Submodular Functions, Optimization, and Applications to Machine Learning

— Spring Quarter, Lecture 12 —

http://www.ee.washington.edu/people/faculty/bilmes/classes/ee596b_spring_2016/

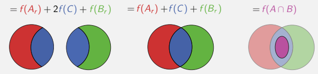
Prof. Jeff Bilmes

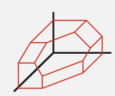
University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

May 11th, 2016

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

$$= f(A_r) + 2f(C) + f(B_r) \qquad = f(A_r) + f(C) + f(B_r) \qquad = f(A \cap B)$$





Cumulative Outstanding Reading

- Read chapters 2 and 3 from Fujishige's book.
- Read chapter 1 from Fujishige's book.

Announcements, Assignments, and Reminders

- Homework 4, soon available at our assignment dropbox (https://canvas.uw.edu/courses/1039754/assignments)
- Homework 3, available at our assignment dropbox (https://canvas.uw.edu/courses/1039754/assignments), due (electronically) Monday (5/2) at 11:55pm.
- Homework 2, available at our assignment dropbox (https://canvas.uw.edu/courses/1039754/assignments), due (electronically) Monday (4/18) at 11:55pm.
- Homework 1, available at our assignment dropbox (https://canvas.uw.edu/courses/1039754/assignments), due (electronically) Friday (4/8) at 11:55pm.
- Weekly Office Hours: Mondays, 3:30-4:30, or by skype or google hangout (set up meeting via our our discussion board (https://canvas.uw.edu/courses/1039754/discussion_topics)).

Class Road Map - IT-I

- L1(3/28): Motivation, Applications, & Basic Definitions
- L2(3/30): Machine Learning Apps (diversity, complexity, parameter, learning target, surrogate).
- L3(4/4): Info theory exs, more apps, definitions, graph/combinatorial examples, matrix rank example, visualization
- L4(4/6): Graph and Combinatorial Examples, matrix rank, Venn diagrams, examples of proofs of submodularity, some useful properties
- L5(4/11): Examples & Properties, Other Defs., Independence
- L6(4/13): Independence, Matroids, Matroid Examples, matroid rank is submodular
- L7(4/18): Matroid Rank, More on Partition Matroid, System of Distinct Reps, Transversals, Transversal Matroid,
- L8(4/20): Transversals, Matroid and representation, Dual Matroids,
- L9(4/25): Dual Matroids, Properties, Combinatorial Geometries, Matroid and Greedy
- L10(4/27): Matroid and Greedy, Polyhedra, Matroid Polytopes,
- L11(5/2): From Matroids to Polymatroids, Polymatroids
- L12(5/4): Polymatroids, Polymatroids and Greedy, Possible Polytopes, Extreme Points
- L13(5/9):
- L14(5/11):
- L15(5/16):
- L16(5/18):
- L17(5/23):
- L18(5/25):
- L19(6/1):
- L20(6/6): Final Presentations maximization.

Finals Week: June 6th-10th, 2016.

P-basis of x given compact set $P \subseteq \mathbb{R}_+^E$

Definition 12.2.1 (subvector)

y is a subvector of x if $y \leq x$ (meaning $y(e) \leq x(e)$ for all $e \in E$).

Definition 12.2.2 (P-basis)

Given a compact set $P \subseteq \mathcal{R}_+^E$, for any $x \in \mathbb{R}_+^E$, a subvector y of x is called a P-basis of x if y maximal in P.
In other words, y is a P-basis of x if y is a maximal P-contained subvector of x.

Here, by y being "maximal", we mean that there exists no $z > y$ (more precisely, no $z \geq y + \epsilon \mathbf{1}_e$ for some $e \in E$ and $\epsilon > 0$) having the properties of y (the properties of y being: in P, and a subvector of x).
In still other words: y is a P-basis of x if:

1. $y \leq x$ (y is a subvector of x); and
2. $y \in P$ and $y + \epsilon \mathbf{1}_e \notin P$ for all $e \in E$ where $y(e) < x(e)$ and $\forall \epsilon > 0$ (y is maximal P-contained).

A vector form of rank

- Recall the definition of rank from a matroid $M = (E, \mathcal{I})$.

$$\text{rank}(A) = \max\{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \qquad (12.1)$$

- vector rank: Given a compact set $P \subseteq \mathbb{R}_+^E$, we can define a form of "vector rank" relative to this P in the following way: Given an $x \in \mathbb{R}^E$, we define the vector rank, relative to P, as:

$$\text{rank}(x) = \max\left(y(E) : y \leq x, y \in P\right) = \max_{y \in P} (x \wedge y)(E) \qquad (12.2)$$

  where $y \leq x$ is componentwise inequality ($y_i \leq x_i, \forall i$), and where $(x \wedge y) \in \mathbb{R}_+^E$ has $(x \wedge y)(i) = \min(x(i), y(i))$.
- If $\mathcal{B}_x$ is the set of P-bases of x, than $\text{rank}(x) = \max_{y \in \mathcal{B}_x} y(E)$.
- If $x \in P$, then $\text{rank}(x) = x(E)$ (x is its own unique self P-basis).
- If $x_{\text{min}} = \min_{x \in P} x(E)$, and $x \leq x_{\text{min}}$ what then? $-\infty$?
- In general, might be hard to compute and/or have ill-defined properties. Next, we look at an object that restrains and cultivates this form of rank.

