Submodular Functions, Optimization, and Applications to Machine Learning

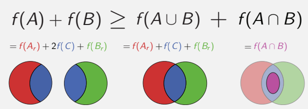
— Spring Quarter, Lecture 9 —

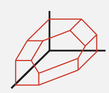
http://j.ee.washington.edu/~bilmes/classes/ee596b_spring_2014/

Prof. Jeff Bilmes

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

April 28th, 2014





Cumulative Outstanding Reading

- Read chapters 1 and 2, and sections 3.1-3.2 from Fujishige's book.
- Good references for today: Schrijver-2003, Oxley-1992/2011, Welsh-1973, Goemans-2010, Cunningham-1984, Edmonds-1969.

Announcements, Assignments, and Reminders

- Weekly Office Hours: Wednesdays, 5:00-5:50, or by skype or google hangout (email me).

Class Road Map - IT-I

- L1 (3/31): Motivation, Applications, & Basic Definitions
- L2: (4/2): Applications, Basic Definitions, Properties
- L3: More examples and properties (e.g., closure properties), and examples, spanning trees
- L4: proofs of equivalent definitions, independence, start matroids
- L5: matroids, basic definitions and examples
- L6: More on matroids, System of Distinct Reps, Transversals, Transversal Matroid, Matroid and representation
- L7: Dual Matroids, other matroid properties, Combinatorial Geometries
- L8: Combinatorial Geometries, matroids and greedy, Polyhedra, Matroid Polytopes,
- L9: From Matroid Polytopes to Polymatroids.
- L10: Polymatroids and Submodularity
- L11: More properties of polymatroids, SFM special cases
- L12:
- L13:
- L14:
- L15:
- L16:
- L17:
- L18:
- L19:
- L20:

Finals Week: June 9th-13th, 2014.

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

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

Theorem 9.2.2

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 $w(I)$.

Matroid Polyhedron in 2D

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{9.10}$$

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

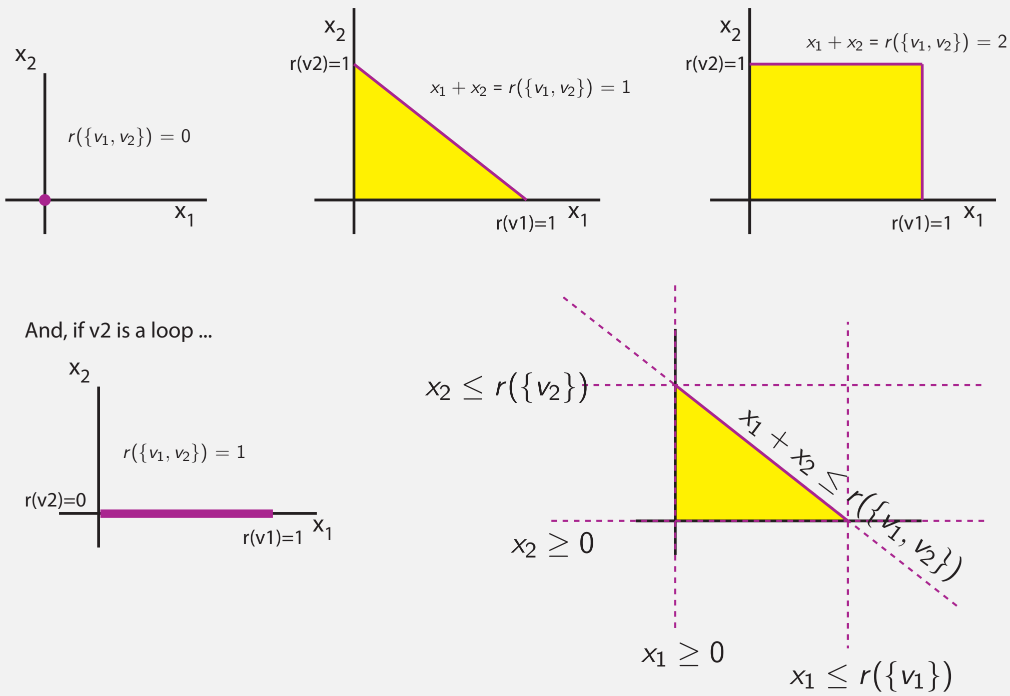
$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{9.11}$$
$$x_1 \leq r(\{v_1\}) \tag{9.12}$$
$$x_2 \leq r(\{v_2\}) \tag{9.13}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{9.14}$$

