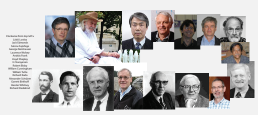
Submodular Functions, Optimization, and Applications to Machine Learning
— Spring Quarter, Lecture 10 —

http://www.ee.washington.edu/people/faculty/bilmes/classes/ee563_spring_2018/

Prof. Jeff Bilmes

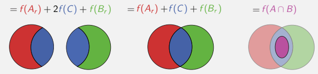
University of Washington, Seattle
Department of Electrical Engineering
http://melodi.ee.washington.edu/~bilmes

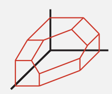
April 30th, 2018



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

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





Cumulative Outstanding Reading

- Read chapter 1 from Fujishige's book.
- Read chapter 2 from Fujishige's book.
- Read chapter 3 from Fujishige's book.
- Read chapter 4 from Fujishige's book.

Announcements, Assignments, and Reminders

- Next homework posted on canvas this evening (will include material from today's lecture).
- As always, if you have any questions about anything, please ask then via our discussion board (`https://canvas.uw.edu/courses/1216339/discussion_topics`). Can meet at odd hours via zoom (send message on canvas to schedule time to chat).

Class Road Map - EE563

- L1(3/26): Motivation, Applications, & Basic Definitions,
- L2(3/28): Machine Learning Apps (diversity, complexity, parameter, learning target, surrogate).
- L3(4/2): Info theory exs, more apps, definitions, graph/combinatorial examples
- L4(4/4): Graph and Combinatorial Examples, Matrix Rank, Examples and Properties, visualizations
- L5(4/9): More Examples/Properties/ Other Submodular Defs., Independence,
- L6(4/11): Matroids, Matroid Examples, Matroid Rank, Partition/Laminar Matroids
- L7(4/16): Laminar Matroids, System of Distinct Reps, Transversals, Transversal Matroid, Matroid Representation, Dual Matroids
- L8(4/18): Dual Matroids, Other Matroid Properties, Combinatorial Geometries, Matroids and Greedy.
- L9(4/23): Polyhedra, Matroid Polytopes, Matroids → Polymatroids
- L10(4/29): Matroids → Polymatroids, Polymatroids, Polymatroids and Greedy,
- L11(4/30):
- L12(5/2):
- L13(5/7):
- L14(5/9):
- L15(5/14):
- L16(5/16):
- L17(5/21):
- L18(5/23):
- L–(5/28): Memorial Day (holiday)
- L19(5/30):
- L21(6/4): Final Presentations maximization.

Last day of instruction, June 1st. Finals Week: June 2-8, 2018.

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.

Independence Polyhedra

- For each $I \in \mathcal{I}$ of a matroid $M = (E, \mathcal{I})$, we can form the incidence vector $\mathbf{1}_I \in \{0,1\}^E \subset [0,1]^E \subset \mathbb{R}_+^E$.
- Taking the convex hull, we get the independent set polytope, that is

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \subseteq [0,1]^E \tag{10.8}$$

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

$$P_r^+ \triangleq \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{10.9}$$

  Examples of P_r^+ are forthcoming.
- Since $\{\mathbf{1}_I : I \in \mathcal{I}\} \subseteq P_{\text{ind. set}} \subseteq P_r^+$, we have $\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \leq \max\{w^\intercal x : x \in P_r^+\}$
- 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.

