Submodular Functions, Optimization, and Applications to Machine Learning

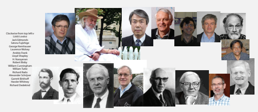
— Spring Quarter, Lecture 11 —

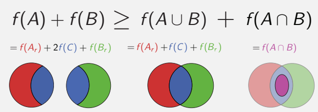
http://www.ee.washington.edu/people/faculty/bilmes/classes/ee596b_spring_2016/

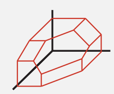
Prof. Jeff Bilmes

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

May 9th, 2016







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):
- 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.

The greedy algorithm

- In combinatorial optimization, the greedy algorithm is often useful as a heuristic that can work quite well in practice.
- The goal is to choose a good subset of items, and the fundamental tenet of the greedy algorithm is to choose next whatever currently looks best, without the possibility of later recall or backtracking.
- Sometimes, this gives the optimal solution (we saw three greedy algorithms that can find the maximum weight spanning tree).
- Greedy is good since it can be made to run very fast $O(n \log n)$.
- Often, however, greedy is heuristic (it might work well in practice, but worst-case performance can be unboundedly poor).
- We will next see that the greedy algorithm working optimally is a defining property of a matroid, and is also a defining property of a polymatroid function.

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

```
Set X ← ∅ ;
while ∃v ∈ E \ X s.t. X ∪ {v} ∈ I do
    v ∈ argmax {w(v) : v ∈ E \ X, X ∪ {v} ∈ I} ;
    X ← X ∪ {v} ;
```

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

Theorem 11.2.7

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 leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Summary of Important (for us) Matroid Definitions

Given an independence system, matroids are defined equivalently by any of the following:

- All maximally independent sets have the same size.
- A monotone non-decreasing submodular integral rank function with unit increments.
- The greedy algorithm achieves the maximum weight independent set for all weight functions.

Matroid Polyhedron in 2D

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{11.10}$$

- Consider this in two dimensions. We have equations of the form:

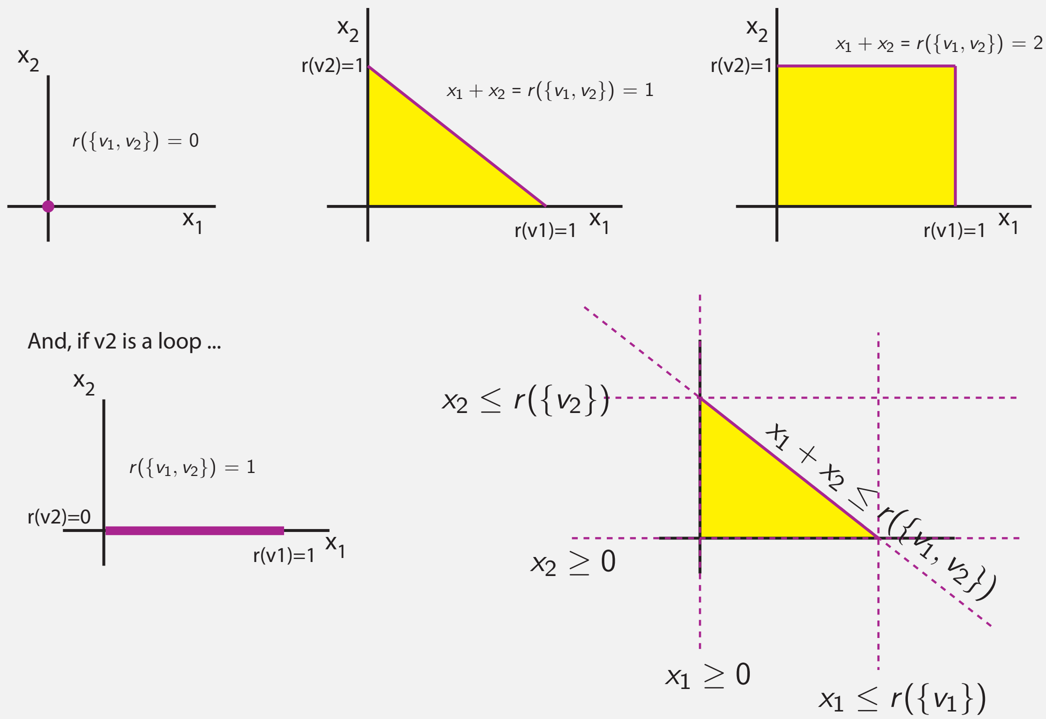
$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{11.11}$$
$$x_1 \leq r(\{v_1\}) \in \{0, 1\} \tag{11.12}$$
$$x_2 \leq r(\{v_2\}) \in \{0, 1\} \tag{11.13}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \in \{0, 1, 2\} \tag{11.14}$$