- Because r is submodular, we have

$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{9.15}$$

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

$$P_{\text{ind. set}} = \operatorname{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{9.10}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{9.11}$$

- 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
$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{9.10}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{9.11}$$
$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{9.12}$$
$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{9.13}$$
$$= r(A) \tag{9.14}$$
- 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
$$P_{\text{ind. set}} = \operatorname{conv}\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\}$$
$$\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{9.19}$$
- 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 9.2.6

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{9.19}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{9.20}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{array}{rlll} \text{maximize} & w^\intercal x & & \\ \text{subject to} & x_v \geq 0 & (v \in V) & \\ & x(U) \leq r(U) & (\forall U \subseteq V) & \end{array} \tag{9.1}$$

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

$$\begin{array}{rlll} \text{minimize} & \sum_{U \subseteq V} y_U r(U), & & \\ \text{subject to} & y_U \geq 0 & (\forall U \subseteq V) & \\ & \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w & & \end{array} \tag{9.2}$$

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

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 \qquad (v \in V) \\ & x(U) \leq r(U) \qquad (\forall U \subseteq V) \end{aligned} \tag{9.3}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r^+ \tag{9.4}$$

where, again, $P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\}$.
- Therefore, since $P_{\text{ind. set}} \subseteq P_r^+$, the above problem can only have a larger solution. I.e.,

$$\max w^\intercal x \text{ s.t. } x \in P_{\text{ind. set}} \leq \max w^\intercal x \text{ s.t. } x \in P_r^+. \tag{9.5}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max \{w(I) : I \in \mathcal{I}\} \leq \max \{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{9.6}$$

$$\leq \max \{w^\intercal x : x \in P_r^+\} \tag{9.7}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min \left\{ \sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \right\} \tag{9.8}$$

- Theorem 8.6.1 states that

$$\max \{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{9.9}$$

for the chain of U_i's and $\lambda_i \geq 0$ that satisfies $w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i}$ (i.e., the r.h.s. of Eq. 9.9 is feasible w.r.t. the dual LP).
- Therefore, we also have

$$\max \{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \geq \alpha_{\min} \tag{9.10}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{9.6}$$
$$= \max\{w^\intercal x : x \in P_r^+\} \tag{9.7}$$
$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U\subseteq V} y_U r(U) : y \geq 0, \sum_{U\subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{9.8}$$

- Therefore, all the inequalities above are equalities.
- And since $w \in \mathbb{R}_+^E$ is an arbitrary direction into the positive orthant, we see that $P_r^+ = P_{\text{ind. set}}$
- That is, we have just proven:

Theorem 9.3.1

$$P_r^+ = P_{\textit{ind. set}} \tag{9.11}$$

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

$$P_{\text{ind. set}} = \text{conv}\{\cup_{I\in\mathcal{I}}\{\mathbf{1}_I\}\} \tag{9.12}$$

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

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{9.13}$$

Theorem 9.3.2

$$P_r^+ = P_{\textit{ind. set}} \tag{9.14}$$

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 9.6, 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 9.3.3

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.

Base Polytope Equivalence

- Consider convex hull of indicator vectors just of the bases of a matroid, rather than all of the independent sets.
- Consider a polytope defined by the following constraints:

$$x \geq 0 \tag{9.15}$$

$$x(A) \leq r(A)\ \forall A \subseteq V \tag{9.16}$$

$$x(V) = r(V) \tag{9.17}$$

- Note the third requirement, $x(V) = r(V)$.
- By essentially the same argument as above (Exercise:), we can shown that the convex hull of the incidence vectors of the bases of a matroid is a polytope that can be described by Eq. 9.15- 9.17 above.
- What does this look like?