Matroid Polyhedron in 2D

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{10.8}$$

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

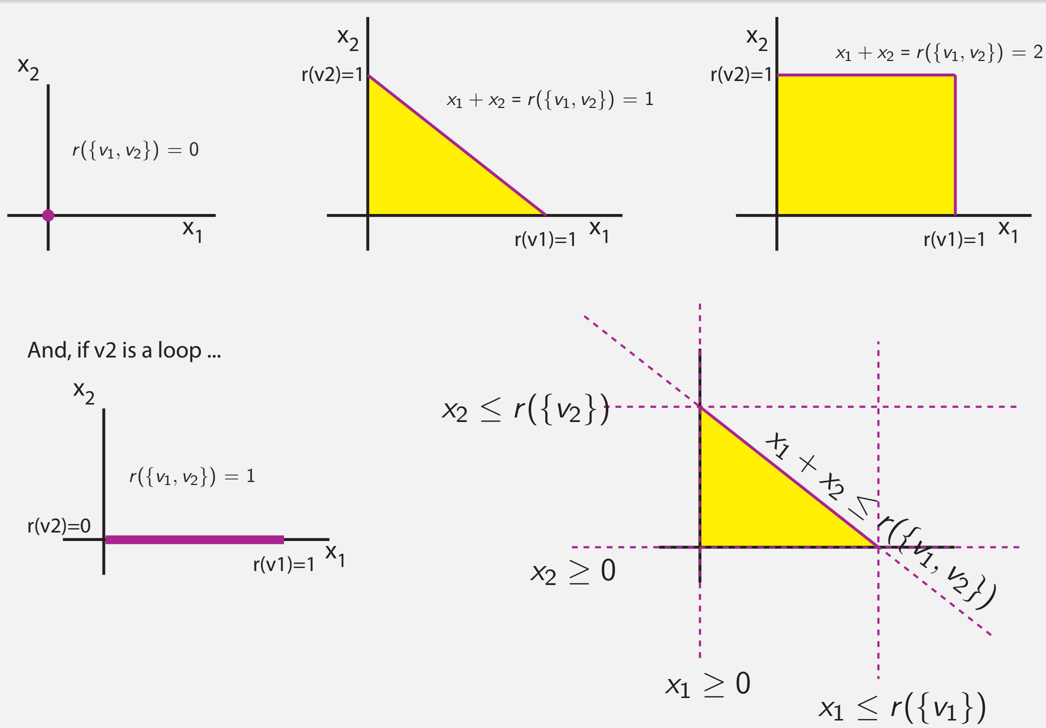
$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{10.9}$$
$$x_1 \leq r(\{v_1\}) \in \{0,1\} \tag{10.10}$$
$$x_2 \leq r(\{v_2\}) \in \{0,1\} \tag{10.11}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \in \{0,1,2\} \tag{10.12}$$

- Because r is submodular, we have

$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{10.13}$$

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

Matroid Polyhedron in 2D



Matroid Polyhedron in 3D

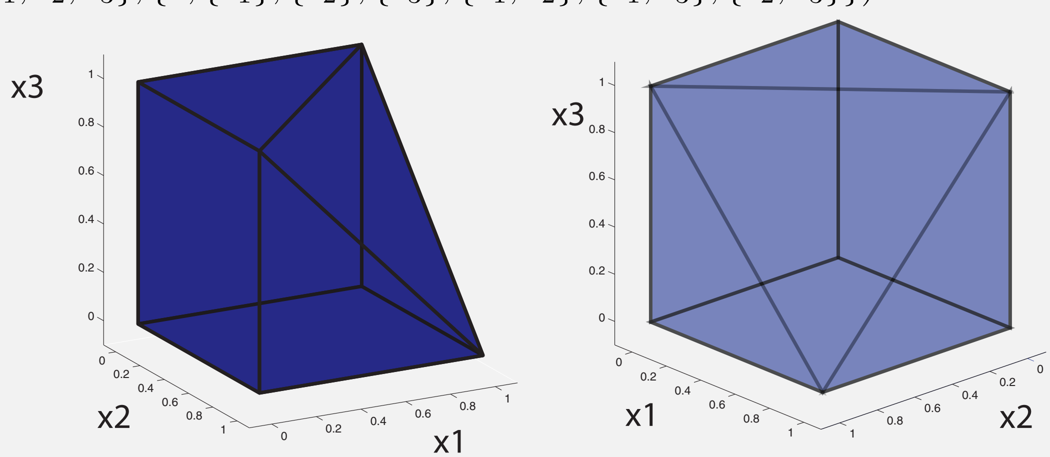
$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{10.8}$$

- Consider three dimensions, $E = \{1, 2, 3\}$. Get equations of the form:

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{10.9}$$
$$x_1 \leq r(\{v_1\}) \tag{10.10}$$
$$x_2 \leq r(\{v_2\}) \tag{10.11}$$
$$x_3 \leq r(\{v_3\}) \tag{10.12}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{10.13}$$
$$x_2 + x_3 \leq r(\{v_2, v_3\}) \tag{10.14}$$
$$x_1 + x_3 \leq r(\{v_1, v_3\}) \tag{10.15}$$
$$x_1 + x_2 + x_3 \leq r(\{v_1, v_2, v_3\}) \tag{10.16}$$

Matroid Polyhedron in 3D

Two view of P_r^+ associated with a matroid $(\{e_1, e_2, e_3\}, \{\emptyset, \{e_1\}, \{e_2\}, \{e_3\}, \{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}\})$.