Polymatroidal polyhedron (or a "polymatroid")

Definition 12.2.1 (polymatroid)

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 $y(E)$

Matroid and Polymatroid: side-by-side

A Matroid is:

1. a set system $(E, \mathcal{I})$
2. empty-set containing $\emptyset \in \mathcal{I}$
3. down closed, $\emptyset \subseteq I' \subseteq I \in \mathcal{I} \Rightarrow I' \in \mathcal{I}$.
4. any maximal set I in $\mathcal{I}$, bounded by another set A, has the same matroid rank (any maximal independent subset $I \subseteq A$ has same size $|I|$).

A Polymatroid is:

1. a compact set $P \subseteq \mathbb{R}_+^E$
2. zero containing, $\mathbf{0} \in P$
3. down monotone, $0 \leq y \leq x \in P \Rightarrow y \in P$
4. any maximal vector y in P, bounded by another vector x, has the same vector rank (any maximal independent subvector $y \leq x$ has same sum $y(E)$).

Polymatroid function and its polyhedron.

Definition 12.2.1

A polymatroid function is a real-valued function f defined on subsets of E which is normalized, non-decreasing, and submodular. That is we have

1. $f(\emptyset) = 0$ (normalized)
2. $f(A) \leq f(B)$ for any $A \subseteq B \subseteq E$ (monotone non-decreasing)
3. $f(A \cup B) + f(A \cap B) \leq f(A) + f(B)$ for any $A, B \subseteq E$ (submodular)

We can define the polyhedron P_f^+ associated with a polymatroid function as follows

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \tag{12.1}$$

$$= \left\{y \in \mathbb{R}^E : y \geq 0, y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \tag{12.2}$$

A polymatroid vs. a polymatroid function's polyhedron

- Summarizing the above, we have:
  - Given a polymatroid function f, its associated polytope is given as

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \tag{12.10}$$

  - We also have the definition of a polymatroidal polytope P (compact subset, zero containing, down-monotone, and $\forall x$ any maximal independent subvector $y \leq x$ has same component sum $y(E)$).
- Is there any relationship between these two polytopes?
- In the next theorem, we show that any P_f^+-basis has the same component sum, when f is a polymatroid function, and P_f^+ satisfies the other properties so that P_f^+ is a polymatroid.

A polymatroid function's polyhedron is a polymatroid.

Theorem 12.2.1

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

$$y^x(E) = \textit{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \qquad (12.10)$$

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

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

$$\text{rank}\left(\frac{1}{\epsilon}\mathbf{1}_{E\setminus B}\right) = f(B) = \max\left\{y(B) : y \in P_f^+\right\} \qquad (12.11)$$

In fact, we will ultimately see a number of important consequences of this theorem (other than just that P_f^+ is a polymatroid)

A polymatroid function's polyhedron is a polymatroid.

Proof.

- Clearly $0 \in P_f^+$ since f is non-negative.
- Also, for any $y \in P_f^+$ then any $x <= y$ is also such that $x \in P_f^+$. So, P_f^+ is down-monotone.
- Now suppose that we are given an $x \in \mathbb{R}^E_+$, and maximal $y^x \in P_f^+$ with $y^x \leq x$ (i.e., y^x is a P_f^+-basis of x).
- Goal is to show that any such y^x has $y^x(E) =$ const, dependent only on x and also f (which defines the polytope) but not dependent on y^x, the particular P_f^+-basis.
- Doing so will thus establish that P_f^+ is a polymatroid.

...

A polymatroid function's polyhedron is a polymatroid.

. . . proof continued.

- First trivial case: could have $y^x = x$, which happens if $x(A) \leq f(A), \forall A \subseteq E$ (i.e., $x \in P_f^+$ strictly). In such case,

$$\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{12.10}$$
$$= x(E) + \min\left(f(E \setminus A) - x(E \setminus A) : A \subseteq E\right) \tag{12.11}$$
$$= x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) \tag{12.12}$$
$$= x(E) \tag{12.13}$$

- When $x \in P_f^+$, $y = x$ is clearly the solution to $\max\left(y(E) : y \leq x, y \in P_f^+\right)$, so this is tight, and rank$(x) = x(E)$.
- This is a value dependent only on x and not on any of its P_f^+-bases.

. . .

A polymatroid function's polyhedron is a polymatroid.

. . . proof continued.

- 2nd trivial case: $x(A) > f(A), \forall A \subseteq E$ (i.e., $x \notin P_f^+$ every direction),
- Then for any order $(a_1, a_2, \dots)$ of the elements and $A_i \triangleq (a_1, a_2, \dots, a_i)$, we have $x(a_i) \geq f(a_i) \geq f(a_i|A_{i-1})$, the second inequality by submodularity. This gives

$$\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{12.10}$$
$$= x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) \tag{12.11}$$
$$= x(E) + \min\left(\sum_i f(a_i|A_{i-1}) - \sum_i x(a_i) : A \subseteq E\right) \tag{12.12}$$
$$= x(E) + \min\left(\sum_i \underbrace{\left(f(a_i|A_{i-1}) - x(a_i)\right)}_{\leq 0} : A \subseteq E\right) \tag{12.13}$$
$$= x(E) + f(E) - x(E) = f(E) = \max(y(E) : y \in P_f^+). \tag{12.14}$$

. . .

A polymatroid function's polyhedron is a polymatroid.

... proof continued.

- Assume neither trivial case. Because $y^x \in P_f^+$, we have that $y^x(A) \leq f(A)$ for all $A \subseteq E$.
- We show that the constant is given by

$$y^x(E) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{12.1}$$

- For any P_f^+-basis y^x of x, and any $A \subseteq E$, we have weak relationship:

$$y^x(E) = y^x(A) + y^x(E \setminus A) \tag{12.2}$$

$$\leq x(A) + f(E \setminus A). \tag{12.3}$$

This follows since $y^x \leq x$ and since $y^x \in P_f^+$.

