Submodular Functions, Optimization, and Applications to Machine Learning

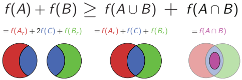
— Spring Quarter, Lecture 8 —

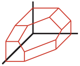
http://j.ee.washington.edu/~bilmes/classes/ee596b_spring_2014/

Prof. Jeff Bilmes

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

April 23rd, 2014





Cumulative Outstanding Reading

- Read chapters 1 and 2, and sections 3.1-3.2 from Fujishige's book.

Announcements, Assignments, and Reminders

- Homework 1 is out, due Wednesday April 23rd, 11:45pm, electronically via our assignment dropbox (`https://canvas.uw.edu/courses/895956/assignments`).
- All homeworks must be done electronically, only PDF file format accepted.
- Weekly Office Hours: Wednesdays, 5:00-5:50, or by skype or google hangout (email me).

Class Road Map - IT-I

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

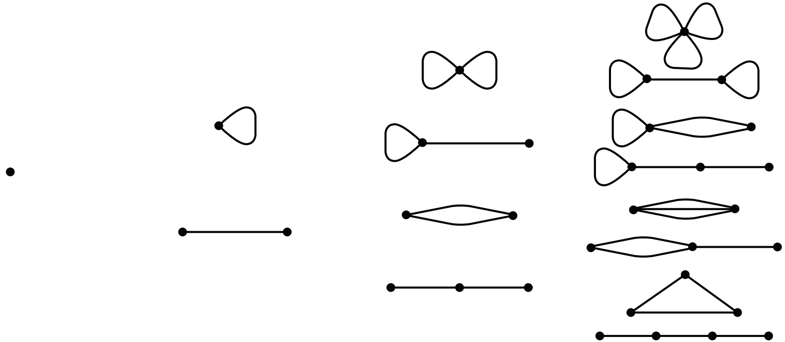
Finals Week: June 9th-13th, 2014.

Matroid Operations

- Matroid restriction/deletion
- Matroid contraction
- Matroid minor (series of deletions & contractions)
- Matroid intersection and its rank (convolution)
- Matroid union and its rank (convolution)

Matroids of three or fewer elements are graphic

- All matroids up to and including three elements are graphic.



(a) The only matroid with zero elements.

(b) The two one-element matroids.

(c) The four two-element matroids.

(d) The eight three-element matroids.

- This is a nice way to show matroids with low ground set sizes. What about matroids that are low rank but with many elements?

Affine Matroids

- Given an $n \times m$ matrix with entries over some field $\mathbb{F}$, we say that a subset $S \subseteq \{1, \dots, m\}$ of indices (with corresponding column vectors $\{v_i : i \in S\}$, with $|S| = k$) is affinely dependent if $m \geq 1$ and there exists elements $\{a_1, \dots, a_k\} \in \mathbb{F}$, not all zero with $\sum_{i=1}^{k} a_i = 0$, such that $\sum_{i=1}^{k} a_i v_i = 0$.
- Otherwise, the set is called affinely independent.
- Concisely: points $\{v_1, v_2, \dots, v_k\}$ are affinely independent if $v_2 - v_1, v_3 - v_1, \dots, v_k - v_1$ are linearly independent.
- Example: in 2D, three collinear points are affinely dependent, three non-collear points are affinely independent, and ≥ 4 non-collinear points are affinely dependent.

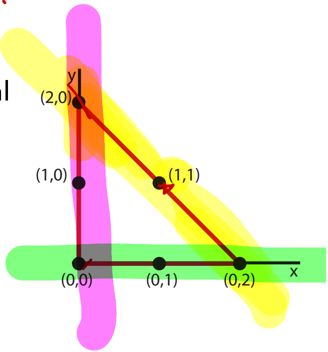
Proposition 8.2.7 (affine matroid)

Let ground set $E = \{1, \dots, m\}$ index column vectors of a matrix, and let $\mathcal{I}$ be the set of subsets X of E such that X indices affinely independent vectors. Then $(E, \mathcal{I})$ is a matroid.

Exercise: prove this.