Maximum weight independent set via greedy weighted rank

Theorem 10.2.5

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.8}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{10.9}$$

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

$$P_{\text{ind. set}} = \text{conv} \{\cup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\} \tag{10.11}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{10.12}$$

Theorem 10.2.5

$$P_r^+ = P_{\textit{ind. set}} \tag{10.13}$$

Greedy solves a linear programming problem

- So we can describe the independence polytope of a matroid using the set of inequalities (an exponential number of them).
- In fact, considering equations starting at Eq **??**, the LP problem with exponential number of constraints $\max \{w^\intercal x : x \in P_r^+\}$ is identical to the maximum weight independent set problem in a matroid, and since greedy solves the latter problem exactly, we have also proven:

Theorem 10.2.5

The LP problem $\max \{w^\intercal x : x \in P_r^+\}$ *can be solved exactly using the greedy algorithm.*

Note that this LP problem has an exponential number of constraints (since P_r^+ is described as the intersection of an exponential number of half spaces).

- This means that if LP problems have certain structure, they can be solved much easier than immediately implied by the equations.

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

Definition 10.2.6 (subvector)

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

Definition 10.2.7 (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 $M = (E, \mathcal{I})$.

$$\text{rank}(A) = \max\{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.23}$$

- 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:

$$\text{rank}(x) = \max(y(E) : y \leq x, y \in P) = \max_{y \in P}(x \wedge y)(E) \tag{10.24}$$

  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 10.3.1 (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 $y(E)$

- Condition 3 restated: That is for any two distinct maximal vectors $y^1, y^2 \in P$, with $y^1 \leq x$ & $y^2 \leq x$, with $y^1 \neq y^2$, we must have $y^1(E) = y^2(E)$.
- Condition 3 restated (again): For every vector $x \in \mathbb{R}_+^E$, every maximal independent (i.e., $\in P$) 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 10.3.1 (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 $y(E)$

- 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

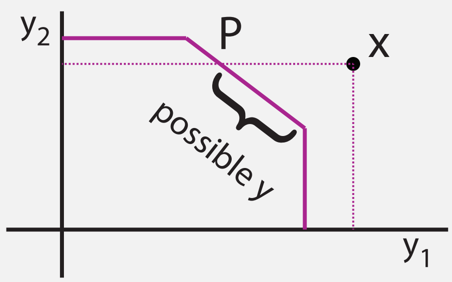
A Matroid is:

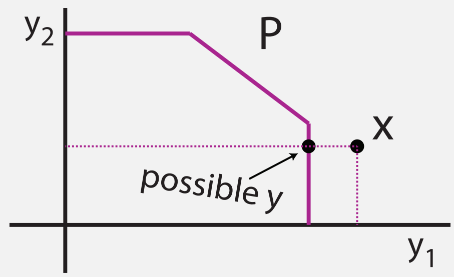
1. a set system $(E, \mathcal{I})$
2. empty-set containing $\emptyset \in \mathcal{I}$
3. down closed, $\emptyset \subseteq I' \subseteq I \in \mathcal{I} \Rightarrow I' \in \mathcal{I}$.
4. any maximal set I in $\mathcal{I}$, bounded by another set A, has the same matroid rank (any maximal independent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

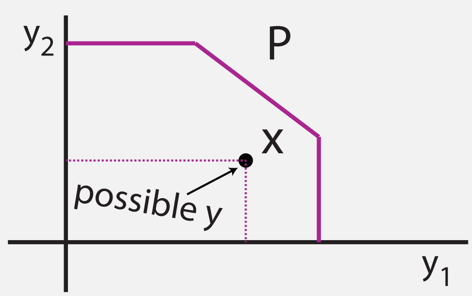




Left: $\exists$ multiple maximal $y \le x$ Right: $\exists$ only one maximal $y \le x$,

- Polymatroid condition here: $\forall$ maximal $y \in P$, with $y \le x$ (which here means $y_1 \le x_1$ and $y_2 \le x_2$), 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 $y \le x$. 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.