- This ensures

$$\max\left(y(E) : y \leq x, y \in P_f^+\right) \leq \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{12.4}$$

- Given an A where equality in Eqn. (12.3) holds, above min result follows.

...

A polymatroid function's polyhedron is a polymatroid.

... proof continued.

- For any $y \in P_f^+$, call a set $B \subseteq E$ tight if $y(B) = f(B)$. The union (and intersection) of tight sets B, C is again tight, since

$$f(B) + f(C) = y(B) + y(C) \tag{12.5}$$

$$= y(B \cap C) + y(B \cup C) \tag{12.6}$$

$$\leq f(B \cap C) + f(B \cup C) \tag{12.7}$$

$$\leq f(B) + f(C) \tag{12.8}$$

which requires equality everywhere above.

- Because $y(A) \leq f(A), \forall A$, this means $y(B \cap C) = f(B \cap C)$ and $y(B \cup C) = f(B \cup C)$, so both also are tight.
- For $y \in P_f^+$, it will be ultimately useful to define this lattice family of tight sets: $\mathcal{D}(y) \triangleq \{A : A \subseteq E,\ y(A) = f(A)\}$.

...

A polymatroid function's polyhedron is a polymatroid.

... proof continued.

- Also, we define $\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\}$.
- Consider again a P_f^+-basis y^x (so maximal).
- Given a $e \in E$, either $y^x(e)$ is cut off due to x (so $y^x(e) = x(e)$) or e is saturated by f, meaning it is an element of some tight set and $e \in \text{sat}(y^x)$.
- Let $E \setminus A = \text{sat}(y^x)$ be the union of all such tight sets (which is also tight, so $y^x(E \setminus A) = f(E \setminus A)$).
- Hence, we have

$$y^x(E) = y^x(A) + y^x(E \setminus A) = x(A) + f(E \setminus A) \tag{12.9}$$

- So we identified the A to be the elements that are non-tight, and achieved the min, as desired. □

A polymatroid is a polymatroid function's polytope

- So, when f is a polymatroid function, P_f^+ is a polymatroid.
- Is it the case that, conversely, for any polymatroid P, there is an associated polymatroidal function f such that $P = P_f^+$?

Theorem 12.3.1

For any polymatroid P (compact subset of $\mathbb{R}_+^E$, zero containing, down-monotone, and $\forall x \in \mathbb{R}_+^E$ any maximal independent subvector $y \leq x$ has same component sum $y(E) = \text{rank}(x)$), there is a polymatroid function $f : 2^E \to \mathbb{R}$ (normalized, monotone non-decreasing, submodular) such that $P = P_f^+$ where $P_f^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\}$.

Tight sets $\mathcal{D}(y)$ are closed, and max tight set $\text{sat}(y)$

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

$$\mathcal{D}(y) \triangleq \{A : A \subseteq E,\ y(A) = f(A)\} \tag{12.10}$$

Theorem 12.3.2

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 11.4.1 □

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

$$\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\} \tag{12.11}$$

Join $\vee$ and meet $\wedge$ for $x, y \in \mathbb{R}_+^E$

- For $x, y \in \mathbb{R}_+^E$, define vectors $x \wedge y \in \mathbb{R}_+^E$ and $x \vee y \in \mathbb{R}_+^E$ such that, for all $e \in E$

$$(x \vee y)(e) = \max(x(e), y(e)) \tag{12.12}$$
$$(x \wedge y)(e) = \min(x(e), y(e)) \tag{12.13}$$

  Hence,

$$x \vee y \triangleq \Big(\max\big(x(e_1), y(e_1)\big), \max\big(x(e_2), y(e_2)\big), \dots, \max\big(x(e_n), y(e_n)\big)\Big)$$

  and similarly

$$x \wedge y \triangleq \Big(\min\big(x(e_1), y(e_1)\big), \min\big(x(e_2), y(e_2)\big), \dots, \min\big(x(e_n), y(e_n)\big)\Big)$$

- From this, we can define things like an lattices, and other constructs.

Vector rank, $\text{rank}(x)$, is submodular

- Recall that the matroid rank function is submodular.
- The vector rank function $\text{rank}(x)$ also satisfies a form of submodularity, namely one defined on the real lattice.

Theorem 12.3.3 (vector rank and submodularity)

Let P be a polymatroid polytope. The vector rank function rank : $\mathbb{R}_+^E \to \mathbb{R}$ with rank$(x) = \max(y(E) : y \leq x, y \in P)$ satisfies, for all $u, v \in \mathbb{R}_+^E$

$$\text{rank}(u) + \text{rank}(v) \geq \text{rank}(u \vee v) + \text{rank}(u \wedge v) \tag{12.14}$$

Vector rank $\text{rank}(x)$ is submodular, proof

Proof of Theorem 12.3.3.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...
- ...and since $b \leq u \vee v$, we get

$$a + b = b \wedge u \wedge v + b = b \wedge u + b \wedge v \tag{12.15}$$

To see this, consider each case where either b is the minimum, or u is minimum with $b \leq v$, or v is minimum with $b \leq u$.

...

Vector rank $\text{rank}(x)$ is submodular, proof

. . . proof of Theorem 12.3.3.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.
- Hence,

$$\begin{aligned} \text{rank}(u \wedge v) + \text{rank}(u \vee v) &= a(E) + b(E) & (12.16) \\ &= (b \wedge u)(E) + (b \wedge v)(E) & (12.17) \\ &\leq \text{rank}(u) + \text{rank}(v) & (12.18) \end{aligned}$$

□

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 12.3.3 and the proof of Theorem 6.6.1 that the standard matroid rank function is submodular.
- Next, we prove Theorem 12.3.1, that any polymatroid polytope P has a polymatroid function f such that $P = P_f^+$.
- Given this result, we can conclude that a polymatroid is really an extremely natural polyhedral generalization of a matroid. This was all realized by Jack Edmonds in the mid 1960s (and published in 1969 in his landmark paper "Submodular Functions, Matroids, and Certain Polyhedra").