- Because r is submodular, we have

$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{11.15}$$

  so since $r(\{v_1, v_2\}) \leq r(\{v_1\}) + r(\{v_2\})$, the last inequality is either touching (so inactive) or active.

Matroid Polyhedron in 2D



Independence Polyhedra

- For each $I \in \mathcal{I}$ of a matroid $M = (E, \mathcal{I})$, we can form the incidence vector $\mathbf{1}_I$.
- Taking the convex hull, we get the independent set polytope, that is

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \subseteq [0,1]^E \tag{11.10}$$

- Since $\{\mathbf{1}_I : I \in \mathcal{I}\} \subseteq P_{\text{ind. set}} \subseteq P_r^+$, we have $\max\{w(I) : I \in \mathcal{I}\} \le \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \le \max\{w^\intercal x : x \in P_r^+\}$
- Now take the rank function r of M, and define the following polyhedron:

$$P_r^+ \triangleq \left\{x \in \mathbb{R}^E : x \ge 0, x(A) \le r(A), \forall A \subseteq E\right\} \tag{11.11}$$

- Now, take any $x \in P_{\text{ind. set}}$, then we have that $x \in P_r^+$ (or $P_{\text{ind. set}} \subseteq P_r^+$). We show this next.

$P_{\text{ind. set}} \subseteq P_r^+$

- If $x \in P_{\text{ind. set}}$, then

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{11.10}$$

  for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,

$$\begin{aligned} x(A) = x^\intercal \mathbf{1}_A &= \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A && (11.11)\\ &\leq \sum_i \lambda_i \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) && (11.12)\\ &= \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) = \max_{I \in \mathcal{I}} |A \cap I| && (11.13)\\ &= r(A) && (11.14)\end{aligned}$$

- Thus, $x \in P_r^+$ and hence $P_{\text{ind. set}} \subseteq P_r^+$.

Matroid Independence Polyhedron

- So recall from a moment ago, that we have that

$$\begin{aligned} P_{\text{ind. set}} &= \text{conv}\left\{\cup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} && (11.19)\end{aligned}$$

- In fact, the two polyhedra are identical (and thus both are polytopes).
- We'll show this in the next few theorems.

Maximum weight independent set via greedy weighted rank

Theorem 11.2.5

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

$$\max\{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{11.19}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{11.20}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{subject to} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{11.19}$$

And its convex dual (note $y \in \mathbb{R}_+^{2^n}$, y_U is a scalar element within this exponentially big vector):

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{subject to} \quad & y_U \geq 0 && (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \end{aligned} \tag{11.20}$$

Thanks to strong duality, the solutions to these are equal to each other.

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \quad (11.22)$$

$$= \max\{w^\intercal x : x \in P_r^+\} \quad (11.23)$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \quad (11.24)$$

Polytope Equivalence (Summarizing the above)

- For each $I \in \mathcal{I}$ of a matroid $M = (E, \mathcal{I})$, we can form the incidence vector $\mathbf{1}_I$.
- Taking the convex hull, we get the independent set polytope, that is

$$P_{\text{ind. set}} = \text{conv}\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\} \quad (11.22)$$

- Now take the rank function r of M, and define the following polytope:

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \quad (11.23)$$

Theorem 11.2.5

$$P_r^+ = P_{\textit{ind. set}} \quad (11.24)$$

Maximal points in a set

- Regarding sets, a subset X of S is a maximal subset of S possessing a given property $\mathfrak{P}$ if X possesses property $\mathfrak{P}$ and no set properly containing X (i.e., any $X' \supset X$ with $X' \setminus X \subseteq V \setminus X$) possesses $\mathfrak{P}$.
- Given any compact (essentially closed & bounded) set $P \subseteq \mathbb{R}^E$, we say that a vector x is maximal within P if it is the case that for any $\epsilon > 0$, and for all directions $e \in E$, we have that

$$x + \epsilon \mathbf{1}_e \notin P \tag{11.34}$$

Review from Lecture 6

- The next slide comes from Lecture 6.