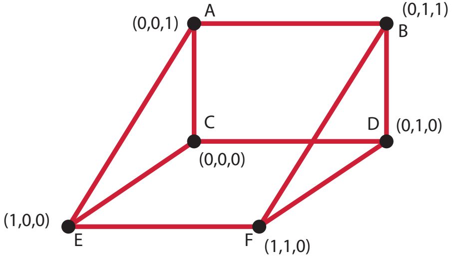
Euclidean Representation of Low-rank Matroids

- Consider the affine matroid with $n \times m = 2 \times 6$ matrix on the field $\mathbb{F} = \mathbb{R}$, and let the elements be $\{(0,0),(1,0),(2,0),(0,1),(0,2),(1,1)\}$.
- We can plot the points in $\mathbb{R}^2$ as on the right:
- Points have rank 1, lines have rank 2, planes have rank 3.
- Flats (points, lines, planes, etc.) have rank equal to one more than their geometric dimension.
- Any two points constitute a line, but lines with only two points are not drawn.
- Lines indicate collinear sets with ≥ 3 points, while any two points have rank 2.
- Dependent sets consist of all subsets with ≥ 4 elements (rank 3), or 3 collinear elements (rank 2). Any two points have rank 2.



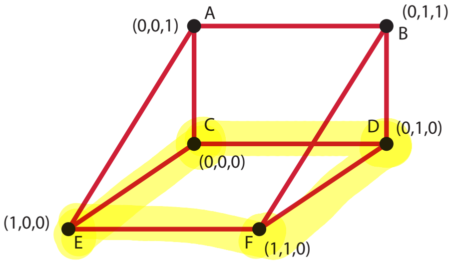
Euclidean Representation of Low-rank Matroids

- As another example on the right, a rank 4 matroid



Euclidean Representation of Low-rank Matroids

- As another example on the right, a rank 4 matroid



- All sets of 5 points are dependent. The only other sets of dependent points are coplanar ones of size 4. Namely: $\{(0,0,0),(0,1,0),(1,1,0),(1,0,0)\}$, $\{(0,0,0),(0,0,1),(0,1,1),(0,1,0)\}$, and $\{(0,0,1),(0,1,1),(1,1,0),(1,0,0)\}$.

Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:

Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;

4

Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;

Euclidean Representation of Low-rank Matroids

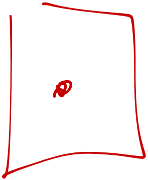
- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;
  3. $|X| \geq 4$ and the points are coplanar; or

Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;
  3. $|X| \geq 4$ and the points are coplanar; or
  4. $|X| \geq 5$ and the points are in space.

Euclidean Representation of Low-rank Matroids

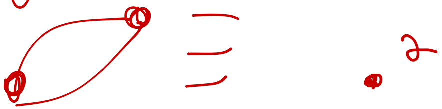
- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;
  3. $|X| \geq 4$ and the points are coplanar; or
  4. $|X| \geq 5$ and the points are in space.
- When they exist, loops are represented in a geometry by a separate box indicating how many loops there are.



Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;
  3. $|X| \geq 4$ and the points are coplanar; or
  4. $|X| \geq 5$ and the points are in space.
- When they exist, loops are represented in a geometry by a separate box indicating how many loops there are.
- Parallel elements, when they exist in a matroid, are indicated by a multiplicity next to a point.

Graph Geom



Euclidean Representation of Low-rank Matroids

- In general, for a matroid $\mathcal{M}$ of rank $m+1$ with $m \leq 3$, then a subset X in a geometric representation in $\mathbb{R}^m$ is dependent if:
  1. $|X| \geq 2$ and the points are identical;
  2. $|X| \geq 3$ and the points are collinear;
  3. $|X| \geq 4$ and the points are coplanar; or
  4. $|X| \geq 5$ and the points are in space.
- When they exist, loops are represented in a geometry by a separate box indicating how many loops there are.
- Parallel elements, when they exist in a matroid, are indicated by a multiplicity next to a point.

Theorem 8.3.1

Any matroid of rank $m \leq 4$ can be represented by an affine matroid in $\mathcal{R}^{m-1}$.

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)
- every plane contains at least three non-collinear points (not dependent unless > 3)

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)
- every plane contains at least three non-collinear points (not dependent unless > 3)
- any three distinct non-collinear points lie on a plane

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)
- every plane contains at least three non-collinear points (not dependent unless > 3)
- any three distinct non-collinear points lie on a plane
- If diagram has at most one plane, then any two distinct lines meet in at most one point.