Proof of Theorem 12.3.1

Proof of Theorem 12.3.1.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max \{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha \to \infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\text{max}}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) \tag{12.19}$$

- Then f is submodular since

$$\begin{aligned} f(A) + f(B) &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_B) && (12.20) \\ &\geq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \vee \alpha_{\text{max}} \mathbf{1}_B) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \wedge \alpha_{\text{max}} \mathbf{1}_B) && (12.21) \\ &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cup B}) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cap B}) && (12.22) \\ &= f(A \cup B) + f(A \cap B) && (12.23) \end{aligned}$$

...

Proof of Theorem 12.3.1

Proof of Theorem 12.3.1.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).
- Hence, f is a polymatroid function.
- Consider the polytope P_f^+ defined as:

$$P_f^+ = \left\{ x \in \mathbb{R}_+^E : x(A) \leq f(A), \; \forall A \subseteq E \right\} \tag{12.24}$$

- Given an $x \in P$, then for any $A \subseteq E$, $x \leq \alpha_{\text{max}} \mathbf{1}_A$, so $x(A) \leq \max \{z(E) : z \in P, z \leq \alpha_{\text{max}} \mathbf{1}_A\} = \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) = f(A)$, therefore $x \in P_f^+$.
- Hence, $P \subseteq P_f^+$.
- We will next show that $P_f^+ \subseteq P$ to complete the proof.

...

Proof of Theorem 12.3.1

Proof of Theorem 12.3.1.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).
- Suppose $x \notin P$. Then, choose y to be a P-basis of x that maximizes the number of y elements strictly less than the corresponding x element. I.e., that maximizes $|N(y)|$, where

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{12.25}$$

- Choose w between y and x, so that

$$y \leq w \triangleq (y + x)/2 \leq x \tag{12.26}$$

  so y is also a P-basis of w.
- Hence, $\text{rank}(x) = \text{rank}(w) = y(E)$, and the set of P-bases of w are also P-bases of x.

Proof of Theorem 12.3.1

Proof of Theorem 12.3.1.

- For any $A \subseteq E$, define $x_A \in \mathbb{R}_+^E$ as

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \tag{12.27}$$

  note this is an analogous definition to $\mathbf{1}_A$ but for a non-unity vector.
- Now, we have

$$y(N(y)) < w(N(y)) \leq f(N(y)) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_{N(y)}) \tag{12.28}$$

  the last inequality follows since $w \leq x \in P_f^+$, and $y \leq w$.
- Thus, $y \wedge x_{N(y)}$ is not a P-basis of $w \wedge x_{N(y)}$ since, over $N(y)$, it is neither tight at w nor tight at the rank (i.e., not a maximal independent subvector on $N(y)$).

Proof of Theorem 12.3.1

Proof of Theorem 12.3.1.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),
- hence $\hat{y}(e) < y(e)$ for some $e \notin N(y)$.
- Thus, $\hat{y}$ is a base of x, which violates the maximality of $|N(y)|$.
- This contradiction means that we must have had $x \in P$.
- Therefore, $P_f^+ = P$.

□

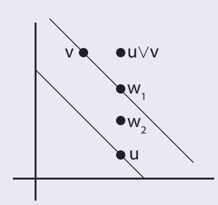
More on polymatroids

Theorem 12.3.4

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*
2. *If $u, v \in P$ (i.e., are independent) and $u(E) < v(E)$, then there exists a vector $w \in P$ such that*

$$u < w \leq u \vee v \qquad (12.29)$$



Corollary 12.3.5

The independent vectors of a polymatroid form a convex polyhedron in $\mathbb{R}_+^E$.

Review

- The next slide comes from lecture 6.

Matroids by bases

In general, besides independent sets and rank functions, there are other equivalent ways to characterize matroids.

Theorem 12.3.3 (Matroid (by bases))

Let E be a set and $\mathcal{B}$ be a nonempty collection of subsets of E. Then the following are equivalent.

1. *$\mathcal{B}$ is the collection of bases of a matroid;*
2. *if $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B' - x + y \in \mathcal{B}$ for some $y \in B \setminus B'$.*
3. *If $B, B' \in \mathcal{B}$, and $x \in B' \setminus B$, then $B - y + x \in \mathcal{B}$ for some $y \in B \setminus B'$.*

Properties 2 and 3 are called "exchange properties."
Proof here is omitted but think about this for a moment in terms of linear spaces and matrices, and (alternatively) spanning trees.

More on polymatroids

For any compact set P, b is a base of P if it is a maximal subvector within P. Recall the bases of matroids. In fact, we can define a polymatroid via vector bases (analogous to how a matroid can be defined via matroid bases).

Theorem 12.3.6

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*
2. *if b, c are bases of P and d is such that $b \wedge c < d < b$, then there exists an f, with $d \wedge c < f \leq c$ such that $d \vee f$ is a base of P*
3. *All of the bases of P have the same rank.*

Note, all three of the above are required for a polymatroid (a matroid analogy would require the equivalent of only the first two).

A word on terminology & notation

- Recall how a matroid is sometimes given as (E, r) where r is the rank function.
- We mention also that the term "polymatroid" is sometimes not used for the polytope itself, but instead but for the pair (E, f),
- But now we see that (E, f) is equivalent to a polymatroid polytope, so this is sensible.

Where are we going with this?