Matroids, independent sets, and bases

- Independent sets: Given a matroid $M = (E, \mathcal{I})$, a subset $A \subseteq E$ is called independent if $A \in \mathcal{I}$ and otherwise A is called dependent.
- A base of $U \subseteq E$: For $U \subseteq E$, a subset $B \subseteq U$ is called a base of U if B is inclusionwise maximally independent subset of U. That is, $B \in \mathcal{I}$ and there is no $Z \in \mathcal{I}$ with $B \subset Z \subseteq U$.
- A base of a matroid: If $U = E$, then a "base of E" is just called a base of the matroid M (this corresponds to a basis in a linear space, or a spanning forest in a graph, or a spanning tree in a connected graph).

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

Definition 11.3.1 (subvector)

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

Definition 11.3.2 (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 $M = (E, \mathcal{I})$.

$$\text{rank}(A) = \max \{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{11.1}$$

- 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:

$$\text{rank}(x) = \max (y(E) : y \le x, y \in P) = \max_{y \in P} (x \wedge y)(E) \tag{11.2}$$

where $y \le x$ is componentwise inequality ($y_i \le x_i, \forall i$), 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 $x \le x_{\text{min}}$ 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 11.3.3 (polymatroid)

A polymatroid is a compact set $P \subseteq \mathbb{R}_+^E$ satisfying

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

- Condition 3 restated: That is for any two distinct maximal vectors $y^1, y^2 \in P$, with $y^1 \le x$ & $y^2 \le x$, with $y^1 \ne y^2$, we must have $y^1(E) = y^2(E)$.
- Condition 3 restated (again): For every vector $x \in \mathbb{R}_+^E$, every maximal independent subvector y of x has the same component sum $y(E) = \text{rank}(x)$.
- Condition 3 restated (yet again): All P-bases of x have the same component sum.

Polymatroidal polyhedron (or a "polymatroid")

Definition 11.3.3 (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 $y(E)$

- 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, than $\text{rank}(x) = y(E)$ for any $y \in \mathcal{B}_x$.

Matroid and Polymatroid: side-by-side

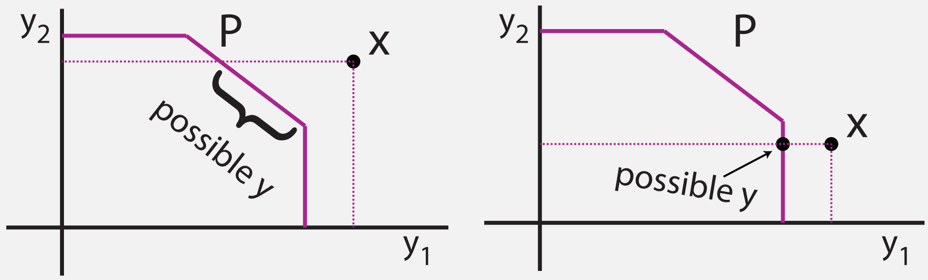
A Matroid is:

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 $y(E)$).

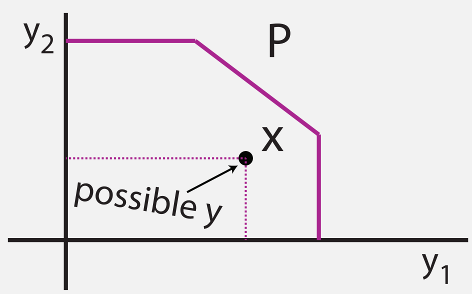
Polymatroidal polyhedron (or a "polymatroid")



Left: ∃ multiple maximal $y \leq x$ Right: ∃ only one maximal $y \leq x$,

- Polymatroid condition here: $\forall$ maximal $y \in P$, with $y \leq x$ (which here means $y_1 \leq x_1$ and $y_2 \leq x_2$), we just have $y(E) = y_1 + y_2 =$ const.
- On the left, we see there are multiple possible maximal $y \in P$ such that $y \leq x$. Each such y must have the same value $y(E)$.
- On the right, there is only one maximal $y \in P$. Since there is only one, the condition on the same value of $y(E), \forall y$ is vacuous.

