Submodular Functions, Optimization, and Applications to Machine Learning

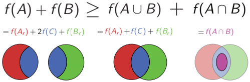
— Spring Quarter, Lecture 10 —

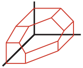
http://www.ee.washington.edu/people/faculty/bilmes/classes/ee596b_spring_2016/

Prof. Jeff Bilmes

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

May 2nd, 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 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, Polymatroid
- L11(5/2):
- 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

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

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

Theorem 10.2.7

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

Review from Lecture 6

- The next slide is from Lecture 6.

Matroids by bases

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

Theorem 10.3.3 (Matroid (by bases))

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

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

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

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume $(E, \mathcal{I})$ is a matroid and $w : E \to \mathcal{R}_+$ is given.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume $(E, \mathcal{I})$ is a matroid and $w : E \to \mathcal{R}_+$ is given.
- Let $A = (a_1, a_2, \dots, a_r)$ be the solution returned by greedy, where $r = r(M)$ the rank of the matroid, and we order the elements as they were chosen (so $w(a_1) \geq w(a_2) \geq \dots \geq w(a_r)$).

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume $(E, \mathcal{I})$ is a matroid and $w : E \to \mathcal{R}_+$ is given.
- Let $A = (a_1, a_2, \ldots, a_r)$ be the solution returned by greedy, where $r = r(M)$ the rank of the matroid, and we order the elements as they were chosen (so $w(a_1) \geq w(a_2) \geq \cdots \geq w(a_r)$).
- A is a base of M, and let $B = (b_1, \ldots, b_r)$ be <u>any</u> another base of M with elements also ordered decreasing by weight, so $w(b_1) \geq w(b_2) \geq \cdots \geq w(b_r)$.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume $(E, \mathcal{I})$ is a matroid and $w : E \to \mathcal{R}_+$ is given.
- Let $A = (a_1, a_2, \dots, a_r)$ be the solution returned by greedy, where $r = r(M)$ the rank of the matroid, and we order the elements as they were chosen (so $w(a_1) \geq w(a_2) \geq \dots \geq w(a_r)$).
- A is a base of M, and let $B = (b_1, \dots, b_r)$ be <u>any</u> another base of M with elements also ordered decreasing by weight, so $w(b_1) \geq w(b_2) \geq \dots \geq w(b_r)$.
- We next show that not only is $w(A) \geq w(B)$ but that $w(a_i) \geq w(b_i)$ for all i.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume otherwise, and let k be the first (smallest) integer such that $w(a_k) < w(b_k)$. Hence $w(a_j) \geq w(b_j)$ for $j < k$.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume otherwise, and let k be the first (smallest) integer such that $w(a_k) < w(b_k)$. Hence $w(a_j) \geq w(b_j)$ for $j < k$.
- Define independent sets $A_{k-1} = \{a_1, \dots, a_{k-1}\}$ and $B_k = \{b_1, \dots, b_k\}$.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume otherwise, and let k be the first (smallest) integer such that $w(a_k) < w(b_k)$. Hence $w(a_j) \geq w(b_j)$ for $j < k$.
- Define independent sets $A_{k-1} = \{a_1, \dots, a_{k-1}\}$ and $B_k = \{b_1, \dots, b_k\}$.
- Since $|A_{k-1}| < |B_k|$, there exists a $b_i \in B_k \setminus A_{k-1}$ where $A_{k-1} \cup \{b_i\} \in \mathcal{I}$ for some $1 \leq i \leq k$.

...

Matroid and the greedy algorithm

proof of Theorem 9.6.1.

- Assume otherwise, and let k be the first (smallest) integer such that $w(a_k) < w(b_k)$. Hence $w(a_j) \geq w(b_j)$ for $j < k$.
- Define independent sets $A_{k-1} = \{a_1, \dots, a_{k-1}\}$ and $B_k = \{b_1, \dots, b_k\}$.
- Since $|A_{k-1}| < |B_k|$, there exists a $b_i \in B_k \setminus A_{k-1}$ where $A_{k-1} \cup \{b_i\} \in \mathcal{I}$ for some $1 \leq i \leq k$.
- But $w(b_i) \geq w(b_k) > w(a_k)$, and so the greedy algorithm would have chosen b_i rather than a_k, contradicting what greedy does.

□

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Given an independence system $(E, \mathcal{I})$, suppose the greedy algorithm leads to an independent set of max weight for every non-negative weight function. We'll show $(E, \mathcal{I})$ is a matroid.

. . .

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Given an independence system $(E, \mathcal{I})$, suppose the greedy algorithm leads to an independent set of max weight for every non-negative weight function. We'll show $(E, \mathcal{I})$ is a matroid.
- Emptyset containing and down monotonicity already holds (since we've started with an independence system).

...

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Given an independence system $(E, \mathcal{I})$, suppose the greedy algorithm leads to an independent set of max weight for every non-negative weight function. We'll show $(E, \mathcal{I})$ is a matroid.
- Emptyset containing and down monotonicity already holds (since we've started with an independence system).
- Let $I, J \in \mathcal{I}$ with $|I| < |J|$. Suppose to the contrary, that $I \cup \{z\} \notin \mathcal{I}$ for all $z \in J \setminus I$.

...

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Given an independence system $(E, \mathcal{I})$, suppose the greedy algorithm leads to an independent set of max weight for every non-negative weight function. We'll show $(E, \mathcal{I})$ is a matroid.
- Emptyset containing and down monotonicity already holds (since we've started with an independence system).
- Let $I, J \in \mathcal{I}$ with $|I| < |J|$. Suppose to the contrary, that $I \cup \{z\} \notin \mathcal{I}$ for all $z \in J \setminus I$.
- Define the following modular weight function w on E, and define $k = |I|$.

$$w(v) = \begin{cases} k+2 & \text{if } v \in I, \\ k+1 & \text{if } v \in J \setminus I, \\ 0 & \text{if } v \in E \setminus (I \cup J) \end{cases} \tag{10.1}$$

...

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Now greedy will, after k iterations, recover I, but it cannot choose any element in $J \setminus I$ by assumption. Thus, greedy chooses a set of weight $k(k+2)$.

. . .

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Now greedy will, after k iterations, recover I, but it cannot choose any element in $J \setminus I$ by assumption. Thus, greedy chooses a set of weight $k(k+2)$.
- On the other hand, J has weight

$$w(J) \geq |J|(k+1) \geq (k+1)(k+1) > k(k+2) \tag{10.2}$$

so J has strictly larger weight but is still independent, contradicting greedy's optimality.

. . .

Matroid and the greedy algorithm

converse proof of Theorem 9.6.1.

- Now greedy will, after k iterations, recover I, but it cannot choose any element in $J \setminus I$ by assumption. Thus, greedy chooses a set of weight $k(k+2)$.
- On the other hand, J has weight

$$w(J) \geq |J|(k+1) \geq (k+1)(k+1) > k(k+2) \tag{10.2}$$

  so J has strictly larger weight but is still independent, contradicting greedy's optimality.
- Therefore, there must be a $z \in J \setminus I$ such that $I \cup \{z\} \in \mathcal{I}$, and since I and J are arbitrary, $(E, \mathcal{I})$ must be a matroid.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.
- If we don't want elements with weight 0, we can stop once (and if) the weight hits zero, thus giving us a maximum weight independent set.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.
- If we don't want elements with weight 0, we can stop once (and if) the weight hits zero, thus giving us a maximum weight independent set.
- We don't need non-negativity, we can use any $w \in \mathbb{R}^E$ and keep going until we have a base.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.
- If we don't want elements with weight 0, we can stop once (and if) the weight hits zero, thus giving us a maximum weight independent set.
- We don't need non-negativity, we can use any $w \in \mathbb{R}^E$ and keep going until we have a base.
- If we stop at a negative value, we'll once again get a maximum weight independent set.

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.
- If we don't want elements with weight 0, we can stop once (and if) the weight hits zero, thus giving us a maximum weight independent set.
- We don't need non-negativity, we can use any $w \in \mathbb{R}^E$ and keep going until we have a base.
- If we stop at a negative value, we'll once again get a maximum weight independent set.
- Exercise: what if we keep going until a base even if we encounter negative values?

Matroid and greedy

- As given, the theorem asked for a modular function $w \in \mathbb{R}_+^E$.
- This will not only return an independent set, but it will return a base if we keep going even if the weights are 0.
- If we don't want elements with weight 0, we can stop once (and if) the weight hits zero, thus giving us a maximum weight independent set.
- We don't need non-negativity, we can use any $w \in \mathbb{R}^E$ and keep going until we have a base.
- If we stop at a negative value, we'll once again get a maximum weight independent set.
- Exercise: what if we keep going until a base even if we encounter negative values?
- We can instead do as small as possible thus giving us a minimum weight independent set/base.