- Consider the right hand side of Theorem 11.4.1: $\min\left(x(A) + f(E \setminus A) : A \subseteq E\right)$
- We are going to study this problem, and approaches that address it, as part of our ultimate goal which is to present strategies for submodular function minimization (that we will ultimately get to, in near future lectures).
- As a bit of a hint on what's to come, recall that we can write it as: $x(E) + \min\left(f(A) - x(A) : A \subseteq E\right)$ where f is a polymatroid function.

Another Interesting Fact: Matroids from polymatroid functions

Theorem 12.3.7

Given integral polymatroid function f, let $(E, \mathcal{F})$ be a set system with ground set E and set of subsets $\mathcal{F}$ such that

$$\forall F \in \mathcal{F}, \ \ \forall \emptyset \subset S \subseteq F, |S| \leq f(S) \tag{12.30}$$

Then $M = (E, \mathcal{F})$ is a matroid.

Proof.

Exercise □

And its rank function is Exercise.

Matroid instance of Theorem 11.4.1

- Considering Theorem 11.4.1, the matroid case is now a special case, where we have that:

Corollary 12.3.8

We have that:

$$\max \{y(E) : y \in P_{\textit{ind. set}}(M), y \leq x\} = \min \{r_M(A) + x(E \setminus A) : A \subseteq E\} \tag{12.31}$$

where r_M is the matroid rank function of some matroid.

Review

- The next two slides come respectively from Lecture 11 and Lecture 10.

Polymatroidal polyhedron (or a "polymatroid")

Definition 12.4.1 (polymatroid)

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 $y(E)$

Maximum weight independent set via greedy weighted rank

Theorem 12.4.5

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{12.19}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{12.20}$$

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 $\max\left\{wx : x \in P_f^+\right\}$, 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, ultimately, even relax things so that $w \in \mathbb{R}^E$?

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- 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 $0 \geq w(e_{k+1})$.
  That is, we have

$$w(e_1) \geq w(e_2) \geq \cdots \geq w(e_k) > 0 \geq w(e_{k+1}) \geq \cdots \geq w(e_m) \tag{12.32}$$

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

$$E_i \stackrel{\text{def}}{=} \{e_1, e_2, \ldots e_i\} \tag{12.34}$$

(note $E_0 = \emptyset$, $f(E_0) = 0$, and $\underline{E \text{ and } E_i \text{ is always sorted w.r.t } w}$).
- The greedy solution is the vector $x \in \mathbb{R}_+^E$ with elements defined as:

$$x(e_1) \stackrel{\text{def}}{=} f(E_1) = f(e_1) = f(e_1|E_0) = f(e_1|\emptyset) \tag{12.35}$$

$$x(e_i) \stackrel{\text{def}}{=} f(E_i) - f(E_{i-1}) = f(e_i|E_{i-1}) \text{ for } i = 2 \ldots k \tag{12.36}$$

$$x(e_i) \stackrel{\text{def}}{=} 0 \text{ for } i = k+1 \ldots m = |E| \tag{12.37}$$

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 (as we will show) 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.

- Consider the LP strong duality equation:

$$\max(wx : x \in P^+_f) = \min\Big(\sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}^{2^E}_+, \sum_{A \subseteq E} y_A \mathbf{1}_A \geq w\Big) \tag{12.38}$$

- Sort E by w, and define the following vector $y \in \mathbb{R}^{2^E}_+$ as

$$y_{E_i} \leftarrow w(e_i) - w(e_{i+1}) \text{ for } i = 1 \ldots (m-1), \tag{12.39}$$
$$y_E \leftarrow w(e_m), \text{ and} \tag{12.40}$$
$$y_A \leftarrow 0 \text{ otherwise} \tag{12.41}$$

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 $A = (a_1, a_2, \dots, a_k)$ based on order $(e_1, e_2, \dots, e_m)$.

		a_1		a_2	a_3			a_4		a_5	$\dots$	
e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	$\dots$	e_m

- Define $e^{-1} : E \to \{1, \dots, m\}$ so that $e^{-1}(e_i) = i$.
- Then, we have $x \in P_f^+$ since for all A:

$$f(A) = \sum_{i=1}^{k} f(a_i | a_{1:i-1}) \tag{12.42}$$

$$\geq \sum_{i=1}^{k} f(a_i | e_{1:e^{-1}(a_i)-1}) \tag{12.43}$$

$$= \sum_{a \in A} f(a | e_{1:e^{-1}(a)-1}) = x(A) \tag{12.44}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Next, y is also feasible for the dual constraints in Eq. 12.38 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.:

$$wx = \sum_{e \in E} w(e)x(e) = \sum_{i=1}^{m} w(e_i) f(e_i | E_{i-1}) = \sum_{i=1}^{m} w(e_i) \Big(f(E_i) - f(E_{i-1}) \Big)$$

$$= \sum_{i=1}^{m-1} f(E_i) \Big(w(e_i) - w(e_{i+1}) \Big) + f(E) w(e_m) = \sum_{A \subseteq E} y_A f(A)$$

...

Polymatroidal polyhedron and greedy

Proof.

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

$$wx = \sum_{i=1}^{m} w_i x_i \tag{12.45}$$

$$= \sum_{i=1}^{m} w_i \Big(f(E_i) - f(E_{i-1}) \Big) \tag{12.46}$$

$$= \sum_{i=1}^{m} w_i f(E_i) - \sum_{i=1}^{m-1} w_{i+1} f(E_i) \tag{12.47}$$

$$= w_m f(E_m) + \sum_{i=1}^{m-1} (w_i - w_{i+1}) f(E_i) \tag{12.48}$$

$\square$

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

Proof.