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)
- every plane contains at least three non-collinear points (not dependent unless > 3)
- any three distinct non-collinear points lie on a plane
- If diagram has at most one plane, then any two distinct lines meet in at most one point.
- If diagram has more than one plane, then: 1) any two distinct planes meeting in more than two points do so in a line; 2) any two distinct lines meeting in a point do so in at most one point and lie in on a common plane; 3) any line not lying on a plane intersects it in at most one point.

Euclidean Rep. of Low-rank Matroids: Conditions

- rank-1 (resp. rank-2, rank-3) flats correspond to points (resp. lines, planes).
- a set of parallel points (could be size 1) does not touch another set of parallel points (could be size 1).
- every line contains at least two points (not dependent unless > 2).
- any two distinct points lie on a line (often not drawn when only two)
- every plane contains at least three non-collinear points (not dependent unless > 3)
- any three distinct non-collinear points lie on a plane
- If diagram has at most one plane, then any two distinct lines meet in at most one point.
- If diagram has more than one plane, then: 1) any two distinct planes meeting in more than two points do so in a line; 2) any two distinct lines meeting in a point do so in at most one point and lie in on a common plane; 3) any line not lying on a plane intersects it in at most one point.
- (see Oxley 2011 for more details).

Euclidean Representation of Low-rank Matroids

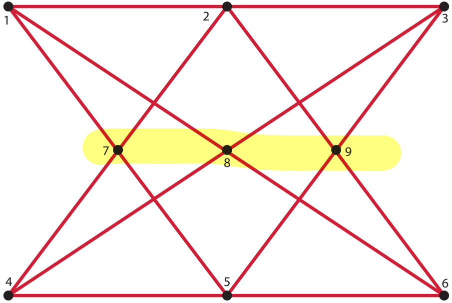
- Very useful for graphically depicting low-rank matrices but which still have rich structure. Also useful for answering questions.

Euclidean Representation of Low-rank Matroids

- Very useful for graphically depicting low-rank matrices but which still have rich structure. Also useful for answering questions.
- Example: Is there a matroid that is not representable (i.e., not linear for some field)?

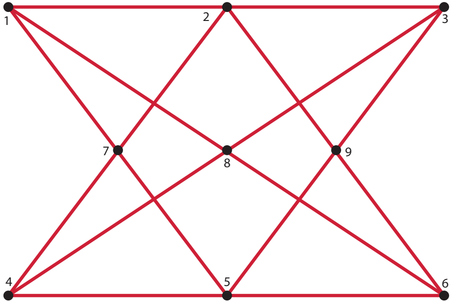
Euclidean Representation of Low-rank Matroids

- Very useful for graphically depicting low-rank matrices but which still have rich structure. Also useful for answering questions.
- Example: Is there a matroid that is not representable (i.e., not linear for some field)? Yes, consider the matroid



Euclidean Representation of Low-rank Matroids

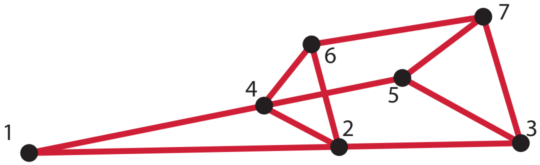
- Very useful for graphically depicting low-rank matrices but which still have rich structure. Also useful for answering questions.
- Example: Is there a matroid that is not representable (i.e., not linear for some field)? Yes, consider the matroid



- Called the non-Pappus matroid. Has rank three, but any matric matroid with the above dependencies would require that $\{7, 8, 9\}$ is dependent, hence requiring an additional line in the above.

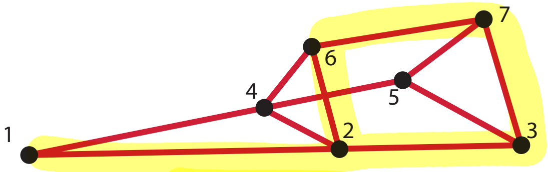
Euclidean Representation of Low-rank Matroids: A test

- Is this a matroid?