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.

Convex Polyhedra

- Convex polyhedra a rich topic, we will only draw what we need.

Convex Polyhedra

- Convex polyhedra a rich topic, we will only draw what we need.

Definition 10.4.1

A subset $P \subseteq \mathbb{R}^E$ is a polyhedron if there exists an $m \times n$ matrix A and vector $b \in \mathbb{R}^m$ (for some $m \geq 0$) such that

$$P = \left\{x \in \mathbb{R}^E : Ax \leq b\right\} \tag{10.3}$$

Convex Polyhedra

- Convex polyhedra a rich topic, we will only draw what we need.

Definition 10.4.1

A subset $P \subseteq \mathbb{R}^E$ is a polyhedron if there exists an $m \times n$ matrix A and vector $b \in \mathbb{R}^m$ (for some $m \geq 0$) such that

$$P = \{x \in \mathbb{R}^E : Ax \leq b\} \tag{10.3}$$

- Thus, P is intersection of finitely many affine halfspaces, which are of the form $a_i x \leq b_i$ where a_i is a row vector and b_i a real scalar.

Convex Polytope

- A polytope is defined as follows

Convex Polytope

- A polytope is defined as follows

Definition 10.4.2

A subset $P \subseteq \mathbb{R}^E$ is a polytope if it is the convex hull of finitely many vectors in $\mathcal{R}^E$. That is, if $\exists$, $x_1, x_2, \ldots, x_k \in \mathcal{R}^E$ such that for all $x \in P$, there exits $\{\lambda_i\}$ with $\sum_i \lambda_i = 1$ and $\lambda_i \geq 0 \; \forall i$ with $x = \sum_i \lambda_i x_i$.

Convex Polytope

- A polytope is defined as follows

Definition 10.4.2

A subset $P \subseteq \mathbb{R}^E$ is a polytope if it is the convex hull of finitely many vectors in $\mathcal{R}^E$. That is, if $\exists$, $x_1, x_2, \ldots, x_k \in \mathcal{R}^E$ such that for all $x \in P$, there exits $\{\lambda_i\}$ with $\sum_i \lambda_i = 1$ and $\lambda_i \geq 0 \; \forall i$ with $x = \sum_i \lambda_i x_i$.

- We define the convex hull operator as follows:

$$\text{conv}(x_1, x_2, \ldots, x_k) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^k \lambda_i x_i : \forall i, \; \lambda_i \geq 0, \text{ and } \sum_i \lambda_i = 1 \right\} \tag{10.4}$$

Convex Polytope - key representation theorem

- A polytope can be defined in a number of ways, two of which include

Theorem 10.4.3

A subset $P \subseteq \mathbb{R}^E$ is a polytope iff it can be described in either of the following (equivalent) ways:

- *P is the convex hull of a finite set of points.*
- *If it is a bounded intersection of halfspaces, that is there exits matrix A and vector b such that*

$$P = \{x : Ax \leq b\} \tag{10.5}$$

Convex Polytope - key representation theorem

- A polytope can be defined in a number of ways, two of which include

Theorem 10.4.3

A subset $P \subseteq \mathbb{R}^E$ is a polytope iff it can be described in either of the following (equivalent) ways:

- *P is the convex hull of a finite set of points.*
- *If it is a bounded intersection of halfspaces, that is there exits matrix A and vector b such that*

$$P = \{x : Ax \leq b\} \tag{10.5}$$

- This result follows directly from results proven by Fourier, Motzkin, Farkas, and Carátheodory.

Linear Programming

Theorem 10.4.4 (weak duality)

Let A be a matrix and b and c vectors, then

$$\max \{c^\intercal x | Ax \leq b\} \leq \min \{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{10.6}$$

Linear Programming

Theorem 10.4.4 (weak duality)

Let A be a matrix and b and c vectors, then

$$\max\{c^\intercal x | Ax \leq b\} \leq \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{10.6}$$

Theorem 10.4.5 (strong duality)

Let A be a matrix and b and c vectors, then

$$\max\{c^\intercal x | Ax \leq b\} = \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{10.7}$$

Linear Programming duality forms

There are many ways to construct the dual. For example,

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \quad (10.8)$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \quad (10.9)$$

$$\min\{c^\intercal x | x \geq 0, Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A \leq c^\intercal\} \quad (10.10)$$

$$\min\{c^\intercal x | Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A = c^\intercal\} \quad (10.11)$$

Linear Programming duality forms

How to form the dual in general? We quote V. Vazirani (2001)

Linear Programming duality forms

How to form the dual in general? We quote V. Vazirani (2001)

> *Intuitively, why is [one set of equations] the dual of [another quite different set of equations]? In our experience, this is not the right question to be asked. As stated in Section 12.1, there is a purely mechanical procedure for obtaining the dual of a linear program. Once the dual is obtained, one can devise intuitive, and possibly physical meaningful, ways of thinking about it. Using this mechanical procedure, one can obtain the dual of a complex linear program in a fairly straightforward manner. Indeed, the LP-duality-based approach derives its wide applicability from this fact.*

Also see the text "Convex Optimization" by Boyd and Vandenberghe, chapter 5, for a great discussion on duality and easy mechanical ways to construct it.

Vector, modular, incidence

- Recall, any vector $x \in \mathbb{R}^E$ can be seen as a normalized modular function, as for any $A \subseteq E$, we have

$$x(A) = \sum_{a \in A} x_a \tag{10.12}$$

Vector, modular, incidence

- Recall, any vector $x \in \mathbb{R}^E$ can be seen as a normalized modular function, as for any $A \subseteq E$, we have

$$x(A) = \sum_{a \in A} x_a \tag{10.12}$$

- Given an $A \subseteq E$, define the the incidence vector $\mathbf{1}_A \in \{0,1\}^E$ on the unit hypercube as follows:

$$\mathbf{1}_A \stackrel{\text{def}}{=} \left\{ x \in \{0,1\}^E : x_i = 1 \text{ iff } i \in A \right\} \tag{10.13}$$

equivalently,

$$\mathbf{1}_A(j) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{if } j \notin A \end{cases} \tag{10.14}$$

Review from Lecture 6

The next slide is review from lecture 6.

Matroid

Slight modification (non unit increment) that is equivalent.

Definition 10.5.3 (Matroid-II)

A set system $(E, \mathcal{I})$ is a Matroid if

(I1') $\emptyset \in \mathcal{I}$

(I2') $\forall I \in \mathcal{I}, J \subset I \Rightarrow J \in \mathcal{I}$ (down-closed or subclusive)

(I3') $\forall I, J \in \mathcal{I}$, with $|I| > |J|$, then there exists $x \in I \setminus J$ such that $J \cup \{x\} \in \mathcal{I}$

Note (I1)=(I1'), (I2)=(I2'), and we get (I3)≡(I3') using induction.

Independence Polyhedra

- For each $I \in \mathcal{I}$ of a matroid $M = (E, \mathcal{I})$, we can form the incidence vector $\mathbf{1}_I$.

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

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{10.15}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{ \bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\} \right\} \tag{10.15}$$

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

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

$$P_{\text{ind. set}} = \text{conv}\left\{ \bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\} \right\} \tag{10.15}$$

- 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:

$$P_r^+ = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{10.16}$$

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

$$P_{\text{ind. set}} = \operatorname{conv}\left\{ \bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\} \right\} \tag{10.15}$$

- 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:

$$P_r^+ = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{10.16}$$

- 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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

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

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{10.18}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.

- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{10.18}$$

$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{10.19}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{10.18}$$

$$\leq \sum_i \lambda_i \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{10.19}$$

$$= \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.20}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{10.18}$$

$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{10.19}$$

$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.20}$$

$$= r(A) \tag{10.21}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{10.17}$$

for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.

- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{10.18}$$

$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{10.19}$$

$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.20}$$

$$= r(A) \tag{10.21}$$

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

Matroid Polyhedron in 2D

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{10.22}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{10.23}$$

$$x_1 \leq r(\{v_1\}) \tag{10.24}$$

$$x_2 \leq r(\{v_2\}) \tag{10.25}$$

$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{10.26}$$

Matroid Polyhedron in 2D

$$P_r^+ = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{10.22}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{10.23}$$

$$x_1 \leq r(\{v_1\}) \tag{10.24}$$

$$x_2 \leq r(\{v_2\}) \tag{10.25}$$

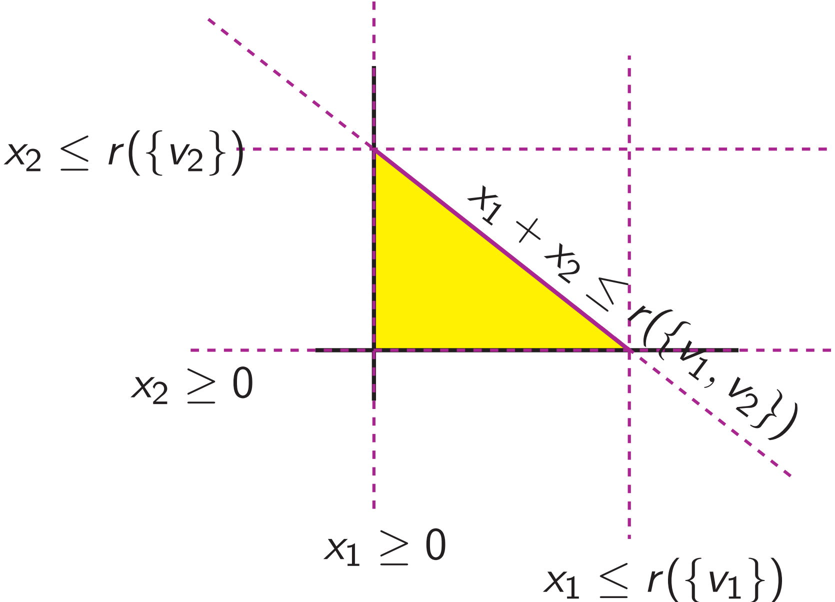
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{10.26}$$

- Because r is submodular, we have

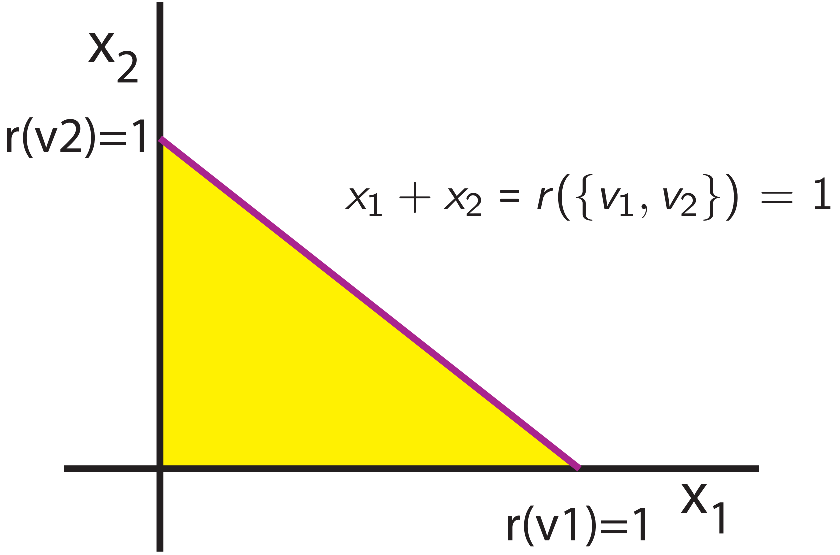
$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{10.27}$$

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

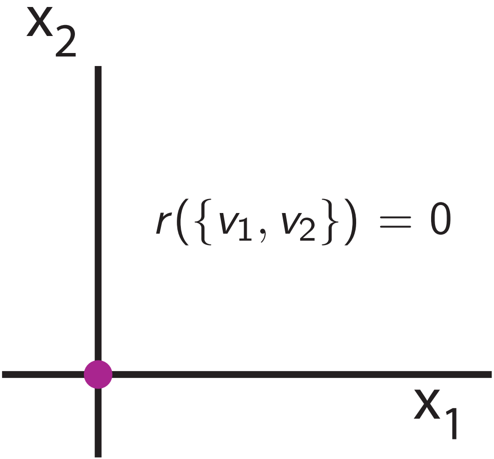
Matroid Polyhedron in 2D



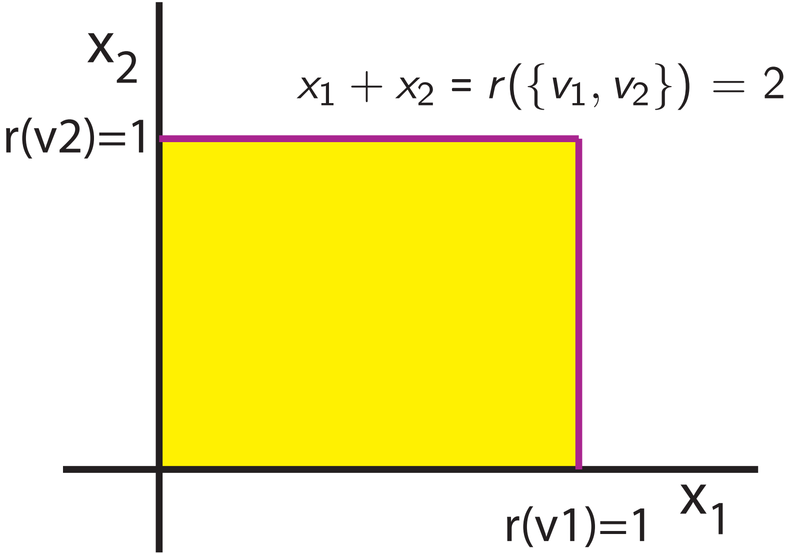
Matroid Polyhedron in 2D



Matroid Polyhedron in 2D

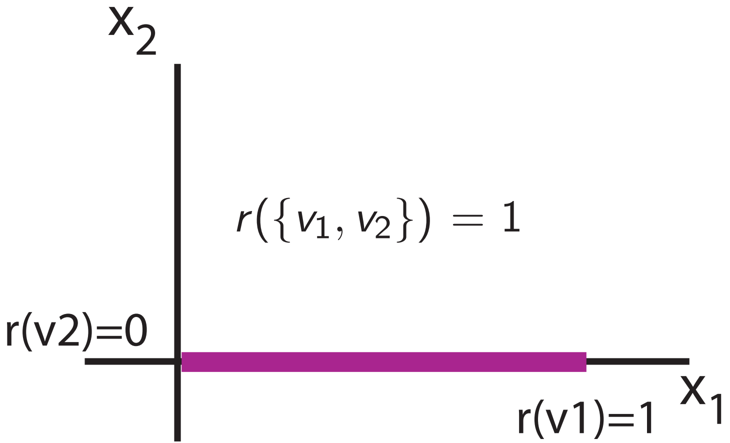


Matroid Polyhedron in 2D

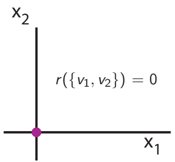


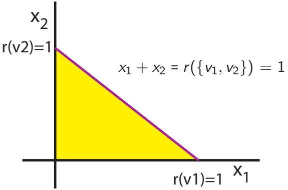
Matroid Polyhedron in 2D

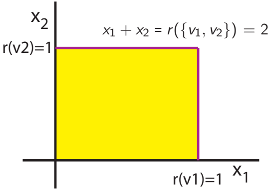
And, if v2 is a loop ...



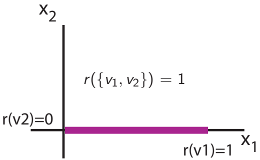
Matroid Polyhedron in 2D

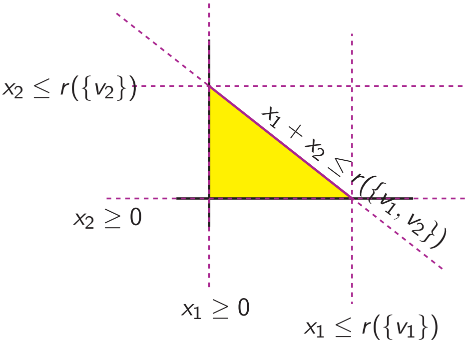




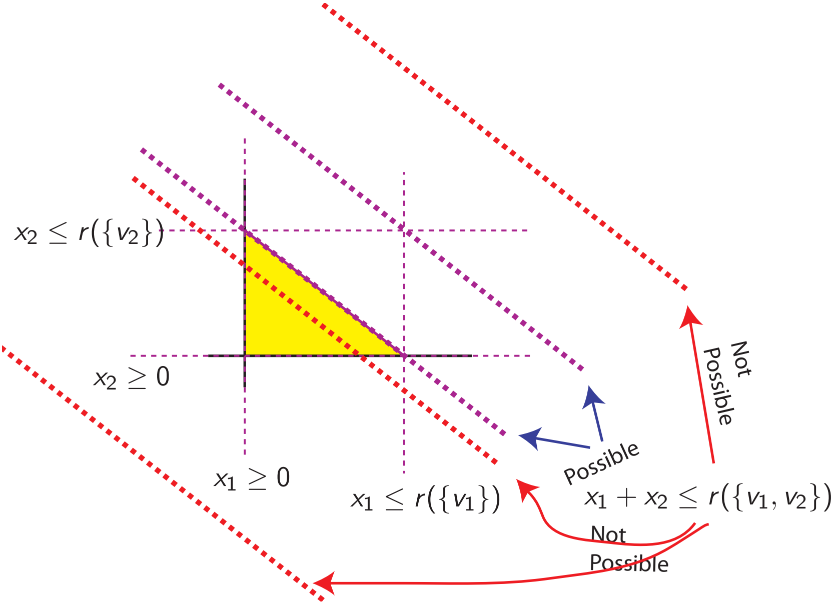


And, if v2 is a loop ...