Spanning set polytope

- Recall, a set A is spanning in a matroid $M = (E, \mathcal{I})$ if $r(A) = r(E)$.
- Consider convex hull of incidence vectors of spanning sets of a matroid M, and call this $P_{\text{spanning}}(M)$.

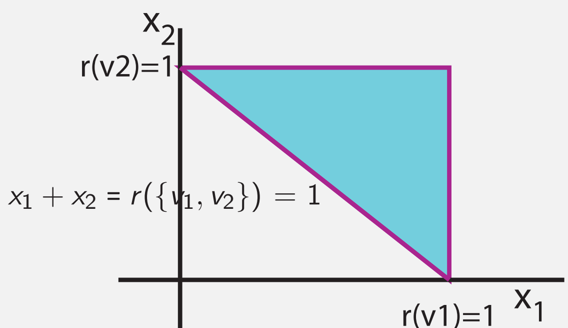
Theorem 9.3.4

The spanning set polytope is determined by the following equations:

$$0 \leq x_e \leq 1 \qquad \text{for } e \in E \tag{9.18}$$

$$x(A) \geq r(E) - r(E \setminus A) \qquad \text{for } A \subseteq E \tag{9.19}$$

- Example of spanning set polytope in 2D.



Spanning set polytope

Proof.

- Recall that any A is spanning in M iff $E \setminus A$ is independent in M^* (the dual matroid).
- For any $x \in \mathbb{R}^E$, we have that

$$x \in P_{\text{spanning}}(M) \Leftrightarrow \mathbf{1} - x \in P_{\text{ind. set}}(M^*) \tag{9.20}$$

as we show next ...

...

Spanning set polytope

... proof continued.

- This follows since if $x \in P_{\text{spanning}}(M)$, we can represent x as a convex combination:

$$x = \sum_i \lambda_i \mathbf{1}_{A_i} \tag{9.21}$$

where A_i is spanning in M.

- Consider

$$\mathbf{1} - x = \mathbf{1}_E - x = \mathbf{1}_E - \sum_i \lambda_i \mathbf{1}_{A_i} = \sum_i \lambda_i \mathbf{1}_{E \setminus A_i}, \tag{9.22}$$

which follows since $\sum_i \lambda_i \mathbf{1} = \mathbf{1}_E$, so $\mathbf{1} - x$ is a convex combination of independent sets in M^* and so $\mathbf{1} - x \in P_{\text{ind. set}}(M^*)$. ...

Spanning set polytope

... proof continued.

- which means, from the definition of $P_{\text{ind. set}}(M^*)$, that

$$\mathbf{1} - x \geq 0 \tag{9.23}$$

$$\mathbf{1}_A - x(A) = |A| - x(A) \leq r_{M^*}(A) \text{ for } A \subseteq E \tag{9.24}$$

And we know the dual rank function is

$$r_{M^*}(A) = |A| + r_M(E \setminus A) - r_M(E) \tag{9.25}$$

- giving

$$x(A) \geq r_M(E) - r_M(E \setminus A) \text{ for all } A \subseteq E \tag{9.26}$$

□

Matroids

where are we going with this?

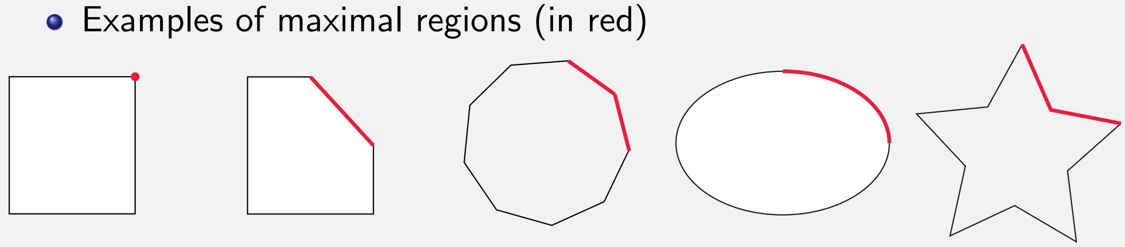
- We've been discussing results about matroids (independence polytope, etc.).
- By now, it is clear that matroid rank functions are special cases of submodular functions. We ultimately will be reviewing submodular function minimization procedures, but in some cases it it worth showing a result for a general submodular function first.
- Henceforth, we will skip between submodular functions and matroids, each lecture talking less about matroids specifically and taking more about submodular functions more generally ...