Euclidean Representation of Low-rank Matroids: A test

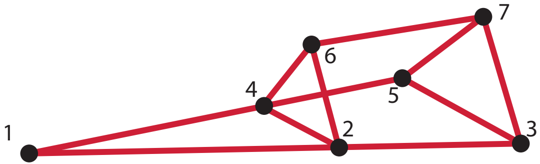
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) =$

Euclidean Representation of Low-rank Matroids: A test

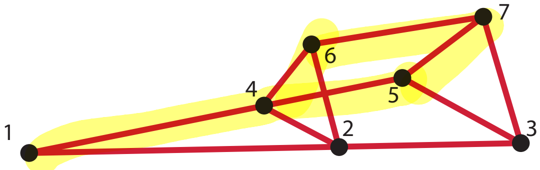
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$

Euclidean Representation of Low-rank Matroids: A test

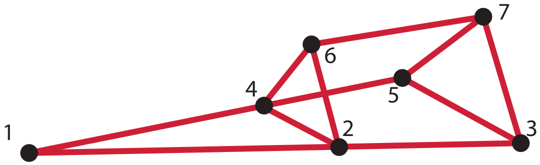
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) =$

Euclidean Representation of Low-rank Matroids: A test

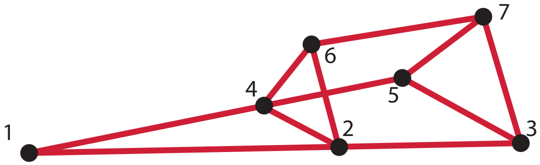
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$

Euclidean Representation of Low-rank Matroids: A test

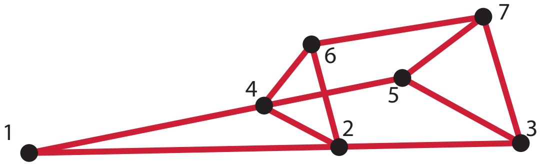
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) =$

Euclidean Representation of Low-rank Matroids: A test

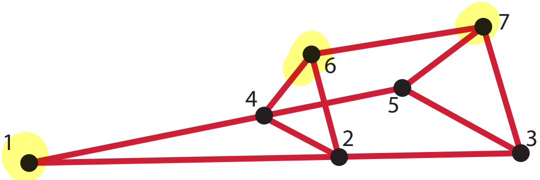
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) = 4$

Euclidean Representation of Low-rank Matroids: A test

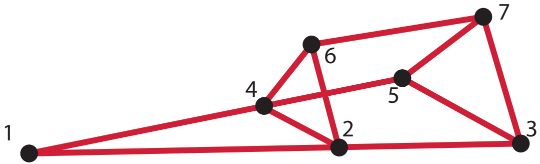
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) = 4$, so we must have, by submodularity, that $r(\{1, 6, 7\}) = r(X \cap Y) \leq r(X) + r(Y) - r(X \cup Y) = 2$.

Euclidean Representation of Low-rank Matroids: A test

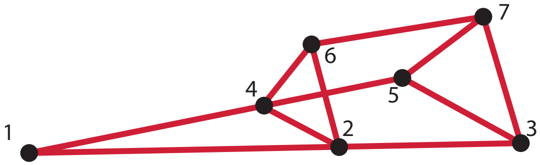
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) = 4$, so we must have, by submodularity, that $r(\{1, 6, 7\}) = r(X \cap Y) \leq r(X) + r(Y) - r(X \cup Y) = 2$.
- However, from the diagram, we have that since 1, 6, 7 are distinct non-collinear points, we have that $r(X \cap Y) =$

Euclidean Representation of Low-rank Matroids: A test

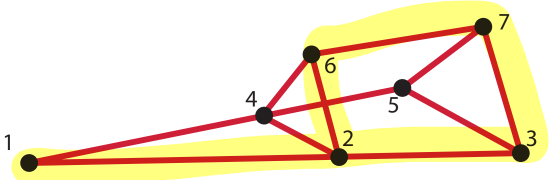
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1, 2, 3, 6, 7\}$, $Y = \{1, 4, 5, 6, 7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) = 4$, so we must have, by submodularity, that $r(\{1, 6, 7\}) = r(X \cap Y) \leq r(X) + r(Y) - r(X \cup Y) = 2$.
- However, from the diagram, we have that since 1, 6, 7 are distinct non-collinear points, we have that $r(X \cap Y) = 3$

Euclidean Representation of Low-rank Matroids: A test

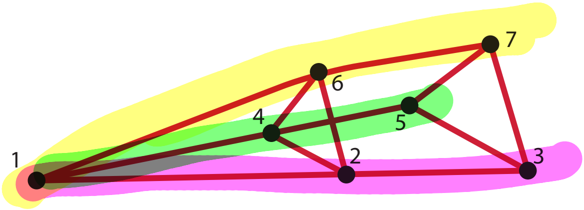
- Is this a matroid?