Matroid Polyhedron in 2D



Matroid Polyhedron in 3D

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{10.28}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{10.29}$$
$$x_1 \leq r(\{v_1\}) \tag{10.30}$$
$$x_2 \leq r(\{v_2\}) \tag{10.31}$$
$$x_3 \leq r(\{v_3\}) \tag{10.32}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{10.33}$$
$$x_2 + x_3 \leq r(\{v_2, v_3\}) \tag{10.34}$$
$$x_1 + x_3 \leq r(\{v_1, v_3\}) \tag{10.35}$$
$$x_1 + x_2 + x_3 \leq r(\{v_1, v_2, v_3\}) \tag{10.36}$$

Matroid Polyhedron in 3D

- Consider the simple cycle matroid on a graph consisting of a 3-cycle, $G = (V, E)$ with matroid $M = (E, \mathcal{I})$ where $I \in \mathcal{I}$ is a forest.

Matroid Polyhedron in 3D

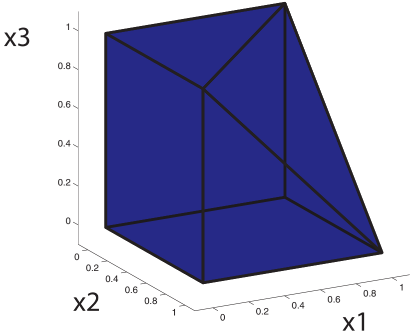
- Consider the simple cycle matroid on a graph consisting of a 3-cycle, $G = (V, E)$ with matroid $M = (E, \mathcal{I})$ where $I \in \mathcal{I}$ is a forest.
- So any set of either one or two edges is independent, and has rank equal to cardinality.

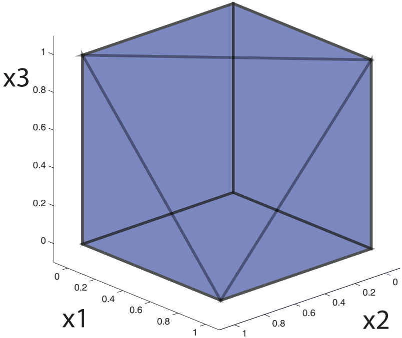
Matroid Polyhedron in 3D

- Consider the simple cycle matroid on a graph consisting of a 3-cycle, $G = (V, E)$ with matroid $M = (E, \mathcal{I})$ where $I \in \mathcal{I}$ is a forest.
- So any set of either one or two edges is independent, and has rank equal to cardinality.
- The set of three edges is dependent, and has rank 2.

Matroid Polyhedron in 3D

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



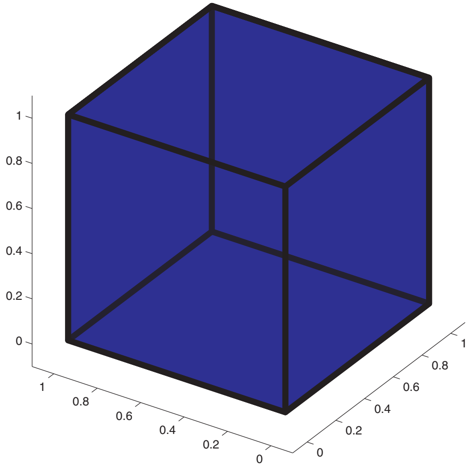


Matroid Polyhedron in 3D

P_r^+ associated with the "free" matroid in 3D.

Matroid Polyhedron in 3D

P_r^+ associated with the "free" matroid in 3D.

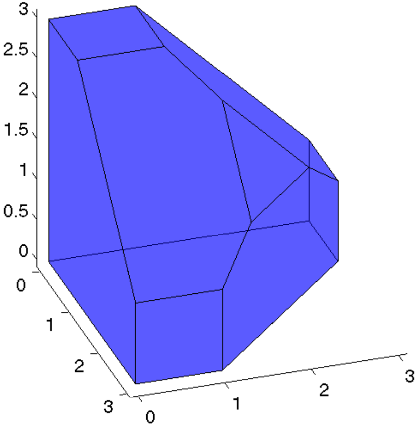


Another Polytope in 3D

Thought question: what kind of polytope might this be?

Another Polytope in 3D

Thought question: what kind of polytope might this be?



Matroid Independence Polyhedron

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

$$\begin{aligned} P_{\text{ind. set}} &= \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \end{aligned} \tag{10.37}$$

Matroid Independence Polyhedron

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

$$\begin{aligned} P_{\text{ind. set}} &= \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \end{aligned} \tag{10.37}$$

- In fact, the two polyhedra are identical (and thus both are polytopes).

Matroid Independence Polyhedron

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

$$\begin{aligned} P_{\text{ind. set}} &= \operatorname{conv}\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \end{aligned} \tag{10.37}$$

- 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 10.5.1

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*

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.38}$$

where $\lambda_i \geq 0$ *satisfy*

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{10.39}$$

Maximum weight independent set via weighted rank

Proof.

- Firstly, note that for any such $w \in \mathbb{R}^E$, we have

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2)\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3)\begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \tag{10.40}$$

...

Maximum weight independent set via weighted rank

Proof.

- Firstly, note that for any such $w \in \mathbb{R}^E$, we have

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3) \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \tag{10.40}$$

- If we can take w in decreasing order ($w_1 \geq w_2 \geq \cdots \geq w_n$), then each coefficient of the vectors is non-negative (except possibly the last one, w_n).

...

Maximum weight independent set via weighted rank

Proof.

- Now, again assuming $w \in \mathbb{R}_+^E$, order the elements of V as $(v_1, v_2, \ldots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \cdots \geq w(v_n)$

...

Maximum weight independent set via weighted rank

Proof.

- Now, again assuming $w \in \mathbb{R}_+^E$, order the elements of V as $(v_1, v_2, \dots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \dots \geq w(v_n)$
- Define the sets U_i based on this order as follows, for $i = 0, \dots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \dots, v_i\} \tag{10.41}$$

Note that

$$\mathbf{1}_{U_0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \mathbf{1}_{U_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \mathbf{1}_{U_\ell} = \begin{pmatrix} \left.\begin{matrix} 1 \\ 1 \\ \vdots \\ 1 \end{matrix}\right\} \ell\times \\ \left.\begin{matrix} 0 \\ 0 \\ \vdots \\ 0 \end{matrix}\right\} (n-\ell)\times \end{pmatrix}, \text{ etc.}$$

Maximum weight independent set via weighted rank

Proof.

- Now, again assuming $w \in \mathbb{R}_+^E$, order the elements of V as $(v_1, v_2, \ldots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \cdots \geq w(v_n)$
- Define the sets U_i based on this order as follows, for $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{10.41}$$

- Define the set I as those elements where the rank increases, i.e.:

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\}. \tag{10.42}$$

Hence, given an i with $v_i \notin I$, $r(U_i) = r(U_{i-1})$.

...

Maximum weight independent set via weighted rank

Proof.

- Now, again assuming $w \in \mathbb{R}_+^E$, order the elements of V as $(v_1, v_2, \ldots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \cdots \geq w(v_n)$
- Define the sets U_i based on this order as follows, for $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{10.41}$$

- Define the set I as those elements where the rank increases, i.e.:

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\}. \tag{10.42}$$

  Hence, given an i with $v_i \notin I$, $r(U_i) = r(U_{i-1})$.
- Therefore, I is the output of the greedy algorithm for $\max\{w(I) | I \in \mathcal{I}\}$. *since items v_i are ordered decreasing by $w(v_i)$, and we only choose the ones that increase the rank, which means they don't violate independence.*

Maximum weight independent set via weighted rank

Proof.

- Now, again assuming $w \in \mathbb{R}_+^E$, order the elements of V as $(v_1, v_2, \dots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \dots \geq w(v_n)$
- Define the sets U_i based on this order as follows, for $i = 0, \dots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \dots, v_i\} \tag{10.41}$$

- Define the set I as those elements where the rank increases, i.e.:

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\}. \tag{10.42}$$

  Hence, given an i with $v_i \notin I$, $r(U_i) = r(U_{i-1})$.
- Therefore, I is the output of the greedy algorithm for $\max \{w(I) | I \in \mathcal{I}\}$.
- And therefore, I is a maximum weight independent set (can even be a base, actually).