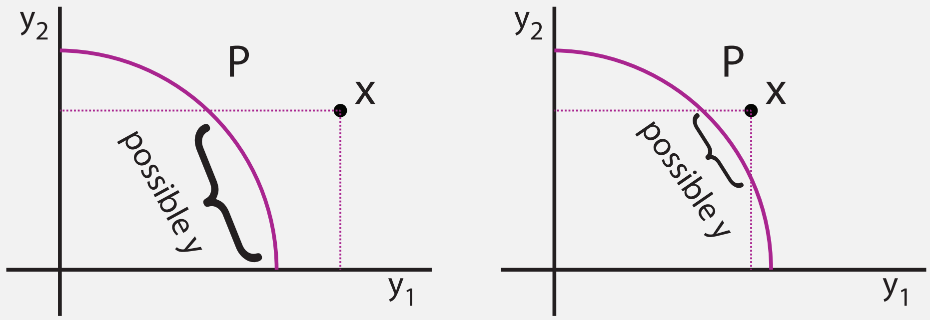
Polymatroidal polyhedron (or a "polymatroid")



$\exists$ only one maximal $y \le x$.

- 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 previous Lectures)

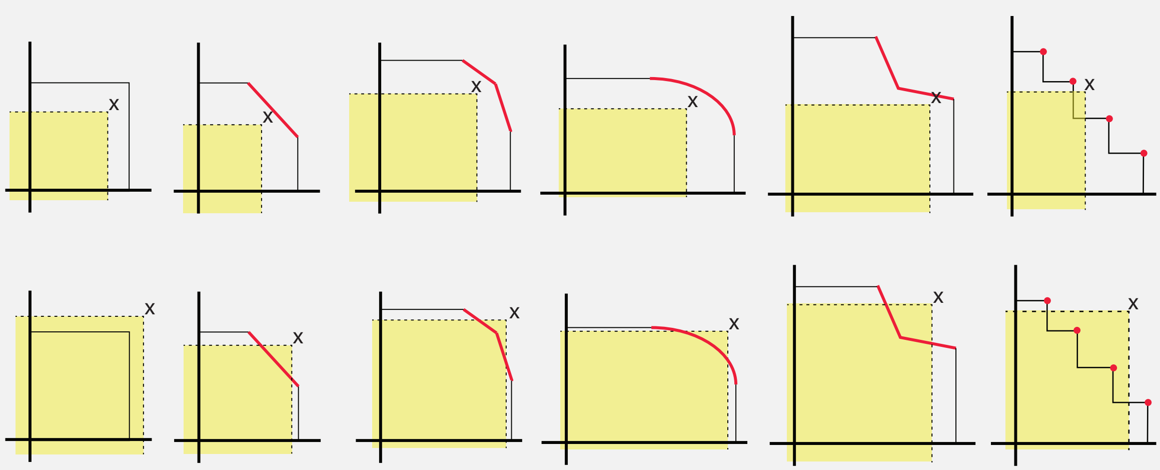
Polymatroid as well? no



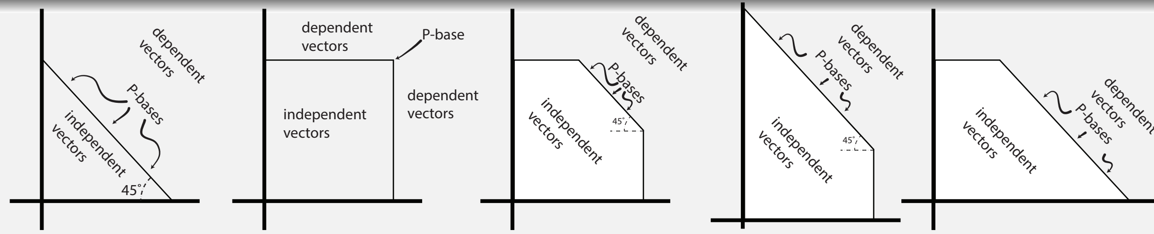
Left and right: $\exists$ multiple maximal $y \leq x$ as indicated.

- 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



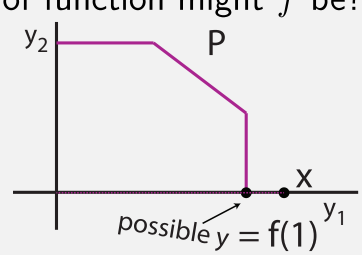
It appears that we have five possible forms of polymatroid in 2D, when neither of the elements $\{v_1, v_2\}$ are self-dependent.

1. On the left: full dependence between v_1 and v_2
2. Next: full independence between v_1 and v_2
3. Next: partial independence between v_1 and v_2
4. Right two: other forms of 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 small enough. What kind of function might f be?