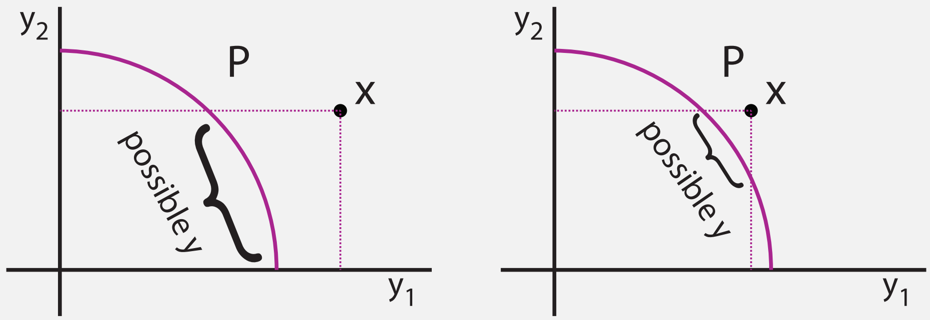
Polymatroidal polyhedron (or a "polymatroid")



∃ only one maximal $y \leq x$.

- If $x \in P$ already, then x is its own P-basis, i.e., it is a self P-basis.
- In a matroid, a base of A is the maximally contained independent set. If A is already independent, then A is a self-base of A (as we saw in Lecture 5)

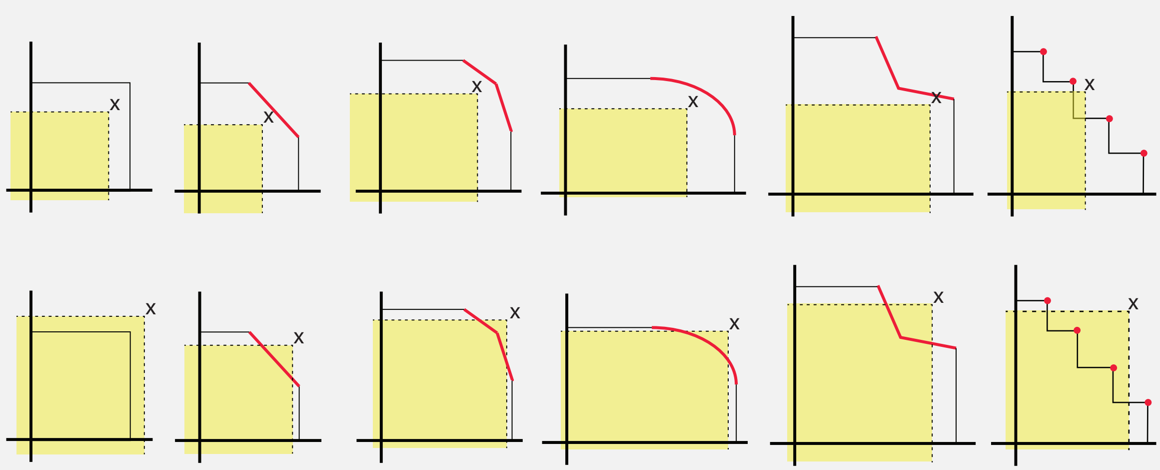
Polymatroid as well? no



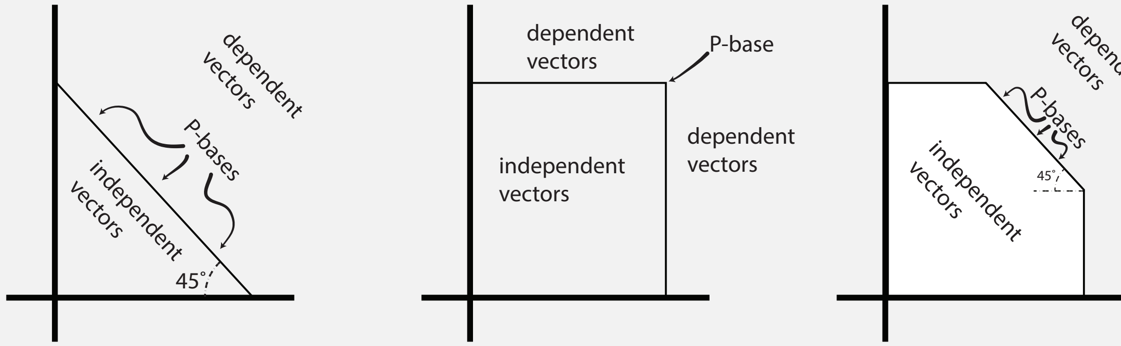
Left and right: $\exists$ multiple maximal $y \leq x$ as indicated.