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

- And the weight of the independent set $w(I)$ is given by

$$w(I) = \sum_{v \in I} w(v) = \tag{10.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

- And the weight of the independent set $w(I)$ is given by

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{10.45}$$

(10.46)

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

- And the weight of the independent set $w(I)$ is given by

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{10.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) \tag{10.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

- And the weight of the independent set $w(I)$ is given by

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{10.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{10.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{10.44}$$

- And the weight of the independent set $w(I)$ is given by

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{10.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.46}$$

- Since we took $v_1, v_2, \dots$ in decreasing order, for all i, and since $w \in \mathbb{R}_+^E$, we have $\lambda_i \geq 0$

□

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^{\intercal} x \\ \text{subject to} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{10.47}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{array}{lll} \text{maximize} & w^\intercal x & \\ \text{subject to} & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{array} \quad (10.47)$$

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

$$\begin{array}{lll} \text{minimize} & \sum_{U \subseteq V} y_U r(U), & \\ \text{subject to} & y_U \geq 0 & (\forall U \subseteq V) \\ & \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w & \end{array} \quad (10.48)$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{subject to} \quad & x_v \geq 0 \qquad (v \in V) \\ & x(U) \leq r(U) \qquad (\forall U \subseteq V) \end{aligned} \tag{10.47}$$

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

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{subject to} \quad & y_U \geq 0 \qquad (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \end{aligned} \tag{10.48}$$

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

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{10.49}$$

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{10.49}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r^+ \tag{10.50}$$

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

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{aligned} \tag{10.49}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r^+ \tag{10.50}$$

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

- Therefore, since $P_{\text{ind. set}} \subseteq P_r^+$, the above problem can only have a larger solution. I.e.,

$$\max w^\intercal x \text{ s.t. } x \in P_{\text{ind. set}} \leq \max w^\intercal x \text{ s.t. } x \in P_r^+. \tag{10.51}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max \{w(I) : I \in \mathcal{I}\} \leq \max \{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{10.52}$$

$$\leq \max \left\{w^\intercal x : x \in P_r^+\right\} \tag{10.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min \left\{ \sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \right\} \tag{10.54}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{10.52}$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \tag{10.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{10.54}$$

- Theorem 10.5.1 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.55}$$

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. 10.55 is feasible w.r.t. the dual LP).

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{10.52}$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \tag{10.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U\subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{10.54}$$

- Theorem 10.5.1 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{10.55}$$

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. 10.55 is feasible w.r.t. the dual LP).

- Therefore, we also have

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \geq \alpha_{\min} \tag{10.56}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \quad (10.52)$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \quad (10.53)$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \quad (10.54)$$

- Therefore, all the inequalities above are equalities.

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{10.52}$$

$$= \max\{w^\intercal x : x \in P_r^+\} \tag{10.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{10.54}$$

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

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{10.52}$$

$$= \max\{w^\intercal x : x \in P_r^+\} \tag{10.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : \forall U, y_U \geq 0; \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{10.54}$$

- 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 10.5.2

$$P_r^+ = P_{\textit{ind. set}} \tag{10.57}$$

Polytope Equivalence (Summarizing the above)

- For each $I \in \mathcal{I}$ of a matroid $M = (E, \mathcal{I})$, we can form the incidence vector $\mathbf{1}_I$.

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{10.58}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \tag{10.58}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{10.59}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \tag{10.58}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{10.59}$$

Theorem 10.5.3

$$P_r^+ = P_{\textit{ind. set}} \tag{10.60}$$

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).

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 10.52, 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:

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 10.52, 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 10.5.4

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

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 10.52, 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 10.5.4

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).

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 10.52, 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 10.5.4

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.

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:

$$x \geq 0 \tag{10.61}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{10.62}$$

$$x(V) = r(V) \tag{10.63}$$

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:

$$x \geq 0 \tag{10.61}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{10.62}$$

$$x(V) = r(V) \tag{10.63}$$

- Note the third requirement, $x(V) = r(V)$.

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:

$$x \geq 0 \tag{10.61}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{10.62}$$

$$x(V) = r(V) \tag{10.63}$$

- 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. 10.61- 10.63 above.

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:

$$x \geq 0 \tag{10.61}$$
$$x(A) \leq r(A) \; \forall A \subseteq V \tag{10.62}$$
$$x(V) = r(V) \tag{10.63}$$

- 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. 10.61- 10.63 above.
- What does this look like?

Spanning set polytope

- Recall, a set A is spanning in a matroid $M = (E, \mathcal{I})$ if $r(A) = r(E)$.

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 $P_{\text{spanning}}(M)$.

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 10.5.5

The spanning set polytope is determined by the following equations:

$$0 \leq x_e \leq 1 \qquad \text{for } e \in E \tag{10.64}$$

$$x(A) \geq r(E) - r(E \setminus A) \qquad \text{for } A \subseteq E \tag{10.65}$$

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 $P_{\text{spanning}}(M)$.

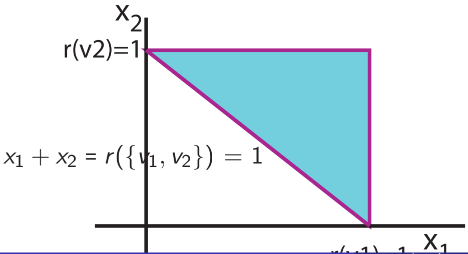
Theorem 10.5.5

The spanning set polytope is determined by the following equations:

$$0 \leq x_e \leq 1 \qquad \text{for } e \in E \tag{10.64}$$

$$x(A) \geq r(E) - r(E \setminus A) \qquad \text{for } A \subseteq E \tag{10.65}$$

- 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).

...

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

$$x \in P_{\text{spanning}}(M) \Leftrightarrow 1 - x \in P_{\text{ind. set}}(M^*) \tag{10.66}$$

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:

$$x = \sum_i \lambda_i \mathbf{1}_{A_i} \tag{10.67}$$

where A_i is spanning in M.

. . .

Spanning set polytope

. . . proof continued.

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

$$x = \sum_i \lambda_i \mathbf{1}_{A_i} \tag{10.67}$$

where A_i is spanning in M.

- Consider

$$\mathbf{1} - x = \mathbf{1}_E - x = \mathbf{1}_E - \sum_i \lambda_i \mathbf{1}_{A_i} = \sum_i \lambda_i \mathbf{1}_{E \setminus A_i}, \tag{10.68}$$

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

$$\mathbf{1} - x \geq 0 \tag{10.69}$$

$$\mathbf{1}_A - x(A) = |A| - x(A) \leq r_{M^*}(A) \text{ for } A \subseteq E \tag{10.70}$$

And we know the dual rank function is

$$r_{M^*}(A) = |A| + r_M(E \setminus A) - r_M(E) \tag{10.71}$$

. . .

Spanning set polytope

. . . proof continued.

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

$$\mathbf{1} - x \geq 0 \quad (10.69)$$

$$\mathbf{1}_A - x(A) = |A| - x(A) \leq r_{M^*}(A) \text{ for } A \subseteq E \quad (10.70)$$

And we know the dual rank function is

$$r_{M^*}(A) = |A| + r_M(E \setminus A) - r_M(E) \quad (10.71)$$

- giving

$$x(A) \geq r_M(E) - r_M(E \setminus A) \text{ for all } A \subseteq E \quad (10.72)$$

□

Matroids

where are we going with this?

- We've been discussing results about matroids (independence polytope, etc.).

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.

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 $\mathfrak{P}$.

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

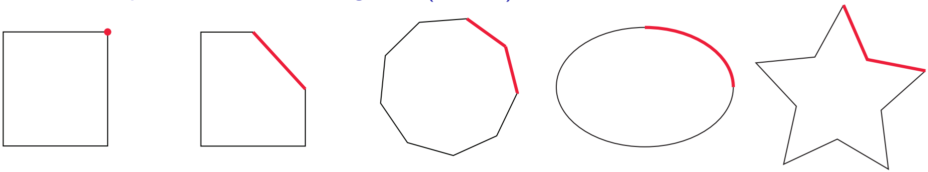
$$x + \epsilon \mathbf{1}_e \notin P \tag{10.73}$$

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

$$x + \epsilon \mathbf{1}_e \notin P \tag{10.73}$$

- Examples of maximal regions (in red)

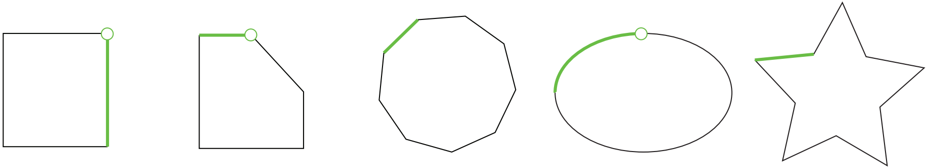


Maximal points in a set

- 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

$$x + \epsilon \mathbf{1}_e \notin P \tag{10.73}$$

- Examples of non-maximal regions (in green)



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 10.6.1 (subvector)

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

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

Definition 10.6.1 (subvector)

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

Definition 10.6.2 (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.