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 $e \in E$, we have that

$$x + \epsilon \mathbf{1}_e \notin P \tag{9.27}$$

- Examples of maximal regions (in red)

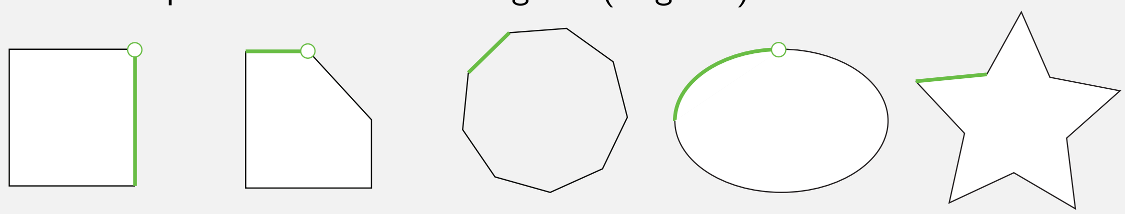


Maximal points in a set

- 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 $e \in E$, we have that

$$x + \epsilon \mathbf{1}_e \notin P \tag{9.27}$$

- Examples of non-maximal regions (in green)



Review

- The next slide comes from Lecture 5.

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).

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

Definition 9.4.1 (subvector)

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

Definition 9.4.2 (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 $M = (E, \mathcal{I})$.

$$\text{rank}(A) = \max \{|I| : I \subseteq A, I \in \mathcal{I}\} \quad (9.28)$$

- vector rank: Given a compact set $P \subseteq \mathcal{R}_+^E$, 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:

$$\text{rank}(x) = \max (y(E) : y \leq x, y \in P) \quad (9.29)$$

  where $y \leq x$ is componentwise inequality ($y_i \leq x_i, \forall i$).
- 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).
- In general, this might be hard to compute and/or have ill-defined properties. We next look at an object that restrains and cultivates this form of rank.

Polymatroidal polyhedron (or a "polymatroid")

Definition 9.4.3 (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 $y(E)$

- 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 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 9.4.3 (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:

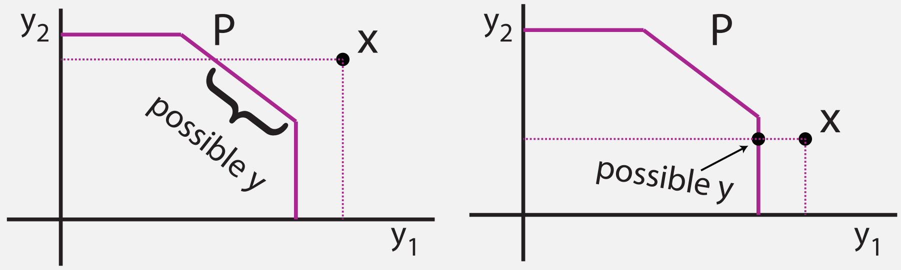
1. a set system $(E, \mathcal{I})$
2. empty-set containing $\emptyset \in \mathcal{I}$
3. down closed, $\emptyset \subseteq I' \subseteq I \in \mathcal{I} \Rightarrow I' \in \mathcal{I}$.
4. any maximal set I in $\mathcal{I}$, bounded by another set A, has the same matroid rank (any maximal independent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

.