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

$$w \stackrel{\text{def}}{=} \sum_{i=1}^{q} \mathbf{1}_{e_i} = \mathbf{1}_{A \cup B} \tag{12.49}$$

- Suppose optimum solution x is given by the greedy procedure.

...

Polymatroids | Polymatroids and Greedy | Possible Polytopes | Extreme Points

Polymatroidal polyhedron and greedy

Proof.

- Then

$$\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k} (f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B) \tag{12.50}$$

- and

$$\sum_{i=1}^{p} x_i = f(E_1) + \sum_{i=2}^{p} (f(E_i) - f(E_{i-1})) = f(E_p) = f(A) \tag{12.51}$$

- and

$$\sum_{i=1}^{q} x_i = f(E_1) + \sum_{i=2}^{q} (f(E_i) - f(E_{i-1})) = f(E_q) = f(A \cup B) \tag{12.52}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Thus, we have

$$x(B) = \sum_{i \in 1,\ldots,k,p+1,\ldots,q} x_i = \sum_{i:e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \quad (12.53)$$

- 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 $x(B) \le f(B)$.
- Thus,

$$x(B) = f(A \cup B) + f(A \cap B) - f(A) = \sum_{i:e_i \in B} x_i \le f(B) \quad (12.54)$$

ensuring the submodularity of f, since A and B are arbitrary.

□

Review from Lecture 9

- The next slide comes from lecture 9.

Matroid and the greedy algorithm

- Let $(E, \mathcal{I})$ be an independence system, and we are given a non-negative modular weight function $w : E \to \mathbb{R}_+$.

Algorithm 1: The Matroid Greedy Algorithm

1 Set $X \leftarrow \emptyset$;
2 **while** $\exists v \in E \setminus X$ *s.t.* $X \cup \{v\} \in \mathcal{I}$ **do**
3 $\quad v \in \operatorname{argmax} \{w(v) : v \in E \setminus X,\ X \cup \{v\} \in \mathcal{I}\}$;
4 $\quad X \leftarrow X \cup \{v\}$;

- Same as sorting items by decreasing weight w, and then choosing items in that order that retain independence.

Theorem 12.4.7

Let $(E, \mathcal{I})$ be an independence system. Then the pair $(E, \mathcal{I})$ is a matroid if and only if for each weight function $w \in \mathcal{R}_+^E$, Algorithm 1 above leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Polymatroidal polyhedron and greedy

- 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 10.5.1)

Theorem 12.4.1

If $f : 2^E \to \mathbb{R}_+$ is given, and P is a polytope in $\mathbb{R}_+^E$ of the form $P = \left\{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\right\}$, 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 $f : 2^V \to R$ (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:

$$f'(A) = \begin{cases} f(A) & \text{if } A \neq \emptyset \\ 0 & \text{if } A = \emptyset \end{cases} \tag{12.55}$$

  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 $f(\emptyset) \geq 0$.
- We can define several polytopes:

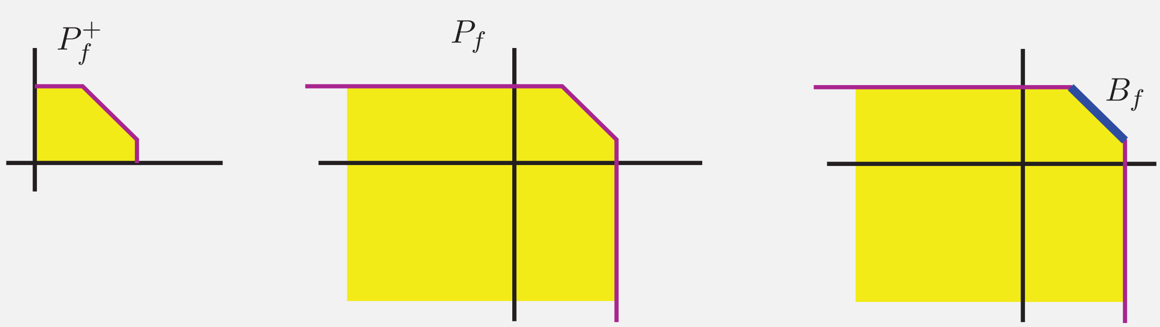
$$P_f = \left\{x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E\right\} \tag{12.56}$$