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

Definition 10.6.1 (subvector)

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

Definition 10.6.2 (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).

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

Definition 10.6.1 (subvector)

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

Definition 10.6.2 (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:

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

Definition 10.6.1 (subvector)

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

Definition 10.6.2 (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

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

Definition 10.6.1 (subvector)

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

Definition 10.6.2 (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 $M = (E, \mathcal{I})$.

$$\text{rank}(A) = \max\{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \quad (10.74)$$

A vector form of rank

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

$$\text{rank}(A) = \max \{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.74}$$

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

$$\text{rank}(x) = \max \left(y(E) : y \leq x, y \in P\right) = \max_{y \in P} (x \wedge y)(E) \tag{10.75}$$

where $y \leq x$ is componentwise inequality ($y_i \leq x_i, \forall i$), and where $(x \wedge y) \in \mathbb{R}_+^E$ has $(x \wedge y)(i) = \min(x(i), y(i))$.

A vector form of rank

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

$$\text{rank}(A) = \max\{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.74}$$

- 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:

$$\text{rank}(x) = \max(y(E) : y \leq x, y \in P) = \max_{y \in P} (x \wedge y)(E) \tag{10.75}$$

  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 $\text{rank}(x) = \max_{y \in \mathcal{B}_x} y(E)$.

A vector form of rank

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

$$\text{rank}(A) = \max\{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.74}$$

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

$$\text{rank}(x) = \max\left(y(E) : y \leq x, y \in P\right) = \max_{y \in P} (x \wedge y)(E) \tag{10.75}$$

  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).

A vector form of rank

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

$$\text{rank}(A) = \max \{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \tag{10.74}$$

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

$$\text{rank}(x) = \max \left(y(E) : y \leq x, y \in P\right) = \max_{y \in P} (x \wedge y)(E) \tag{10.75}$$

  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).
- 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 10.6.3 (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 $y(E)$

Polymatroidal polyhedron (or a "polymatroid")

Definition 10.6.3 (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 $y^1(E) = y^2(E)$.

Polymatroidal polyhedron (or a "polymatroid")

Definition 10.6.3 (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 $y(E) = \text{rank}(x)$.

Polymatroidal polyhedron (or a "polymatroid")

Definition 10.6.3 (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 10.6.3 (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.

Polymatroidal polyhedron (or a "polymatroid")

Definition 10.6.3 (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:

A Polymatroid is:

Matroid and Polymatroid: side-by-side

A Matroid is:

1. a set system $(E, \mathcal{I})$

A Polymatroid is:

1. a compact set $P \subseteq \mathbb{R}_+^E$

Matroid and Polymatroid: side-by-side

A Matroid is:

1. a set system $(E, \mathcal{I})$
2. empty-set containing $\emptyset \in \mathcal{I}$

A Polymatroid is:

1. a compact set $P \subseteq \mathbb{R}_+^E$
2. zero containing, $\mathbf{0} \in P$

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, $\emptyset \subseteq I' \subseteq I \in \mathcal{I} \Rightarrow I' \in \mathcal{I}$.

A Polymatroid is:

1. a compact set $P \subseteq \mathbb{R}_+^E$
2. zero containing, $\mathbf{0} \in P$
3. down monotone, $0 \leq y \leq x \in P \Rightarrow y \in P$

Matroid and Polymatroid: side-by-side

A Matroid is:

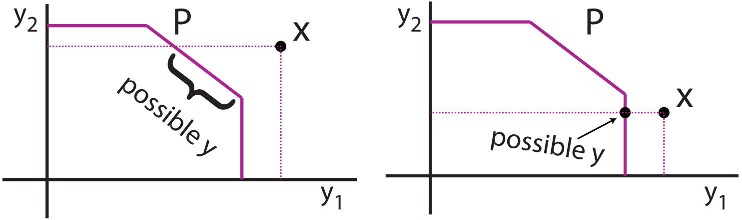
1. a set system $(E, \mathcal{I})$
2. empty-set containing $\emptyset \in \mathcal{I}$
3. down closed, $\emptyset \subseteq I' \subseteq I \in \mathcal{I} \Rightarrow I' \in \mathcal{I}$.
4. any maximal set I in $\mathcal{I}$, bounded by another set A, has the same matroid rank (any maximal independent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

.

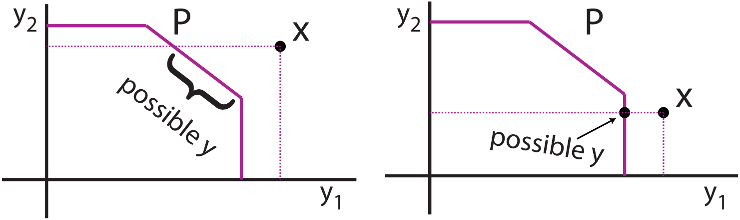
Polymatroidal polyhedron (or a "polymatroid")



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 $y(E) = y_1 + y_2 = \text{const}$.

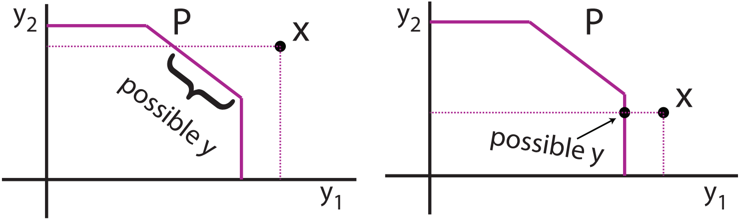
Polymatroidal polyhedron (or a "polymatroid")



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 $y(E) = y_1 + y_2 = \text{const}$.
- 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 $y(E)$.

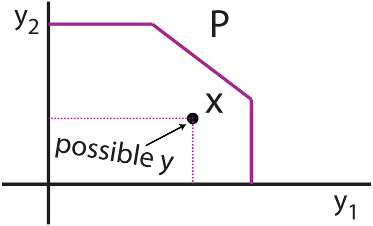
Polymatroidal polyhedron (or a "polymatroid")



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.

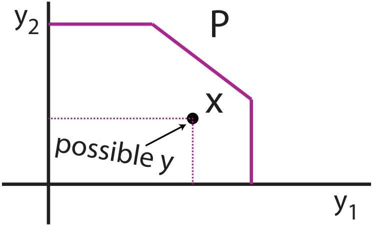
Polymatroidal polyhedron (or a "polymatroid")



$\exists$ 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.

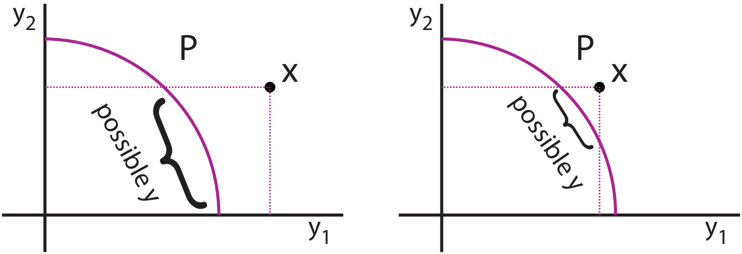
Polymatroidal polyhedron (or a “polymatroid”)



$\exists$ 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)

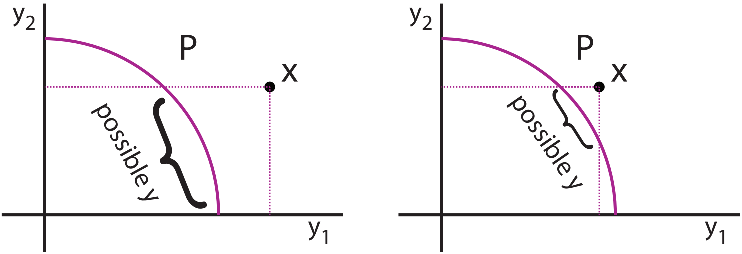
Polymatroid as well?



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.