- On the left, we see there are multiple possible maximal such $y \in P$ that are $y \leq x$. Each such y must have the same value $y(E)$, but since the equation for the curve is $y_1^2 + y_2^2 =$ const. $\neq y_1 + y_2$, we see this is not a polymatroid.
- On the right, we have a similar situation, just the set of potential values that must have the $y(E)$ condition changes, but the values of course are still not constant.

Other examples: Polymatroid or not?



Some possible polymatroid forms in 2D



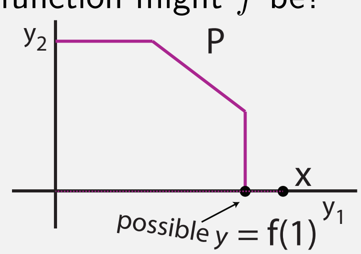
It appears that we have three possible forms of polymatroid in 2D, when neither of the elements $\{v_1, v_2\}$ are self-dependent.

1. On the left: full dependence between v_1 and v_2
2. In the middle: full independence between v_1 and v_2
3. On the right: partial independence between v_1 and v_2

- The P-bases (or single P-base in the middle case) are as indicated.
- Independent vectors are those within or on the boundary of the polytope. Dependent vectors are exterior to the polytope.
- The set of P-bases for a polytope is called the base polytope.

Polymatroidal polyhedron (or a "polymatroid")

- Note that if x contains any zeros (i.e., suppose that $x \in \mathbb{R}_+^E$ has $E \setminus S$ s.t. $x(E \setminus S) = 0$, so S indicates the non-zero elements, or $S = \text{supp}(x)$), then this also forces $y(E \setminus S) = 0$, so that $y(E) = y(S)$. This is true either for $x \in P$ or $x \notin P$.
- Therefore, in this case, it is the non-zero elements of x, corresponding to elements S (i.e., the support $\text{supp}(x)$ of x), determine the common component sum.
- For the case of either $x \notin P$ or right at the boundary of P, we might give a "name" to this component sum, lets say $f(S)$ for any given set S of non-zero elements of x. We could name rank$(\frac{1}{\epsilon}\mathbf{1}_S) \triangleq f(S)$ for ϵ very small. What kind of function might f be?



Polymatroid function and its polyhedron.

Definition 11.3.4

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

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \tag{11.3}$$

$$= \left\{y \in \mathbb{R}^E : y \geq 0, y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \tag{11.4}$$

Associated polyhedron with a polymatroid function

$$P_f^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\right\} \tag{11.5}$$

- Consider this in three dimensions. We have equations of the form:

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{11.6}$$

$$x_1 \leq f(\{v_1\}) \tag{11.7}$$

$$x_2 \leq f(\{v_2\}) \tag{11.8}$$

$$x_3 \leq f(\{v_3\}) \tag{11.9}$$

$$x_1 + x_2 \leq f(\{v_1, v_2\}) \tag{11.10}$$

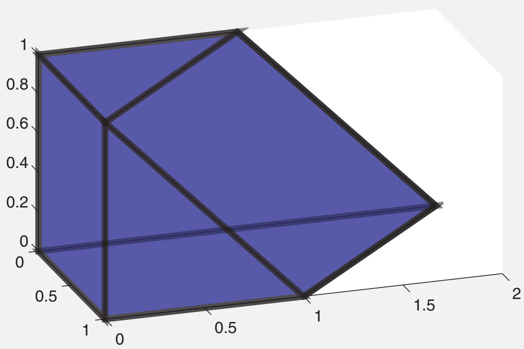
$$x_2 + x_3 \leq f(\{v_2, v_3\}) \tag{11.11}$$

