcal{B}(C)}$, and suppose $\exists c \in C$ such that $c \notin C'$.
- Hence, $\forall B \in \mathcal{B}(C')$ we have $c \notin B$, and $B + c$ contains a single circuit for any $B$, namely $C(B, c)$. 

**Prof. Jeff Bilmes**
EE596b/Spring 2014/Submodularity - Lecture 12 - May 12th, 2014

---

**Prof. Jeff Bilmes**
EE596b/Spring 2014/Submodularity - Lecture 12 - May 12th, 2014
Union of matroid bases of a set

**Lemma 12.8.1**

Let $\mathcal{B}(C)$ be the set of bases of $C$. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free set $C \subseteq E$, we have that:

$$\bigcup_{B \in \mathcal{B}(C)} B = C.$$  \hfill(12.51)

**Proof.**

- Define $C' \triangleq \bigcup_{B \in \mathcal{B}(C)} B$, and suppose $\exists c \in C$ such that $c \notin C'$.
- Hence, $\forall B \in \mathcal{B}(C')$ we have $c \notin B$, and $B + c$ contains a single circuit for any $B$, namely $C(B, c)$.
- Then choose $c' \in C(B, c)$ with $c' \neq c$. 

---

Prof. Jeff Bilmes  
EE596b/Spring 2014/Submodularity - Lecture 12 - May 12th, 2014  
F43/47 (pg.143/167)
Lemma 12.8.1

Let $\mathcal{B}(C)$ be the set of bases of $C$. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free set $C \subseteq E$, we have that:

$$
\bigcup_{B \in \mathcal{B}(C)} B = C. \tag{12.51}
$$

Proof.

- Define $C' \triangleq \bigcup_{B \in \mathcal{B}(C)}$, and suppose $\exists c \in C$ such that $c \notin C'$.
- Hence, $\forall B \in \mathcal{B}(C')$ we have $c \notin B$, and $B + c$ contains a single circuit for any $B$, namely $C(B, c)$.
- Then choose $c' \in C(B, c)$ with $c' \neq c$.
- Then $B + c - c'$ is independent size $|B|$ subset of $C$ and hence spans $C$, and thus is a $c$-containing member of $\mathcal{B}(C)$, contradicting $c \notin C'$. 
The \texttt{sat} function $\equiv$ Polymatroid Closure

- Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
The \textit{sat} function $\equiv$ Polymatroid Closure

• Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
• We wish to generalize closure to polymatroids.
The \textit{sat} function $= \text{Polymatroid Closure}$

- Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function $f$. 

The \textit{sat} function $= \text{Polymatroid Closure}$

- Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function $f$.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.
The \textit{sat} function $= \text{Polymatroid Closure}$

- Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function $f$.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.
- That is, we saw in Lecture 7 that for any $A, B \in \mathcal{D}(x)$, we have that $A \cup B \in \mathcal{D}(x)$ and $A \cap B \in \mathcal{D}(x)$, which can constitute a join and meet.
The \textit{sat} function $\equiv$ Polymatroid Closure

- Thus, in a matroid, closure (span) of a set $A$ are all items that $A$ spans (eq. that depend on $A$).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function $f$.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.
- That is, we saw in Lecture 7 that for any $A, B \in \mathcal{D}(x)$, we have that $A \cup B \in \mathcal{D}(x)$ and $A \cap B \in \mathcal{D}(x)$, which can constitute a join and meet.
- Recall, for a given $x \in P_f$, we have defined this tight family as

\[
\mathcal{D}(x) = \{ A : A \subseteq E, x(A) = f(A) \} \quad (12.52)
\]
The `sat` function = Polymatroid Closure

- Now given \( x \in P_f^+ \):
  \[
  D(x) = \{ A : A \subseteq E, x(A) = f(A) \} \tag{12.53}
  \]
  \[
  = \{ A : f(A) - x(A) = 0 \} \tag{12.54}
  \]
The \textit{sat} function $=\text{Polymatroid Closure}$

- Now given $x \in P_f^+$:

\[ D(x) = \{ A : A \subseteq E, x(A) = f(A) \} = \{ A : f(A) - x(A) = 0 \} \quad (12.53) \]

- Since $x \in P_f^+$ and $f$ is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $D(x)$ are the zero-valued minimizers (if any) of $f'(A)$. 