- Check rank's submodularity: Let $X = \{1,2,3,6,7\}$, $Y = \{1,4,5,6,7\}$. So $r(X) = 3$, and $r(Y) = 3$, and $r(X \cup Y) = 4$, so we must have, by submodularity, that $r(\{1,6,7\}) = r(X \cap Y) \leq r(X) + r(Y) - r(X \cup Y) = 2$.
- However, from the diagram, we have that since 1, 6, 7 are distinct non-collinear points, we have that $r(X \cap Y) = 3$

Euclidean Representation of Low-rank Matroids: A test

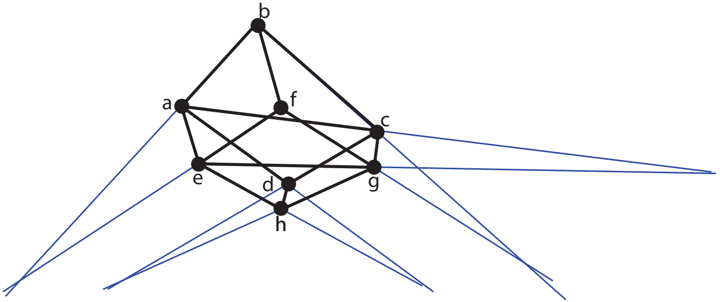
- Is this a matroid?



- If we extend the line from 6-7 to 1, then is it a matroid?
- Hence, not all 2D or 3D graphs of points and lines are matroids.

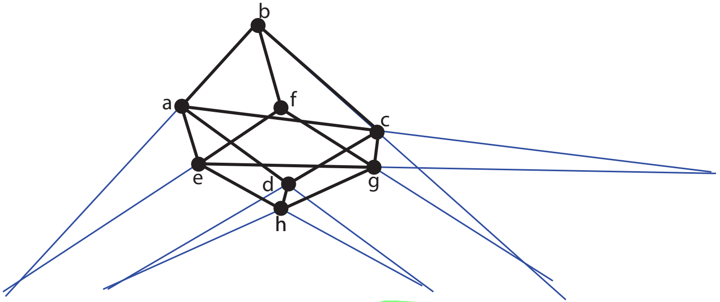
Matroid?

- Consider the following geometry on $|V| = 8$ points with $V = \{a, b, c, d, e, f, g, h\}$.



Matroid?

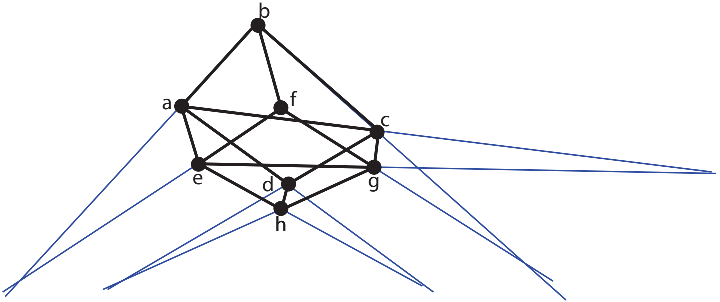
- Consider the following geometry on $|V| = 8$ points with $V = \{a, b, c, d, e, f, g, h\}$.



- Note, we are given that the points $\{b, d, h, f\}$ are not coplanar. However, the following sets of points are coplanar: $\{a, b, e, f\}$, $\{d, c, g, h\}$, $\{a, d, h, e\}$, $\{b, c, g, f\}$, $\{b, c, d, a\}$, $\{f, g, h, e\}$, and $\{a, c, g, e\}$.

Matroid?

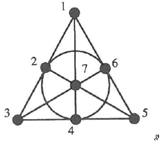
- Consider the following geometry on $|V| = 8$ points with $V = \{a, b, c, d, e, f, g, h\}$.