Polymatroid function and its polyhedron.

Definition 10.3.2

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\} \quad (10.1)$$

$$= \{y \in \mathbb{R}^E : y \geq 0, y(A) \leq f(A) \text{ for all } A \subseteq E\} \quad (10.2)$$

Associated polyhedron with a polymatroid function

$$P_f^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\} \quad (10.3)$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \quad (10.4)$$

$$x_1 \leq f(\{v_1\}) \quad (10.5)$$

$$x_2 \leq f(\{v_2\}) \quad (10.6)$$

$$x_3 \leq f(\{v_3\}) \quad (10.7)$$

$$x_1 + x_2 \leq f(\{v_1, v_2\}) \quad (10.8)$$

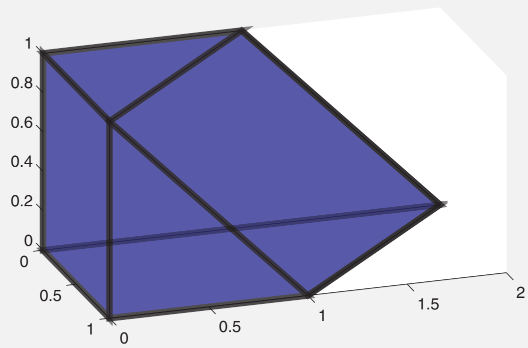
$$x_2 + x_3 \leq f(\{v_2, v_3\}) \quad (10.9)$$

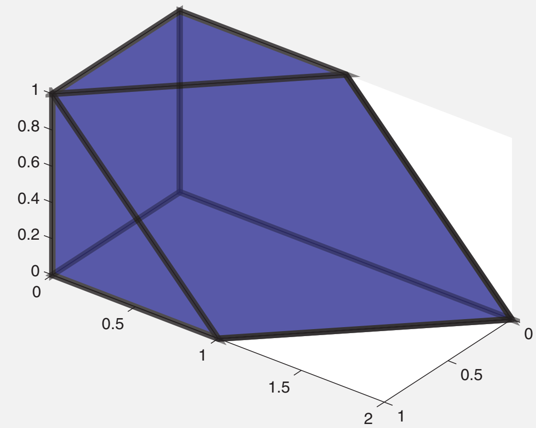
$$x_1 + x_3 \leq f(\{v_1, v_3\}) \quad (10.10)$$

$$x_1 + x_2 + x_3 \leq f(\{v_1, v_2, v_3\}) \quad (10.11)$$

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):

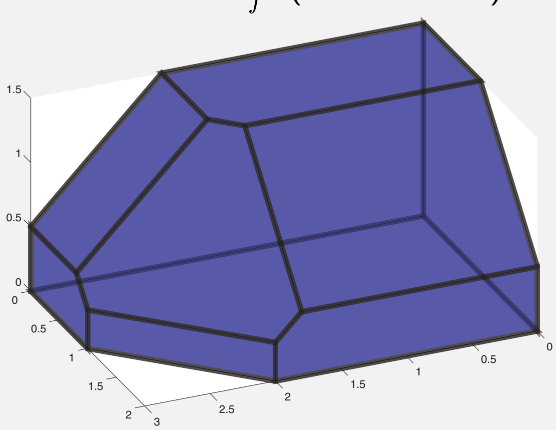


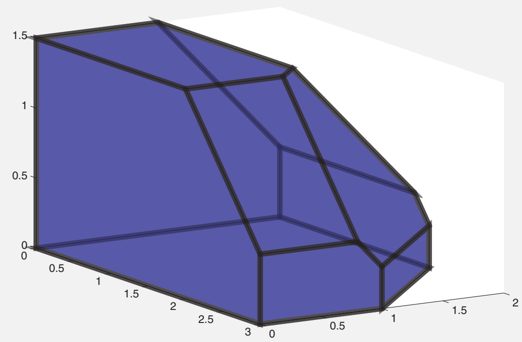


- which axis is which?