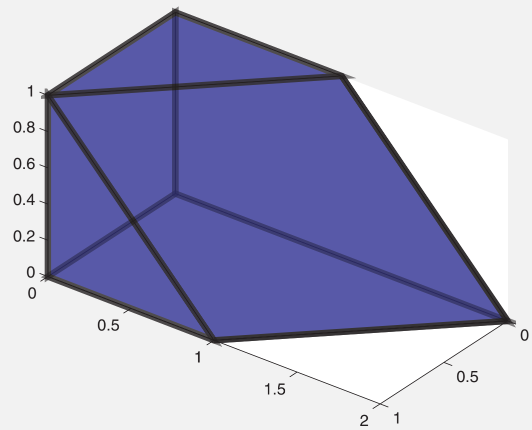
$$x_1 + x_3 \leq f(\{v_1, v_3\}) \tag{11.12}$$

$$x_1 + x_2 + x_3 \leq f(\{v_1, v_2, v_3\}) \tag{11.13}$$

Associated polyhedron with a polymatroid function

- Consider the asymmetric graph cut function on the simple chain graph $v_1 - v_2 - v_3$. That is, $f(S) = |\{(v,s) \in E(G) : v \in V, s \in S\}|$ is count of any edges within S or between S and $V \setminus S$, so that $\delta(S) = f(S) + f(V \setminus S) - f(V)$ is the standard graph cut.
- Observe: P_f^+ (at two views):

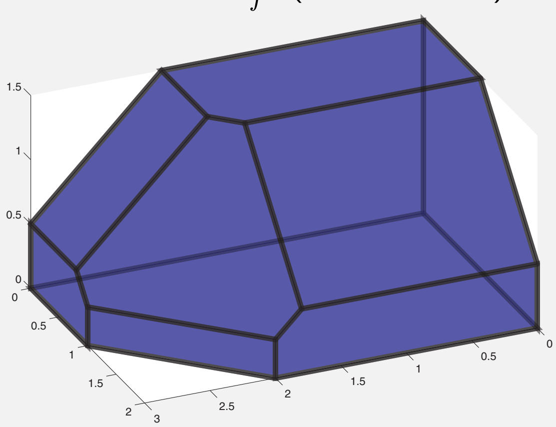


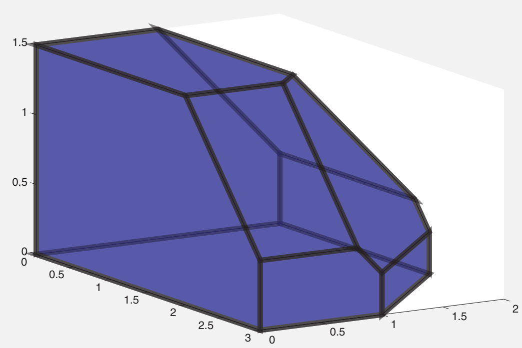


- which axis is which?

Associated polyhedron with a polymatroid function

- Consider: $f(\emptyset) = 0$, $f(\{v_1\}) = 1.5$, $f(\{v_2\}) = 2$, $f(\{v_1, v_2\}) = 2.5$, $f(\{v_3\}) = 3$, $f(\{v_3, v_1\}) = 3.5$, $f(\{v_3, v_2\}) = 4$, $f(\{v_3, v_2, v_1\}) = 4.3$.
- Observe: P_f^+ (at two views):

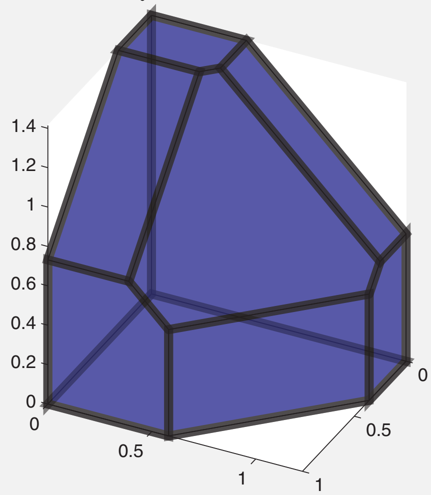


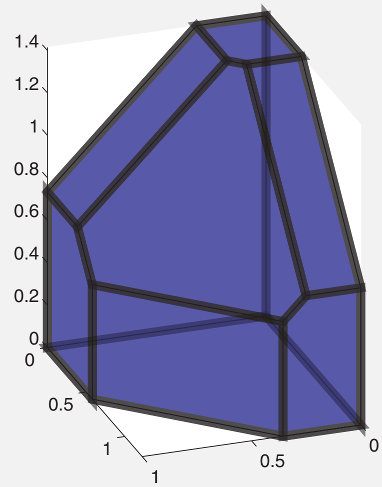


- which axis is which?