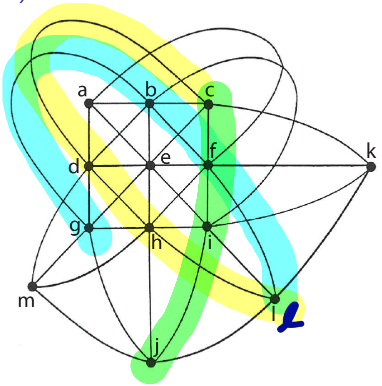


- Note, we are given that the points $\{b, d, h, f\}$ are not coplanar. However, the following sets of points are coplanar: $\{a, b, e, f\}$, $\{d, c, g, h\}$, $\{a, d, h, e\}$, $\{b, c, g, f\}$, $\{b, c, d, a\}$, $\{f, g, h, e\}$, and $\{a, c, g, e\}$.
- Exercise: Is this a matroid? Exercise: If so, is it representable?

Projective Geometries: Other Examples

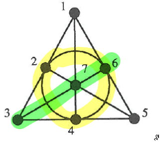
- Other examples can be more complex, consider the following two matroids (from Oxley, 2011):

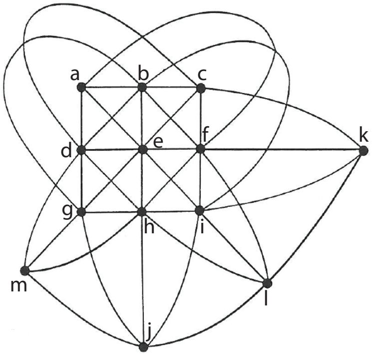




Projective Geometries: Other Examples

- Other examples can be more complex, consider the following two matroids (from Oxley, 2011):

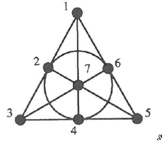


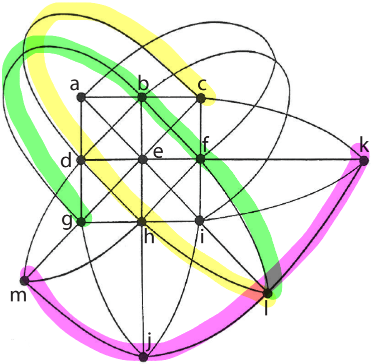


- Right: a matroid (and a 2D depiction of a geometry) over the field $\mathrm{GF}(3) = \{0, 1, 2\}$ mod 3 and is "coordinatizable" in $\mathrm{GF}(3)^3$.

Projective Geometries: Other Examples

- Other examples can be more complex, consider the following two matroids (from Oxley, 2011):





- Right: a matroid (and a 2D depiction of a geometry) over the field GF$(3) = \{0, 1, 2\}$ mod 3 and is "coordinatizable" in GF$(3)^3$.
- Hence, lines (in 2D) which are rank 2 sets may be curved; planes (in 3D) can be twisted.

Matroids, Representation and Equivalence: Summary

- Matroids with $|V| \leq 3$ are graphic.

Matroids, Representation and Equivalence: Summary

- Matroids with $|V| \leq 3$ are graphic.
- Matroids with $r(V) \leq 4$ can be geometrically represented in $\mathbb{R}^3$.

Matroids, Representation and Equivalence: Summary

- Matroids with $|V| \leq 3$ are graphic.
- Matroids with $r(V) \leq 4$ can be geometrically represented in $\mathbb{R}^3$.
- Not all matroids are linear (i.e., matric) matroids.

Matroids, Representation and Equivalence: Summary

- Matroids with $|V| \leq 3$ are graphic.
- Matroids with $r(V) \leq 4$ can be geometrically represented in $\mathbb{R}^3$.
- Not all matroids are linear (i.e., matric) matroids.
- Matroids can be seen as related to projective geometries (and are sometimes called combinatorial geometries).

Matroids, Representation and Equivalence: Summary

- Matroids with $|V| \leq 3$ are graphic.
- Matroids with $r(V) \leq 4$ can be geometrically represented in $\mathbb{R}^3$.
- Not all matroids are linear (i.e., matric) matroids.
- Matroids can be seen as related to projective geometries (and are sometimes called combinatorial geometries).
- Exists much research on different subclasses of matroids, and if/when they are contained in (or isomorphic to) each other.

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)
- Welsh, "Matroid Theory", 1975.