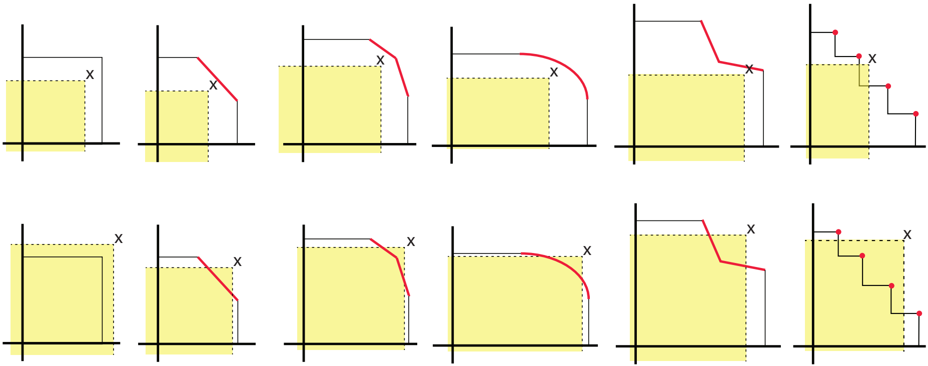
Polymatroid as well? no



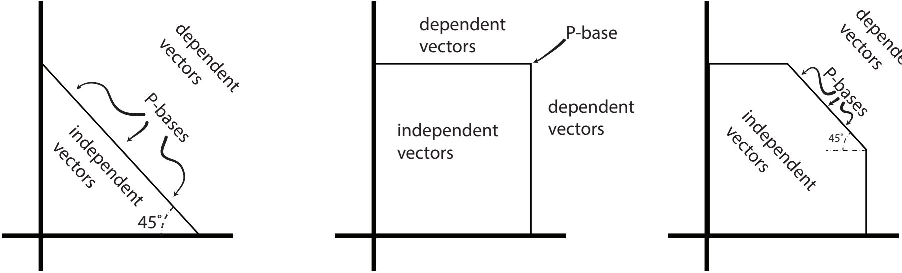
Left and right: $\exists$ multiple maximal $y \leq x$ as indicated.

- 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?



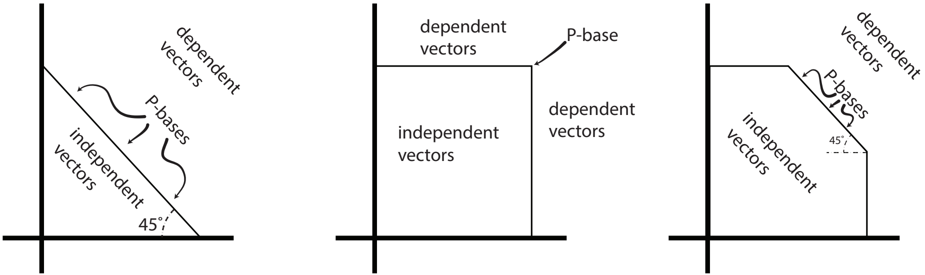
Some possible polymatroid forms in 2D



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

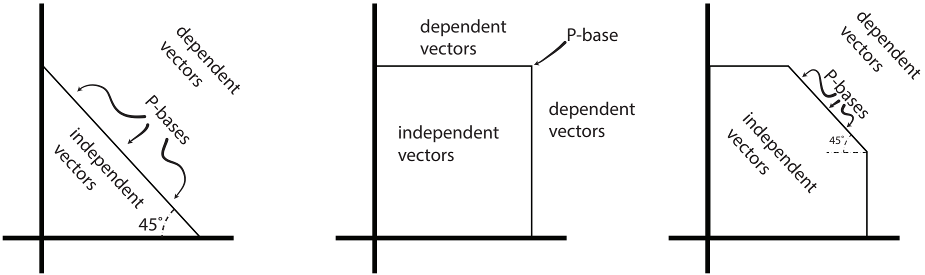
Some possible polymatroid forms in 2D



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

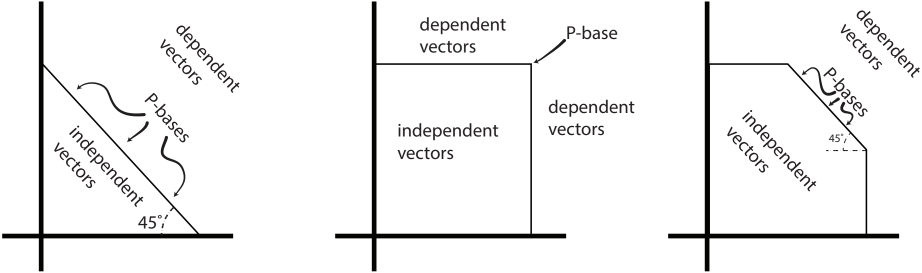
Some possible polymatroid forms in 2D



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

Some possible polymatroid forms in 2D

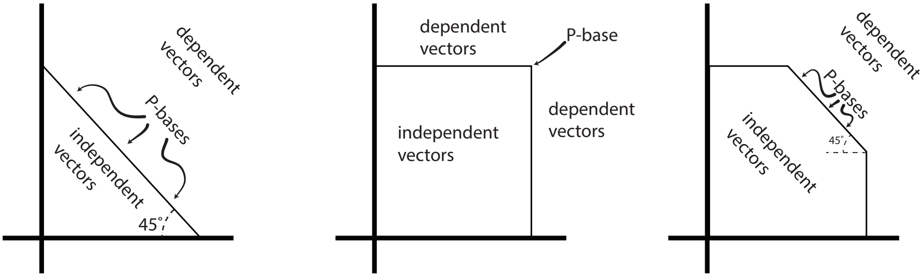


It appears that we have three possible forms of polymatroid in 2D, when neither of the elements $\{v_1, v_2\}$ are self-dependent.

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.

Some possible polymatroid forms in 2D

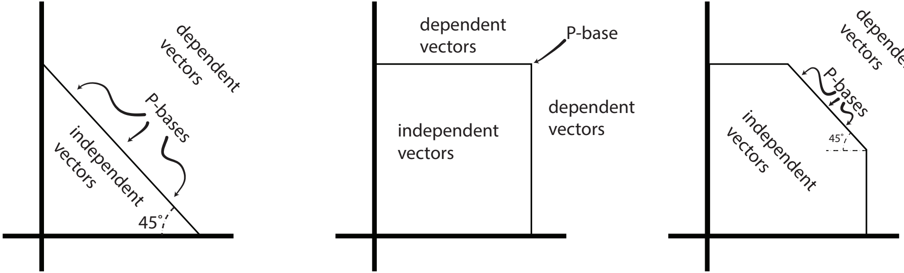


It appears that we have three possible forms of polymatroid in 2D, when neither of the elements $\{v_1, v_2\}$ are self-dependent.

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.

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")

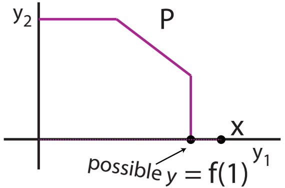
- Note that if x contains any zeros (i.e., suppose that $x \in \mathbb{R}_+^E$ has $E \setminus S$ s.t. $x(E \setminus S) = 0$, so S indicates the non-zero elements, or $S = \text{supp}(x)$), then this also forces $y(E \setminus S) = 0$, so that $y(E) = y(S)$. This is true either for $x \in P$ or $x \notin P$.

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.

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 10.6.4

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

$$P_f^+ = \left\{ y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E \right\} \tag{10.76}$$

$$= \left\{ y \in \mathbb{R}^E : y \geq 0, y(A) \leq f(A) \text{ for all } A \subseteq E \right\} \tag{10.77}$$

Associated polyhedron with a polymatroid function

$$P_f^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\right\} \tag{10.78}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{10.79}$$

$$x_1 \leq f(\{v_1\}) \tag{10.80}$$

$$x_2 \leq f(\{v_2\}) \tag{10.81}$$

$$x_3 \leq f(\{v_3\}) \tag{10.82}$$

$$x_1 + x_2 \leq f(\{v_1, v_2\}) \tag{10.83}$$

$$x_2 + x_3 \leq f(\{v_2, v_3\}) \tag{10.84}$$

$$x_1 + x_3 \leq f(\{v_1, v_3\}) \tag{10.85}$$

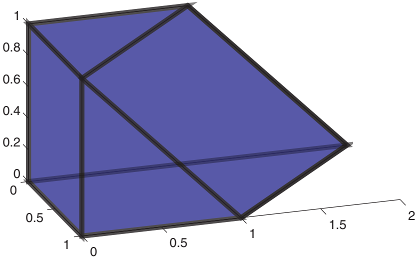
$$x_1 + x_2 + x_3 \leq f(\{v_1, v_2, v_3\}) \tag{10.86}$$

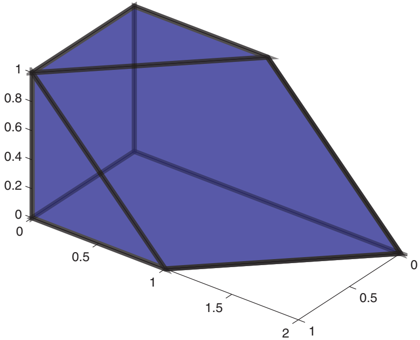
Associated polyhedron with a polymatroid function

- 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.