The \textit{sat} function $\equiv$ Polymatroid Closure

Now given $x \in P_f^+$:

$$
\mathcal{D}(x) = \{ A : A \subseteq E, x(A) = f(A) \} \tag{12.53}
$$

$$
= \{ A : f(A) - x(A) = 0 \} \tag{12.54}
$$

Since $x \in P_f^+$ and $f$ is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $\mathcal{D}(x)$ are the zero-valued minimizers (if any) of $f'(A)$.

The zero-valued minimizers of $f'$ are thus closed under union and intersection.
Now given $x \in P_f^+$:

$$D(x) = \{A : A \subseteq E, x(A) = f(A)\}$$

(12.53)

$$= \{A : f(A) - x(A) = 0\}$$

(12.54)

Since $x \in P_f^+$ and $f$ is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $D(x)$ are the zero-valued minimizers (if any) of $f'(A)$.

The zero-valued minimizers of $f'$ are thus closed under union and intersection.

In fact, this is true for all minimizers of a submodular function as stated in the next theorem.
Minimizers of a Submodular Function form a lattice

**Theorem 12.9.1**

For arbitrary submodular \( f \), the minimizers are closed under union and intersection. That is, let \( M = \arg \min_{X \subseteq E} f(X) \) be the set of minimizers of \( f \). Let \( A, B \in M \). Then \( A \cup B \in M \) and \( A \cap B \in M \).
Minimizers of a Submodular Function form a lattice

Theorem 12.9.1

For arbitrary submodular $f$, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \arg\min_{X \subseteq E} f(X)$ be the set of minimizers of $f$. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.
Minimizers of a Submodular Function form a lattice

**Theorem 12.9.1**

For arbitrary submodular \( f \), the minimizers are closed under union and intersection. That is, let \( \mathcal{M} = \arg\min_{X \subseteq E} f(X) \) be the set of minimizers of \( f \). Let \( A, B \in \mathcal{M} \). Then \( A \cup B \in \mathcal{M} \) and \( A \cap B \in \mathcal{M} \).

**Proof.**

Since \( A \) and \( B \) are minimizers, we have \( f(A) = f(B) \leq f(A \cap B) \) and \( f(A) = f(B) \leq f(A \cup B) \).
Minimizers of a Submodular Function form a lattice

Theorem 12.9.1

For arbitrary submodular $f$, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \arg\min_{X \subseteq E} f(X)$ be the set of minimizers of $f$. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

Since $A$ and $B$ are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.

By submodularity, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \quad (12.55)$$
Minimizers of a Submodular Function form a lattice

**Theorem 12.9.1**

For arbitrary submodular \( f \), the minimizers are closed under union and intersection. That is, let \( \mathcal{M} = \arg\min_{X \subseteq E} f(X) \) be the set of minimizers of \( f \). Let \( A, B \in \mathcal{M} \). Then \( A \cup B \in \mathcal{M} \) and \( A \cap B \in \mathcal{M} \).

**Proof.**

Since \( A \) and \( B \) are minimizers, we have \( f(A) = f(B) \leq f(A \cap B) \) and \( f(A) = f(B) \leq f(A \cup B) \).

By submodularity, we have

\[
 f(A) + f(B) \geq f(A \cup B) + f(A \cap B)
\]

(12.55)

Hence, we must have \( f(A) = f(B) = f(A \cup B) = f(A \cap B) \).
Minimizers of a Submodular Function form a lattice

**Theorem 12.9.1**

For arbitrary submodular $f$, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \text{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of $f$. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

**Proof.**

Since $A$ and $B$ are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.