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):

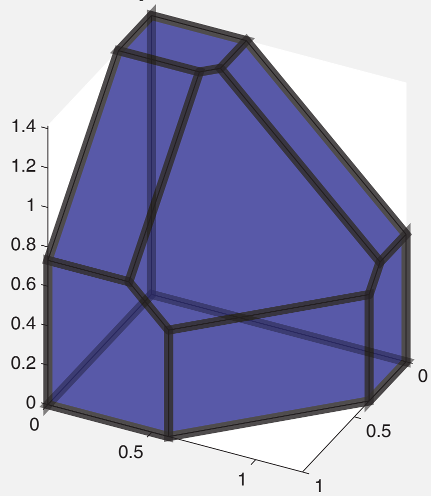


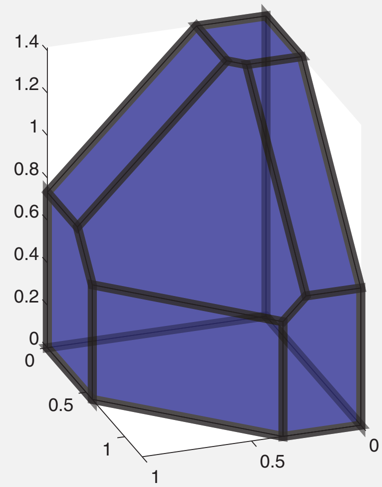


- which axis is which?

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):

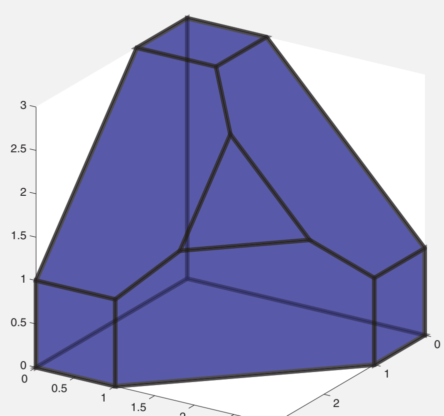


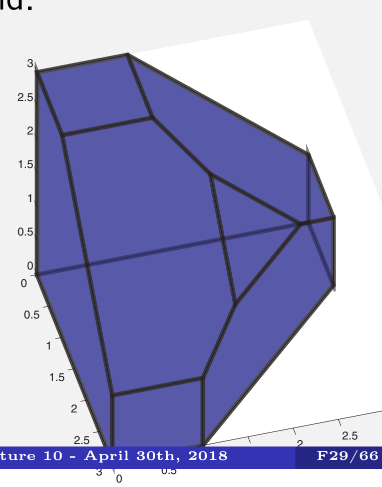


- which axis is which?

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \le f(A) \text{ for all } A \subseteq E\} \tag{10.12}$$

  - We also have the definition of a polymatroidal polytope P (compact subset, zero containing, down-monotone, and $\forall x$ any maximal independent subvector $y \le x$ 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 10.4.1

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

$$\begin{aligned} y^x(E) = \text{rank}(x) &= \max\left(y(E) : y \le x, y \in P_f^+\right) \\ &= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \end{aligned} \tag{10.13}$$

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:

$$\text{rank}\left(\frac{1}{\epsilon}\mathbf{1}_{E \setminus B}\right) = f(E \setminus B) = \max\left\{y(E \setminus B) : y \in P_f^+\right\} \tag{10.14}$$

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 $x \leq y$ 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 <u>only</u> 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,

$$
\begin{aligned}
&\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) && (10.15)\\
&\quad = x(E) + \min\left(f(E \setminus A) - x(E \setminus A) : A \subseteq E\right) && (10.16)\\
&\quad = x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) && (10.17)\\
&\quad = x(E) && (10.18)
\end{aligned}
$$

- 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, a self basis, unique P_f^+-base.

...

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

$$\min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{10.19}$$

$$= x(E) + \min\left(f(A) - x(A) : A \subseteq E\right) \tag{10.20}$$

$$= x(E) + \min\left(\sum_i f(a_i|A_{i-1}) - \sum_i x(a_i) : A \subseteq E\right) \tag{10.21}$$

$$= x(E) + \min\left(\sum_i \underbrace{\left(f(a_i|A_{i-1}) - x(a_i)\right)}_{\leq 0} : A \subseteq E\right) \tag{10.22}$$

$$= x(E) + f(E) - x(E) = f(E) = \max(y(E) : y \in P_f^+). \tag{10.23}$$

. . .

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

$$y^x(E) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{10.24}$$

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

$$y^x(E) = y^x(A) + y^x(E \setminus A) \tag{10.25}$$

$$\leq x(A) + f(E \setminus A). \tag{10.26}$$

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

- This ensures

$$\max\left(y(E) : y \leq x, y \in P_f^+\right) \leq \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{10.27}$$

- Given an A where equality in Eqn. (10.26) 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

$$f(B) + f(C) = y(B) + y(C) \tag{10.28}$$
$$= y(B \cap C) + y(B \cup C) \tag{10.29}$$
$$\leq f(B \cap C) + f(B \cup C) \tag{10.30}$$
$$\leq f(B) + f(C) \tag{10.31}$$

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

. . .