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)
- Welsh, "Matroid Theory", 1975.
- Oxley, "Matroid Theory", 1992 (and 2011) (perhaps best "single source" on matroids right now).

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)
- Welsh, "Matroid Theory", 1975.
- Oxley, "Matroid Theory", 1992 (and 2011) (perhaps best "single source" on matroids right now).
- Crapo & Rota, "On the Foundations of Combinatorial Theory: Combinatorial Geometries", 1970 (while this is old, it is very readable).

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)
- Welsh, "Matroid Theory", 1975.
- Oxley, "Matroid Theory", 1992 (and 2011) (perhaps best "single source" on matroids right now).
- Crapo & Rota, "On the Foundations of Combinatorial Theory: Combinatorial Geometries", 1970 (while this is old, it is very readable).
- Lawler, "Combinatorial Optimization: Networks and Matroids", 1976.

Matroid Further Reading

- "The Coming of the Matroids", William Cunningham, 2012 (a nice history)
- Welsh, "Matroid Theory", 1975.
- Oxley, "Matroid Theory", 1992 (and 2011) (perhaps best "single source" on matroids right now).
- Crapo & Rota, "On the Foundations of Combinatorial Theory: Combinatorial Geometries", 1970 (while this is old, it is very readable).
- Lawler, "Combinatorial Optimization: Networks and Matroids", 1976.
- Schrijver, "Combinatorial Optimization", 2003

The greedy algorithm

- In combinatorial optimization, the greedy algorithm is often useful as a heuristic that can work quite well in practice.

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.

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).

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 $O(n \log n)$.

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).

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 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 $w : E \to \mathbb{R}_+$.

Matroid and the greedy algorithm

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

Algorithm 1: The Matroid Greedy Algorithm

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

Matroid and the greedy algorithm

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

Algorithm 1: The Matroid Greedy Algorithm

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

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

Matroid and the greedy algorithm

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

Algorithm 1: The Matroid Greedy Algorithm

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

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

Theorem 8.4.1

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

Review

- The next slide is from Lecture 5.

Matroids by bases

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

Theorem 8.4.1 (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 8.4.1.

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

...

Matroid and the greedy algorithm

proof of Theorem 8.4.1.

- 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 8.4.1.

- 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.

...

Matroid and the greedy algorithm

proof of Theorem 8.4.1.

- 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 any another base of M with elements also ordered decreasing by weight.
- 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 8.4.1.

- 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 8.4.1.

- 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 8.4.1.

- 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 $b_i \notin A_{k-1}$ 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 8.4.1.

- 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 $A_{k-1} = \{a_1, \ldots, a_{k-1}\}$ and $B_k = \{b_1, \ldots, b_k\}$.
- Since $|A_{k-1}| < |B_k|$, there exists a $b_i \notin A_{k-1}$ 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 8.4.1.

- 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 8.4.1.

- 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 8.4.1.

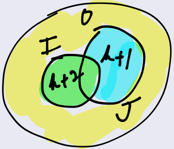
- Given an independence system $(E, \mathcal{I})$, suppose the greedy algorithm leads to an independent set of max weight for every non-negative weight function. We'll show $(E, \mathcal{I})$ is a matroid.
- Emptyset containing and down monotonicity already holds (since we've started with an independence 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 8.4.1.

- 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 $k = |I|$.



$$w(v) = \begin{cases} k+2 & \text{if } v \in I, \\ k+1 & \text{if } v \in J \setminus I, \\ 0 & \text{if } v \in E \setminus (I \cup J) \end{cases} \tag{8.1}$$

...

Matroid and the greedy algorithm

converse proof of Theorem 8.4.1.

- 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 8.4.1.

- 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

$$w(J) \geq |J|(k+1) \geq (k+1)(k+1) > k(k+2) \tag{8.2}$$

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

...

Matroid and the greedy algorithm

converse proof of Theorem 8.4.1.