By submodularity, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \quad (12.55)$$

Hence, we must have $f(A) = f(B) = f(A \cup B) = f(A \cap B)$.

Thus, the minimizers of a submodular function form a lattice, and there is a maximal and a minimal minimizer of every submodular function.
The \textbf{sat} function \(\equiv\) Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in \(\mathcal{D}(x)\), also called the polymatroid closure or \textbf{sat} (saturation function).
The \textit{sat} function $\equiv$ Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $D(x)$, also called the polymatroid closure or \textit{sat} (saturation function).
- For some $x \in P_f$, we have defined:

$$\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{ A : A \in D(x) \}$$

(12.56)
The \textit{sat} function $\equiv$ Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or \textit{sat} (\textit{saturation} function).
- For some $x \in P_f$, we have defined:

$$\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \quad (12.56)$$

$$= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \quad (12.57)$$

Hence, \text{sat}(x) is the maximal (zero-valued) minimizer of the submodular function $f(x) \triangleq f(A) - x(A)$. Eq. (12.58) says that \text{sat} consists of any point $x$ that is $P_f$ saturated (any additional positive movement, in that dimension, leaves $P_f$). We'll revisit this in a few slides. First, we see how \text{sat} generalizes matroid closure.
The \textit{sat} function $\equiv$ Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or \textit{sat} (\textit{saturation} function).

- For some $x \in P_f$, we have defined:

  $$\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{ A : A \in \mathcal{D}(x) \} \quad (12.56)$$

  $$= \bigcup \{ A : A \subseteq E, x(A) = f(A) \} \quad (12.57)$$

  $$= \{ e : e \in E, \forall \alpha > 0, x + \alpha 1_e \notin P_f \} \quad (12.58)$$
The \textit{sat} function $\equiv$ Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $D(x)$, also called the polymatroid closure or \textit{sat} (saturation function).

- For some $x \in P_f$, we have defined:

\[
\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{A : A \in D(x)\} = \bigcup \{A : A \subseteq E, x(A) = f(A)\} = \{e : e \in E, \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f\}.
\]

- Hence, \textit{sat}(x) is the maximal (zero-valued) minimizer of the submodular function $f_x(A) \triangleq f(A) - x(A)$. 

The sat function \(=\) Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in \(\mathcal{D}(x)\), also called the polymatroid closure or sat (saturation function).
- For some \(x \in P_f\), we have defined:

\[
\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \\
= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \\
= \{e : e \in E, \forall \alpha > 0, x + \alpha 1_e \notin P_f\} 
\]

Eq. (12.58) says that sat consists of any point \(x\) that is \(P_f\) saturated (any additional positive movement, in that dimension, leaves \(P_f\)). We’ll revisit this in a few slides.
Matroid closure is generalized by the unique maximal element in \( D(x) \), also called the polymatroid closure or \( \text{sat} \) (saturation function).

For some \( x \in P_f \), we have defined:

\[
\text{cl}(x) \overset{\text{def}}{=} \text{sat}(x) \overset{\text{def}}{=} \bigcup \{ A : A \in D(x) \}
\]

\[
= \bigcup \{ A : A \subseteq E, x(A) = f(A) \} \tag{12.57}
\]

\[
= \{ e : e \in E, \forall \alpha > 0, x + \alpha 1_e \notin P_f \} \tag{12.58}
\]

Hence, \( \text{sat}(x) \) is the maximal (zero-valued) minimizer of the submodular function \( f_x(A) \triangleq f(A) - x(A) \).

Eq. (12.58) says that \( \text{sat} \) consists of any point \( x \) that is \( P_f \) saturated (any additional positive movement, in that dimension, leaves \( P_f \)). We’ll revisit this in a few slides.

First, we see how \( \text{sat} \) generalizes matroid closure.