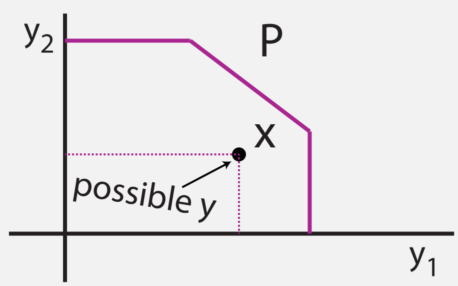
Polymatroidal polyhedron (or a "polymatroid")



Left: $\exists$ multiple maximal $y \leq x$ Right: $\exists$ 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.

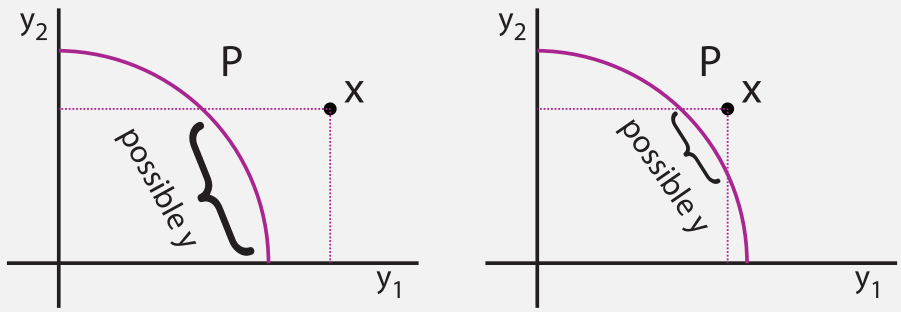
Polymatroidal polyhedron (or a "polymatroid")



$\exists$ 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)