Matroids → Polymatroids | Polymatroids | Polymatroids and Greedy

A polymatroid function's polyhedron is a polymatroid.

. . . proof continued.

- Also, we define $\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\}$, so $y(\text{sat}(y)) = f(\text{sat}(y))$.
- 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 $e \in \text{sat}(y^x)$ (since if $e \in T \in \mathcal{D}(y^x)$, then $e \in \text{sat}(y^x)$).
- 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

$$y^x(E) = y^x(A) + y^x(E \setminus A) = x(A) + f(E \setminus A) \tag{10.32}$$

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

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 10.4.2

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 $y \in P_f^+$:

$$\mathcal{D}(y) \triangleq \{A : A \subseteq E,\ y(A) = f(A)\} \tag{10.33}$$

Theorem 10.4.3

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 ?? □

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

$$\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\} \tag{10.34}$$

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 $e \in E$

$$(x \vee y)(e) = \max(x(e), y(e)) \tag{10.35}$$

$$(x \wedge y)(e) = \min(x(e), y(e)) \tag{10.36}$$

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, namely one defined on the real lattice.

Theorem 10.4.4 (vector rank and submodularity)

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

$$\textit{rank}(u) + \textit{rank}(v) \ge \textit{rank}(u \vee v) + \textit{rank}(u \wedge v) \tag{10.37}$$

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

Proof of Theorem 10.4.4.

- 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

$$a + b = b \wedge u \wedge v + b = b \wedge u + b \wedge v \tag{10.38}$$

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 10.4.4.

- 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,

$$\text{rank}(u \wedge v) + \text{rank}(u \vee v) = a(E) + b(E) \tag{10.39}$$

$$= (b \wedge u)(E) + (b \wedge v)(E) \tag{10.40}$$

$$\leq \text{rank}(u) + \text{rank}(v) \tag{10.41}$$

$\square$

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 10.4.4 and the proof of Theorem 6.5.1 that the standard matroid rank function is submodular.
- Next, we prove Theorem 10.4.2, 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 10.4.2

Proof of Theorem 10.4.2.

- 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 $\alpha_{\text{max}} = \lim_{\alpha \to \infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_E)$.
- 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 $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) \tag{10.42}$$

- Then f is submodular since

$$\begin{aligned}
f(A) + f(B) &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_B) && (10.43)\\
&\geq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \vee \alpha_{\text{max}} \mathbf{1}_B) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_A \wedge \alpha_{\text{max}} \mathbf{1}_B) && (10.44)\\
&= \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cup B}) + \text{rank}(\alpha_{\text{max}} \mathbf{1}_{A \cap B}) && (10.45)\\
&= f(A \cup B) + f(A \cap B) && (10.46)
\end{aligned}$$

...

Proof of Theorem 10.4.2

Proof of Theorem 10.4.2.

- 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.
- Definition: for any $A \subseteq E$, define $x_A \in \mathbb{R}_+^E$ as

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \tag{10.47}$$

note this is an analogous definition to $\mathbf{1}_A$ but for a not necessarily unity vector x.

- Hence $x_A(A) = x(A)$ and $x_A(E \setminus A) = 0$.
- Consider the polytope P_f^+ defined as:

$$P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \le f(A), \forall A \subseteq E\} \tag{10.48}$$

...

Matroids → Polymatroids | Polymatroids | Polymatroids and Greedy

Proof of Theorem 10.4.2

Proof of Theorem 10.4.2.

- Given an $x \in P$, then for any $A \subseteq E$, $x_A \le \alpha_{\text{max}} \mathbf{1}_A$, and $x(A) \le \alpha_{\text{max}} |A|$.
- Therefore,

$$\begin{aligned} x(A) &\le \max \{z(A) : z \in P, z_A \le \alpha_{\text{max}} \mathbf{1}_A\} && (10.49) \\ &= \max \{z(A) : z \in P, z \le \alpha_{\text{max}} \mathbf{1}_A\} && (10.50) \\ &\le \max \{z(E) : z \in P, z \le \alpha_{\text{max}} \mathbf{1}_A\} && (10.51) \\ &= \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) && (10.52) \\ &= f(A) && (10.53) \end{aligned}$$

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 10.4.2

Proof of Theorem 10.4.2.