$$P_f^+ = P_f \cap \left\{x \in \mathbb{R}^E : x \geq 0\right\} \tag{12.57}$$

$$B_f = P_f \cap \left\{x \in \mathbb{R}^E : x(E) = f(E)\right\} \tag{12.58}$$

- 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 f

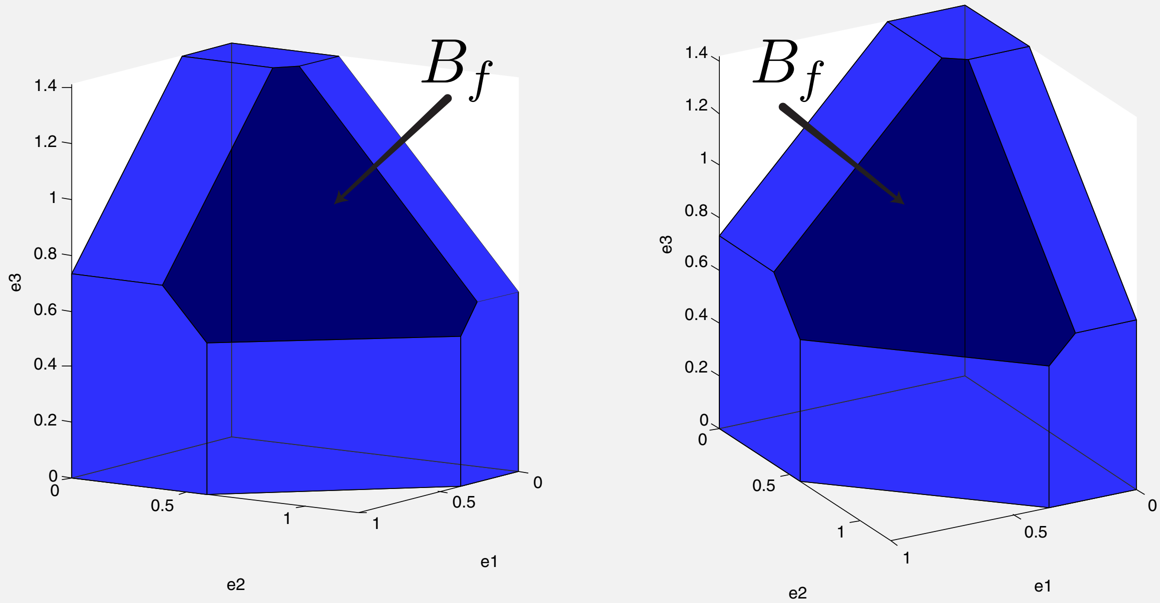


$$P_f^+ = P_f \cap \left\{x \in \mathbb{R}^E : x \geq 0\right\} \tag{12.59}$$

$$P_f = \left\{x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E\right\} \tag{12.60}$$

$$B_f = P_f \cap \left\{x \in \mathbb{R}^E : x(E) = f(E)\right\} \tag{12.61}$$

Base Polytope in 3D



$$P_f = \left\{x \in \mathbb{R}^E : x(S) \leq f(S), \forall S \subseteq E\right\} \tag{12.62}$$

$$B_f = P_f \cap \left\{x \in \mathbb{R}^E : x(E) = f(E)\right\} \tag{12.63}$$

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:

$$\textit{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f\right) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{12.64}$$

Essentially the same theorem as Theorem 11.4.1. 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:

$$\textit{rank}(0) = \max\left(y(E) : y \leq 0, y \in P_f\right) = \min\left(f(A) : A \subseteq E\right) \tag{12.65}$$

Proof of Theorem 12.5.1

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 $\exists T \in \mathcal{D}(x)$ 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^*.

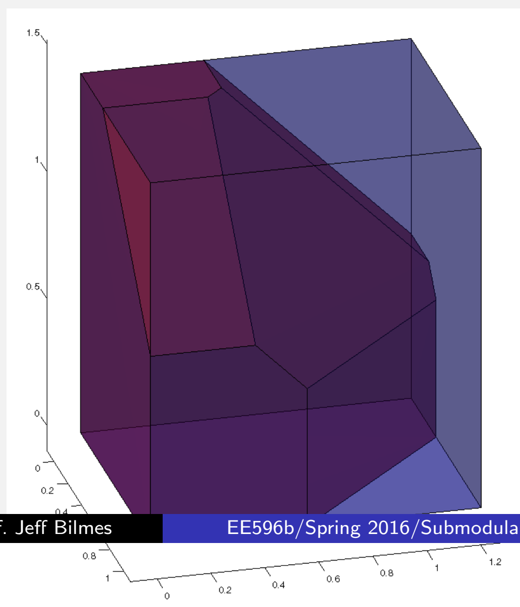
□

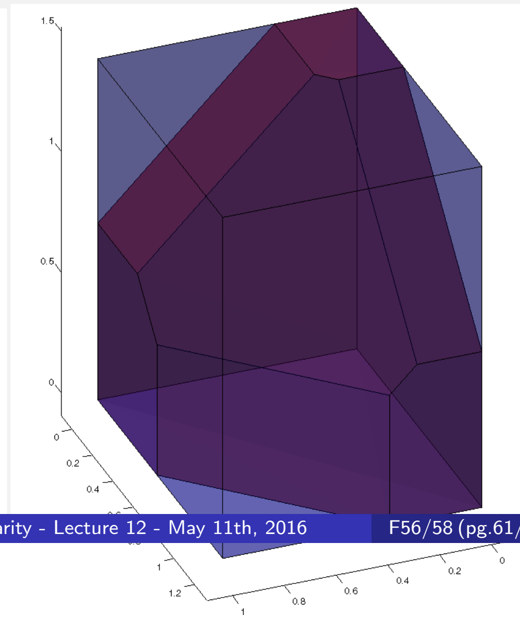
Greedy and P_f

- 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 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.
- Consider P_f^+ and also $C_f^+ \stackrel{\text{def}}{=} \{x : x \in \mathbb{R}_+^E, x(e) \le f(e), \forall e \in E\}$
- Then ordering $A = (a_1, \dots, a_{|A|})$ arbitrarily with $A_i = \{a_1, \dots, a_i\}$, $f(A) = \sum_i f(a_i|A_{i-1}) \le \sum_i f(a_i)$, 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 $e' \in E \setminus \{e\}$).
- 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.,

$$B_f = P_f \cap \{x \in \mathbb{R}_+^E : x(E) = f(E)\} \tag{12.66}$$

- 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 $E_i = (e_1, e_2, \ldots, e_i)$.
- As before, a vector x is generated by E_i using the greedy procedure as follows

$$x(e_1) = f(E_1) = f(e_1) \tag{12.67}$$
$$x(e_j) = f(E_j) - f(E_{j-1}) = f(e_j|E_{j-1}) \text{ for } 2 \leq j \leq i \tag{12.68}$$
$$x(e) = 0 \text{ for } e \in E \setminus E_i \tag{12.69}$$

- 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

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

$$x(E_j) = f(E_j) \text{ for } 1 \leq j \leq i \leq m \tag{12.70}$$
$$x(e) = 0 \text{ for } e \in E \setminus E_i \tag{12.71}$$