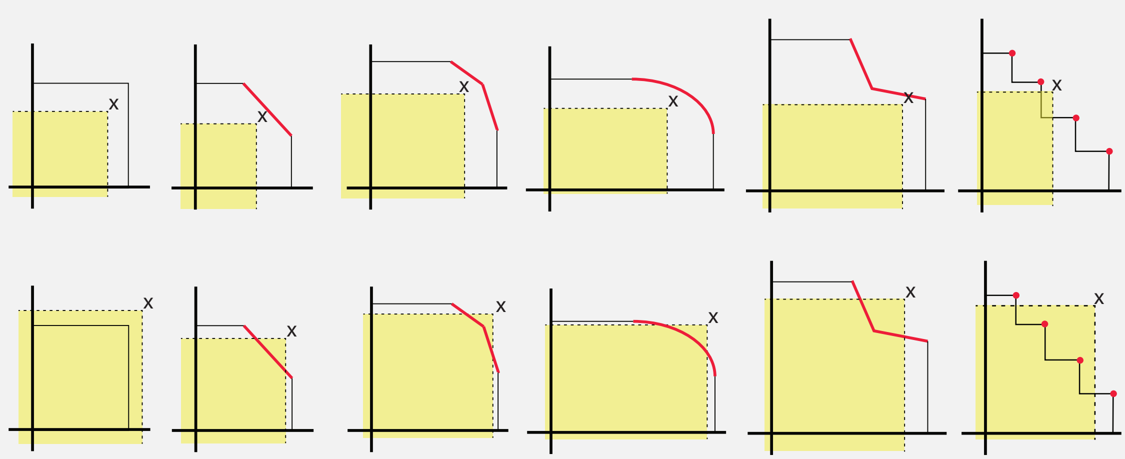
Polymatroid as well? no



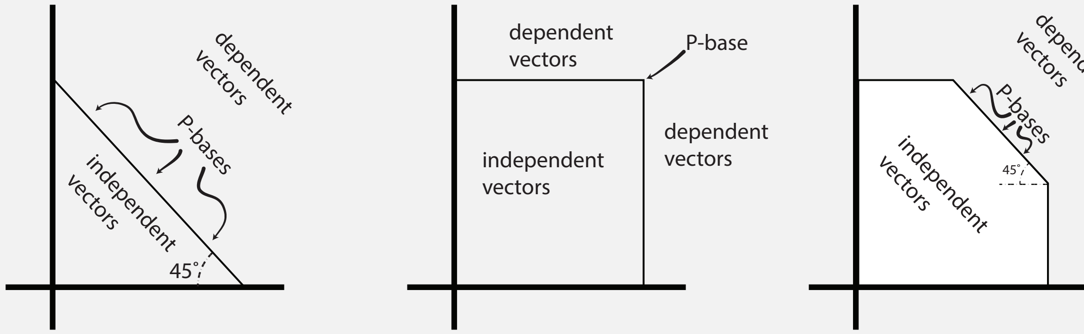
Left and right: $\exists$ multiple maximal $y \leq x$ as indicated.

- 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



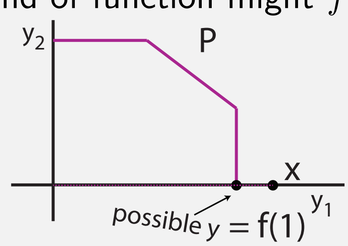
It appears that we have three possible forms of polymatroid in 2D, when neither of the elements $\{v_1, v_2\}$ are self-dependent.

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 9.4.4

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

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \quad (9.30)$$

$$= \left\{y \in \mathbb{R}^E : y \geq 0, y(A) \leq f(A) \text{ for all } A \subseteq E\right\} \quad (9.31)$$

Associated polyhedron with a polymatroid function

$$P_f^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\right\} \quad (9.32)$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \quad (9.33)$$

$$x_1 \leq f(\{v_1\}) \quad (9.34)$$

$$x_2 \leq f(\{v_2\}) \quad (9.35)$$

$$x_3 \leq f(\{v_3\}) \quad (9.36)$$

$$x_1 + x_2 \leq f(\{v_1, v_2\}) \quad (9.37)$$

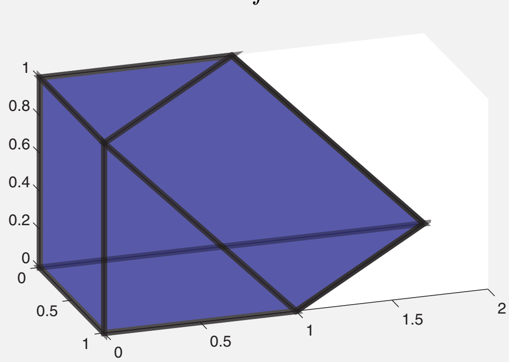
$$x_2 + x_3 \leq f(\{v_2, v_3\}) \quad (9.38)$$

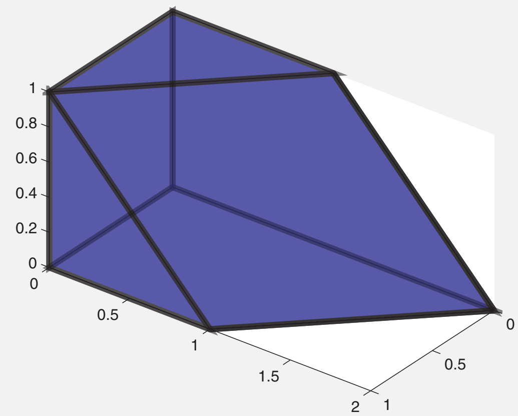
$$x_1 + x_3 \leq f(\{v_1, v_3\}) \quad (9.39)$$

$$x_1 + x_2 + x_3 \leq f(\{v_1, v_2, v_3\}) \quad (9.40)$$

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):