- 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

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{10.54}$$

- Choose w between y and x, so that

$$y \leq w \triangleq (y + x)/2 \leq x \tag{10.55}$$

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 10.4.2

Proof of Theorem 10.4.2.

- Now, we have

$$y(N(y)) < w(N(y)) \leq f(N(y)) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_{N(y)}) \tag{10.56}$$

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 10.4.2

Proof of Theorem 10.4.2.

- 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, $P_f^+ = P$.

□

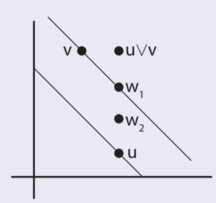
More on polymatroids

Theorem 10.4.5

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*

$$u < w \leq u \vee v \tag{10.57}$$



Corollary 10.4.6

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 10.4.3 (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 10.4.7

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).

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 ??: $\min\left(x(A) + f(E \setminus A) : A \subseteq E\right)$
- 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 10.4.8

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

$$\forall F \in \mathcal{F}, \quad \forall \emptyset \subset S \subseteq F, |S| \leq f(S) \tag{10.58}$$

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

Proof.

Exercise □

And its rank function is Exercise.

Matroid instance of Theorem ??

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

Corollary 10.4.9

We have that:

$$\max\{y(E) : y \in P_{\textit{ind. set}}(M), y \leq x\} = \min\{r_M(A) + x(E \setminus A) : A \subseteq E\} \tag{10.59}$$

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 10.5.1 (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 10.5.5

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.8}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{10.9}$$

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 independence 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. $E = (e_1, e_2, \dots, e_m)$ 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

$$w(e_1) \geq w(e_2) \geq \cdots \geq w(e_k) > 0 \geq w(e_{k+1}) \geq \cdots \geq w(e_m) \tag{10.60}$$

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

$$E_i \stackrel{\text{def}}{=} \{e_1, e_2, \dots e_i\} \tag{10.62}$$

  (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:

$$x(e_1) \stackrel{\text{def}}{=} f(E_1) = f(e_1) = f(e_1|E_0) = f(e_1|\emptyset) \tag{10.63}$$

$$x(e_i) \stackrel{\text{def}}{=} f(E_i) - f(E_{i-1}) = f(e_i|E_{i-1}) \text{ for } i = 2 \dots k \tag{10.64}$$

$$x(e_i) \stackrel{\text{def}}{=} 0 \text{ for } i = k+1 \dots m = |E| \tag{10.65}$$

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 10.5.1) 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 10.5.1

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:

$$\max(wx : x \in P_f^+) = \min\Big(\sum_{A\subseteq E} y_A f(A) : y \in \mathbb{R}_+^{2^E}, \sum_{A\subseteq E} y_A \mathbf{1}_A \geq w\Big) \tag{10.66}$$

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

$$y_{E_i} \leftarrow w(e_i) - w(e_{i+1}) \text{ for } i = 1 \dots (m-1), \tag{10.67}$$

$$y_E \leftarrow w(e_m), \text{ and} \tag{10.68}$$

$$y_A \leftarrow 0 \text{ otherwise} \tag{10.69}$$

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 $(e_1, e_2, \ldots, e_m)$.

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

$$f(A) = \sum_{i=1}^{k} f(a_i | a_{1:i-1}) \tag{10.70}$$

$$\geq \sum_{i=1}^{k} f(a_i | e_{1:e^{-1}(a_i)-1}) \tag{10.71}$$

$$= \sum_{a \in A} f(a | e_{1:e^{-1}(a)-1}) = x(A) \tag{10.72}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Next, y is also feasible for the dual constraints in Eq. 10.66 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:

$$wx = \sum_{i=1}^{m} w_i x_i \tag{10.73}$$

$$= \sum_{i=1}^{m} w_i \Big(f(E_i) - f(E_{i-1}) \Big) \tag{10.74}$$

$$= \sum_{i=1}^{m} w_i f(E_i) - \sum_{i=1}^{m-1} w_{i+1} f(E_i) \tag{10.75}$$

$$= w_m f(E_m) + \sum_{i=1}^{m-1} (w_i - w_{i+1}) f(E_i) \tag{10.76}$$

□

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 $w \in \mathbb{R}^E$