Associated polyhedron with a polymatroid function

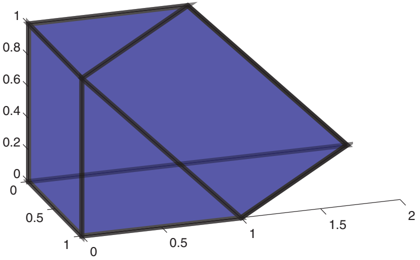
- Consider the asymmetric graph cut function on the simple chain graph $v_1 - v_2 - v_3$. That is, $f(S) = |\{(v,s) \in E(G) : v \in V, s \in S\}|$ is count of any edges within S or between S and $V \setminus S$, so that $\delta(S) = f(S) + f(V \setminus S) - f(V)$ is the standard graph cut.
- Observe: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

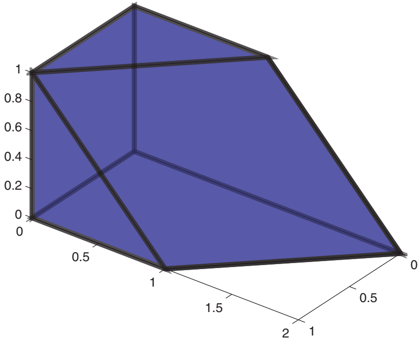




Associated polyhedron with a polymatroid function

- 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):





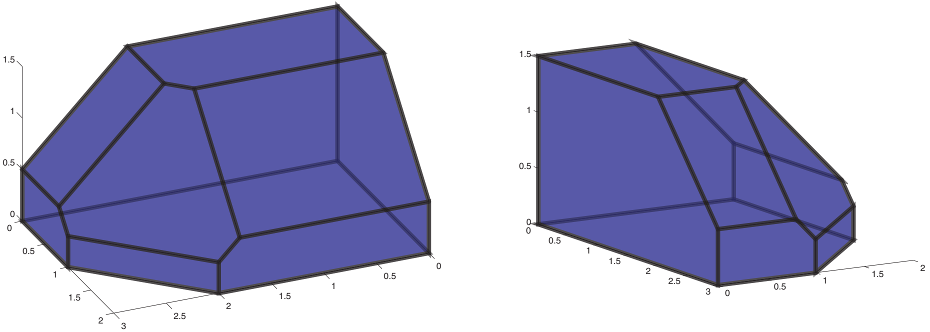
- which axis is which?

Associated polyhedron with a polymatroid function

- Consider: $f(\emptyset) = 0$, $f(\{v_1\}) = 1.5$, $f(\{v_2\}) = 2$, $f(\{v_1, v_2\}) = 2.5$, $f(\{v_3\}) = 3$, $f(\{v_3, v_1\}) = 3.5$, $f(\{v_3, v_2\}) = 4$, $f(\{v_3, v_2, v_1\}) = 4.3$.

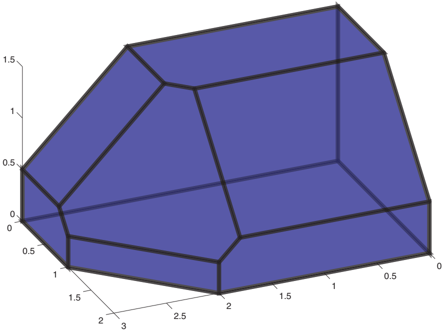
Associated polyhedron with a polymatroid function

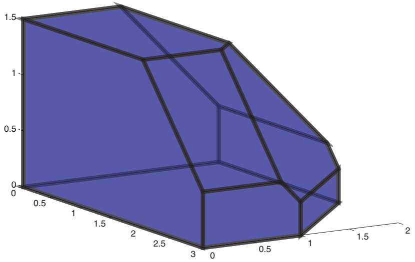
- Consider: $f(\emptyset) = 0$, $f(\{v_1\}) = 1.5$, $f(\{v_2\}) = 2$, $f(\{v_1, v_2\}) = 2.5$, $f(\{v_3\}) = 3$, $f(\{v_3, v_1\}) = 3.5$, $f(\{v_3, v_2\}) = 4$, $f(\{v_3, v_2, v_1\}) = 4.3$.
- Observe: P_f^+ (at two views):



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):





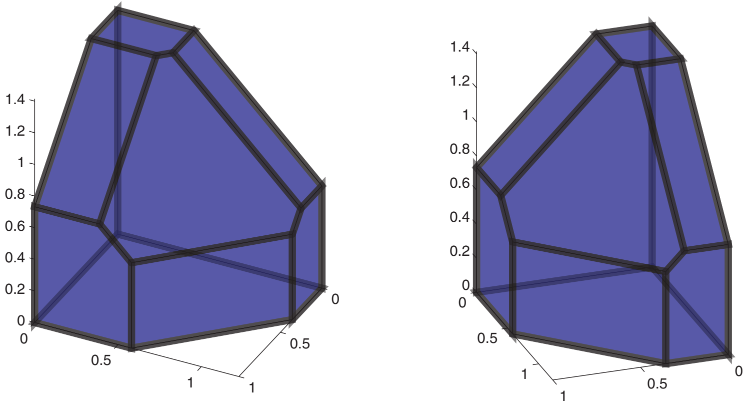
- which axis is which?

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 $f(S) = \sqrt{w(S)}$.

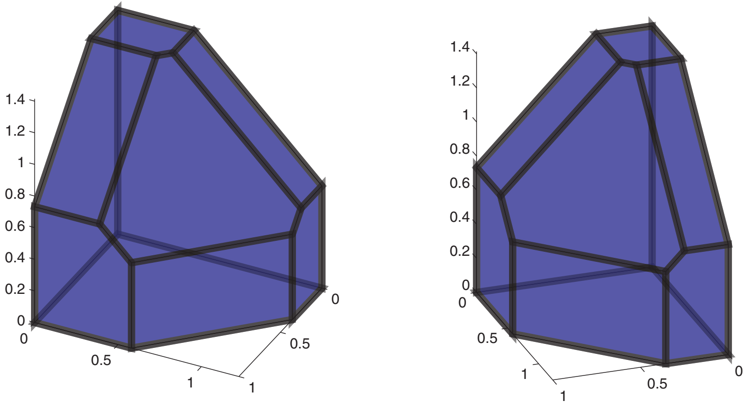
Associated polyhedron with a polymatroid function

- 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):



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 $g(3) = 5.5$.

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?

Associated polytope with a non-submodular function

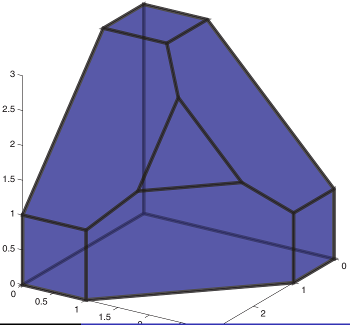
- Consider function on integers: $g(0) = 0$, $g(1) = 3$, $g(2) = 4$, and $g(3) = 5.5$. Is $f(S) = g(|S|)$ submodular? $f(S) = g(|S|)$ is not submodular since $f(\{e_1, e_3\}) + f(\{e_1, e_2\}) = 4 + 4 = 8$ but $f(\{e_1, e_2, e_3\}) + f(\{e_1\}) = 5.5 + 3 = 8.5$.

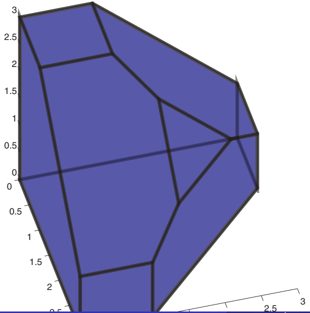
Associated polytope with a non-submodular function

- 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, $1 = g(1 + 1) - g(1) < g(2 + 1) - g(2) = 1.5$.

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:

A polymatroid vs. a polymatroid function's polyhedron

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \le f(A) \text{ for all } A \subseteq E\} \tag{10.87}$$

A polymatroid vs. a polymatroid function's polyhedron

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\} \tag{10.87}$$

  - 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 $y(E)$).

A polymatroid vs. a polymatroid function's polyhedron

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \leq f(A) \text{ for all } A \subseteq E\} \tag{10.87}$$

  - 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?

A polymatroid vs. a polymatroid function's polyhedron

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

$$P_f^+ = \{y \in \mathbb{R}_+^E : y(A) \le f(A) \text{ for all } A \subseteq E\} \tag{10.87}$$

  - 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 10.6.5

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

$$y^x(E) = \text{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \quad (10.88)$$

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

A polymatroid function's polyhedron is a polymatroid.

Theorem 10.6.5

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

$$y^x(E) = \text{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{10.88}$$

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:

$$f(B) = \max\left\{y(B) : y \in P_f^+\right\} \tag{10.89}$$

A polymatroid function's polyhedron is a polymatroid.

Theorem 10.6.5

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

$$y^x(E) = \textit{rank}(x) = \max\left(y(E) : y \leq x, y \in P_f^+\right) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \quad (10.88)$$

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:

$$f(B) = \max\left\{y(B) : y \in P_f^+\right\} \quad (10.89)$$

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