Associated polyhedron with a polymatroid function

- Consider modular function $w : V \to \mathbb{R}_+$ as $w = (1, 1.5, 2)^\intercal$, and then the submodular function $f(S) = \sqrt{w(S)}$.
- Observe: P_f^+ (at two views):

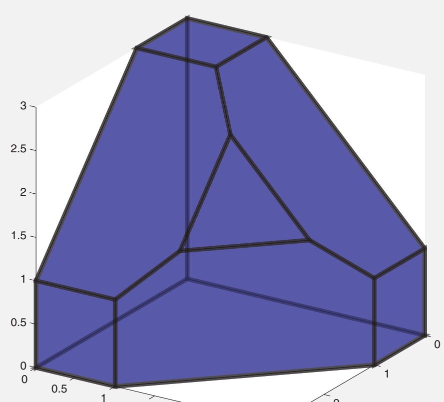


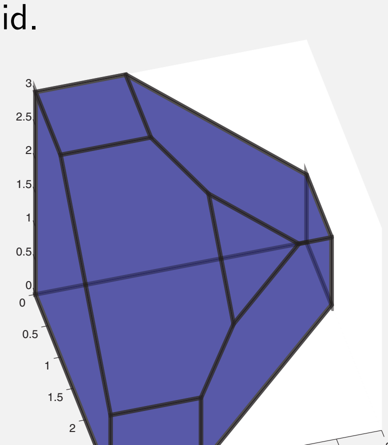


- which axis is which?

Associated polytope with a non-submodular function

- Consider function on integers: $g(0) = 0$, $g(1) = 3$, $g(2) = 4$, and $g(3) = 5.5$. Is $f(S) = g(|S|)$ submodular? $f(S) = g(|S|)$ is not submodular since $f(\{e_1, e_3\}) + f(\{e_1, e_2\}) = 4 + 4 = 8$ but $f(\{e_1, e_2, e_3\}) + f(\{e_1\}) = 5.5 + 3 = 8.5$. Alternatively, consider concavity violation, $1 = g(1 + 1) - g(1) < g(2 + 1) - g(2) = 1.5$.
- Observe: P_f^+ (at two views), maximal independent subvectors not constant rank, hence not a polymatroid.





A polymatroid vs. a polymatroid function's polyhedron

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

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \quad (11.14)$$

  - 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 11.4.1

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

$$y^x(E) = \text{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \quad (11.15)$$

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

Taking $E \setminus B = \text{supp}(x)$ (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:

$$\text{rank}\left(\frac{1}{\epsilon}\mathbf{1}_{E\setminus B}\right) = f(B) = \max\left\{y(B) : y \in P_f^+\right\} \quad (11.16)$$

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-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,

$$\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{11.17}$$
$$= x(E) + \min\left(f(E \setminus A) - x(E \setminus A) : A \subseteq E\right) \tag{11.18}$$
$$= x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) \tag{11.19}$$
$$= x(E) \tag{11.20}$$

...

A polymatroid function's polyhedron is a polymatroid.

... proof continued.

- 2nd trivial case: when $x(A) > f(A), \forall A \subseteq E$ (i.e., $x \notin P_f^+$ strictly),
- 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

$$\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{11.21}$$

$$= x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) \tag{11.22}$$

$$= x(E) + \min\left(\sum_i f(a_i|A_{i-1}) - \sum_i x(a_i) : A \subseteq E\right) \tag{11.23}$$

$$= x(E) + \min\left(\sum_i \underbrace{\left(f(a_i|A_{i-1}) - x(a_i)\right)}_{\leq 0} : A \subseteq E\right) \tag{11.24}$$

$$= x(E) + f(E) - x(E) = f(E) \tag{11.25}$$

...