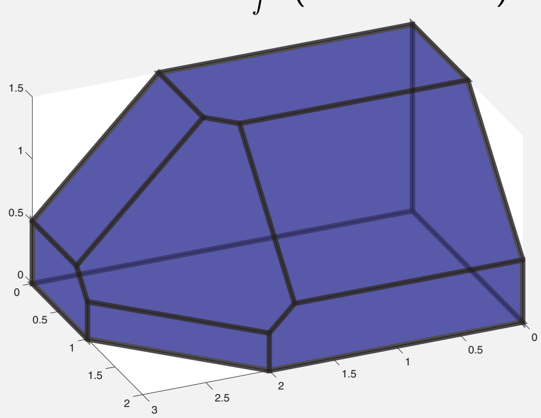


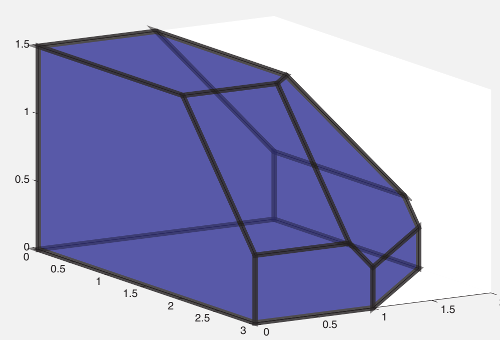


- which axis is which?

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):

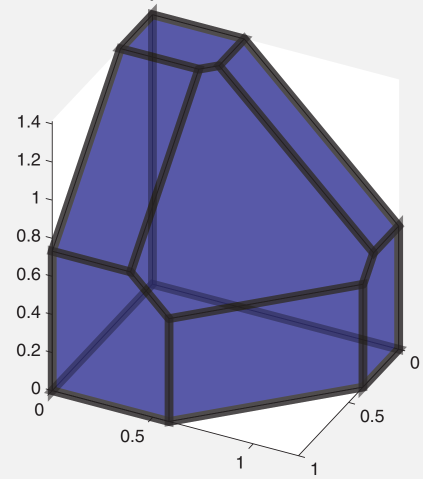


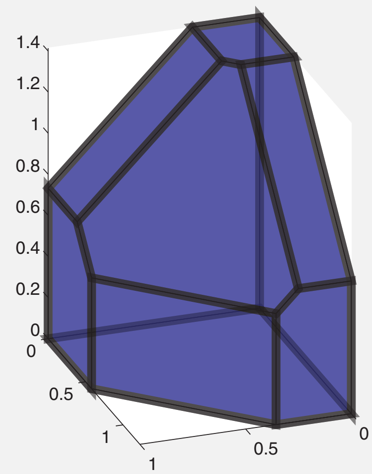


- which axis is which?

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):

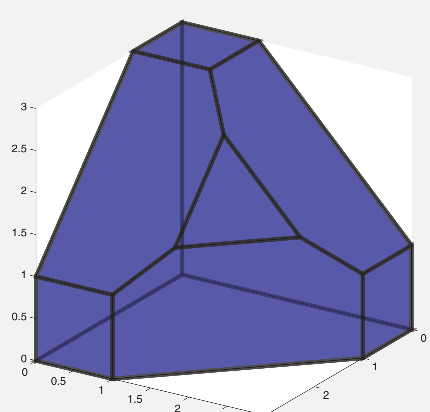


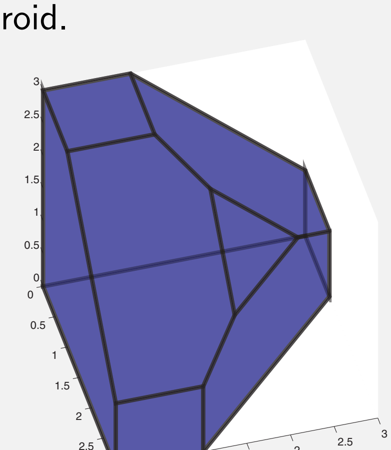


- which axis is which?

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

$$P_f^+ = \left\{y \in \mathbb{R}_+^E : y(A) \le f(A) \text{ for all } A \subseteq E\right\} \quad (9.41)$$

  - 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 9.4.5

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

$$y^x(E) = \text{rank}(x) = \max\left(y(E) : y \le x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \quad (9.42)$$

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

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

$$f(B) = \max\left\{y(B) : y \in P_f^+\right\} \quad (9.43)$$

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