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 $x(E_1) = x(e_1) = f(e_1)$
- $x(E_2) = x(e_1) + x(e_2) = f(e_1, e_2)$ 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,

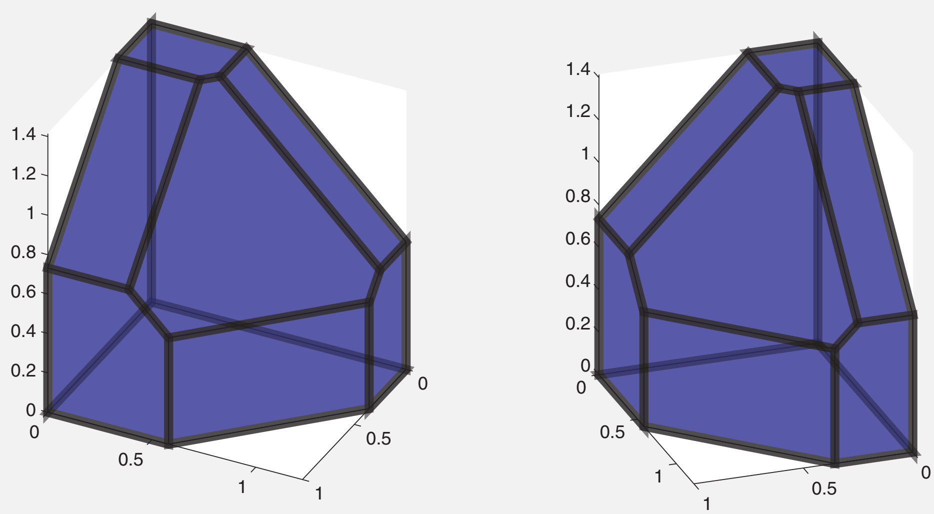
$$x(E_j) = f(E_j) \quad \text{for } 1 \le j \le i \tag{12.72}$$

$$x(A) \le f(A), \forall A \subseteq E \tag{12.73}$$

- 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)

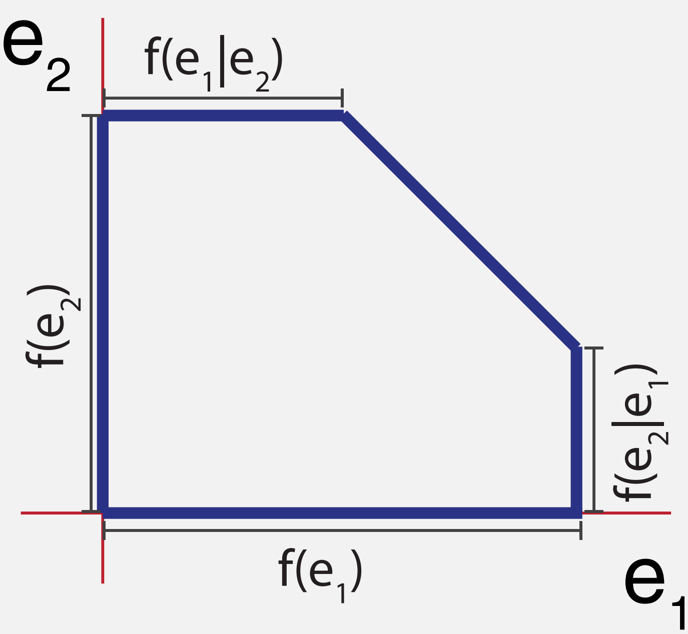
Corollary 12.6.2

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

- Note, $\text{sat}(x) = \text{cl}(x) = \cup(A : x(A) = f(A))$ 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 12.3.2)
- 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.

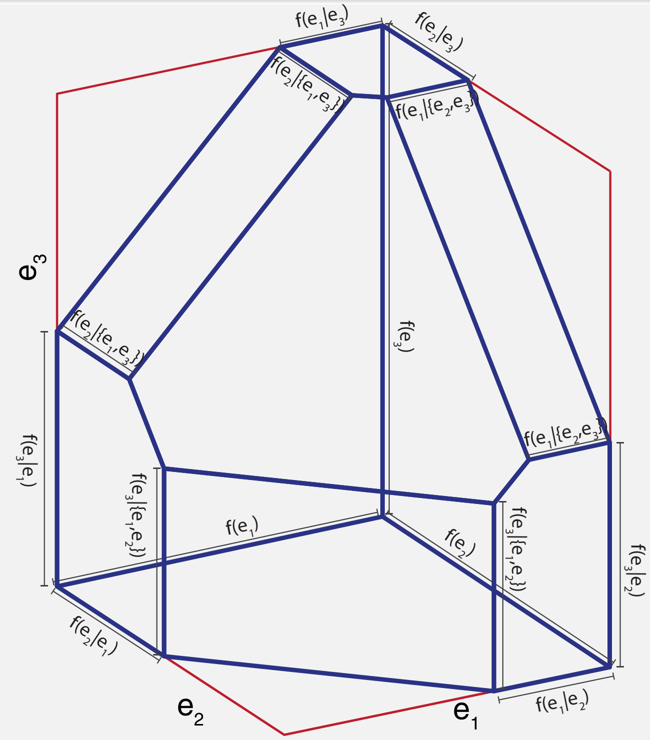
Polymatroid with labeled edge lengths

- 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.



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 $x^\intercal w = x(e_1)w(e_1) + x(e_2)w(e_2)$.
- If $w(e_2) > w(e_1)$ the upper extreme point indicated maximizes $x^\intercal w$ over $x \in P_f^+$.
- If $w(e_2) < w(e_1)$ the lower extreme point indicated maximizes $x^\intercal w$ over $x \in P_f^+$.