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

$$y^x(E) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{11.26}$$

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

$$y^x(E) = y^x(A) + y^x(E \setminus A) \tag{11.27}$$

$$\leq x(A) + f(E \setminus A). \tag{11.28}$$

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

- Given one A where equality holds, the 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

$$f(B) + f(C) = y(B) + y(C) \tag{11.29}$$
$$= y(B \cap C) + y(B \cup C) \tag{11.30}$$
$$\leq f(B \cap C) + f(B \cup C) \tag{11.31}$$
$$\leq f(B) + f(C) \tag{11.32}$$

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, 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

$$y^x(E) = y^x(A) + y^x(E \setminus A) = x(A) + f(E \setminus A) \tag{11.33}$$

- 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 11.4.2

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 $y \in P_f^+$:

$$\mathcal{D}(y) \triangleq \{A : A \subseteq E,\ y(A) = f(A)\} \tag{11.34}$$

Theorem 11.4.3

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 $y \in \mathbb{R}_+^E$.

$$\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\} \tag{11.35}$$

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 $e \in E$

$$(x \vee y)(e) = \max(x(e), y(e)) \tag{11.36}$$
$$(x \wedge y)(e) = \min(x(e), y(e)) \tag{11.37}$$

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, rank(x), is submodular

- Recall that the matroid rank function is submodular.
- The vector rank function rank(x) also satisfies a form of submodularity.

Theorem 11.4.4 (vector rank and submodularity)

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

$$\text{rank}(u) + \text{rank}(v) \geq \text{rank}(u \vee v) + \text{rank}(u \wedge v) \tag{11.38}$$

Vector rank rank(x) is submodular, proof

Proof of Theorem 11.4.4.

- Let a be a P-basis of $u \wedge v$, so 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 rank$(b) = b(E) =$ rank$(u \vee v)$, so b is a P-basis of $u \vee v$.
- 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, $a = b \wedge (u \wedge v) \ldots$
- ...and since $b \leq u \vee v$, we get

$$a + b = b + b \wedge u \wedge v = b \wedge u + b \wedge v \tag{11.39}$$

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 rank(x) is submodular, proof

...proof of Theorem 11.4.4.

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

$$\text{rank}(u \wedge v) + \text{rank}(u \vee v) = a(E) + b(E) \tag{11.40}$$

$$= (b \wedge u)(E) + (b \wedge v)(E) \tag{11.41}$$

$$\leq \text{rank}(u) + \text{rank}(v) \tag{11.42}$$

□

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 11.4.4 and the proof of Theorem **??** that the standard matroid rank function is submodular.
- Next, we prove Theorem 11.4.2, 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 11.4.2

Proof of Theorem 11.4.2.

- 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 $\alpha_{\text{max}} = \text{rank}(\infty \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_E)$.
- Hence, for any $x \in P$, $x(e) \leq \alpha_{\text{max}}, \forall e \in E$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) \tag{11.43}$$

- Then f is submodular since

$$\begin{aligned} f(A) + f(B) &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_B) && (11.44) \\ &\geq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \vee \alpha_{\text{max}} \mathbf{1}_B) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \wedge \alpha_{\text{max}} \mathbf{1}_B) && (11.45) \\ &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cup B}) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cap B}) && (11.46) \\ &= f(A \cup B) + f(A \cap B) && (11.47) \end{aligned}$$

...

Proof of Theorem 11.4.2

Proof of Theorem 11.4.2.

- 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:

$$P_f^+ = \left\{x \in \mathbb{R}_+^E : x(A) \leq f(A), \; \forall A \subseteq E\right\} \tag{11.48}$$

- Given an $x \in P$, then for any $A \subseteq E$, $x(A) \leq \max\{z(E) : z \in P, z \leq \alpha_{\text{max}} \mathbf{1}_A\} = \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) = f(A)$, 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 11.4.2

Proof of Theorem 11.4.2.

- 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

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{11.49}$$

- Choose w between y and x, so that

$$y \leq w \triangleq (y + x)/2 \leq x \tag{11.50}$$

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

Proof of Theorem 11.4.2

Proof of Theorem 11.4.2.

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

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \quad (11.51)$$

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

- Now, we have

$$y(N(y)) < w(N(y)) \le f(N(y)) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_{N(y)}) \quad (11.52)$$

the last inequality follows since $w \le x \in P_f^+$, and $y \le w$.

- 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 11.4.2

Proof of Theorem 11.4.2.

- 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, $P_f^+ = P$. □

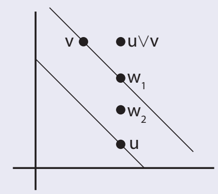
More on polymatroids

Theorem 11.4.5

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*

$$u < w \leq u \vee v \quad (11.53)$$



Corollary 11.4.6

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

Review

- The next slide comes from lecture 5.

Matroids by bases

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

Theorem 11.4.3 (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 11.4.7

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).