- 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

$$w(J) \geq |J|(k+1) \geq (k+1)(k+1) > k(k+2) \tag{8.2}$$

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 $I \cup \{z\} \cancel{\notin} \in \mathcal{I}$, 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 8.5.1

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

$$P = \{x : Ax \leq b\} \tag{8.3}$$

Convex Polyhedra

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

Definition 8.5.1

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

$$P = \{x : Ax \leq b\} \tag{8.3}$$

- 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

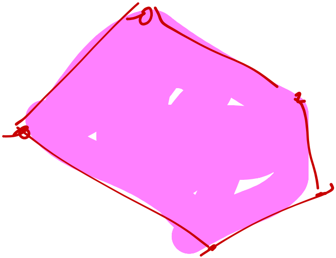
- A polytope is defined as follows

Convex Polytope

- A polytope is defined as follows

Definition 8.5.2

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 $\exists$, $x_1, x_2, \dots, x_k \in \mathcal{R}^E$ 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 8.5.2

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:

$$\text{conv}(x_1, x_2, \ldots, x_k) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{k} \lambda_i x_i : \forall i, \; \lambda_i \geq 0, \text{ and } \sum_i \lambda_i = 1 \right\} \tag{8.4}$$

Convex Polytope - key representation theorem

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

Theorem 8.5.3

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*

$$P = \{x : Ax \leq b\} \tag{8.5}$$

Convex Polytope - key representation theorem

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

Theorem 8.5.3

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*

$$P = \{x : Ax \leq b\} \tag{8.5}$$

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

Linear Programming

Theorem 8.5.4 (weak duality)

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

$$\max\{c^\intercal x | Ax \le b\} \le \min\{y^\intercal b : y \ge 0, y^\intercal A = c^\intercal\} \tag{8.6}$$

Linear Programming

Theorem 8.5.4 (weak duality)

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

$$\max\{c^\intercal x | Ax \leq b\} \leq \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{8.6}$$

Theorem 8.5.5 (strong duality)

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

$$\max\{c^\intercal x | Ax \leq b\} = \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{8.7}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{8.8}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{8.8}$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \tag{8.9}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{8.8}$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \tag{8.9}$$

$$\min\{c^\intercal x | x \geq 0, Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A \leq c^\intercal\} \tag{8.10}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{8.8}$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \tag{8.9}$$

$$\min\{c^\intercal x | x \geq 0, Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A \leq c^\intercal\} \tag{8.10}$$

$$\min\{c^\intercal x | Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A = c^\intercal\} \tag{8.11}$$

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.

Vector, modular, incidence

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

$$x(A) = \sum_{a \in A} x_a \tag{8.12}$$

Vector, modular, incidence

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

$$x(A) = \sum_{a \in A} x_a \tag{8.12}$$

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

$$\mathbf{1}_A \stackrel{\text{def}}{=} \left\{ x \in \{0,1\}^E : x_i = 1 \text{ iff } i \in A \right\} \tag{8.13}$$

equivalently,

$$\mathbf{1}_A(j) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{if } j \notin A \end{cases} \tag{8.14}$$

Matroid

Slight modification (non unit increment) that is equivalent.

Definition 8.6.1 (Matroid-II)

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

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

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

(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

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{8.15}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{ \bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\} \right\} \tag{8.15}$$

- 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

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{8.15}$$

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

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{8.16}$$

$$\left(\sum_{a \in A} x(a) \le r(A) \right)$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\bigcup_{I \in \mathcal{I}} \{\mathbf{1}_I\}\right\} \tag{8.15}$$

- 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:

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{8.16}$$

- 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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

- Clearly, for such x, $x \geq 0$.

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{8.18}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{8.18}$$

$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.19}$$

$|I_j|$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{8.18}$$

$$\leq \sum_i \lambda_i \max_{j:I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.19}$$

$$= \max_{j:I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.20}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{8.18}$$

$$\leq \sum_i \lambda_i \max_{j:I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.19}$$

$$= \max_{j:I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.20}$$

$$= r(A) \tag{8.21}$$

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

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

$$x = \sum_i \lambda_i \mathbf{1}_{I_i} \tag{8.17}$$

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

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

$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{8.18}$$

$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.19}$$

$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{8.20}$$

$$= r(A) \tag{8.21}$$

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

Matroid Polyhedron in 2D

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{8.22}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{8.23}$$

$$x_1 \leq r(\{v_1\}) \tag{8.24}$$

$$x_2 \leq r(\{v_2\}) \tag{8.25}$$

$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{8.26}$$

Matroid Polyhedron in 2D

$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{8.22}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{8.23}$$

$$x_1 \leq r(\{v_1\}) \tag{8.24}$$

$$x_2 \leq r(\{v_2\}) \tag{8.25}$$

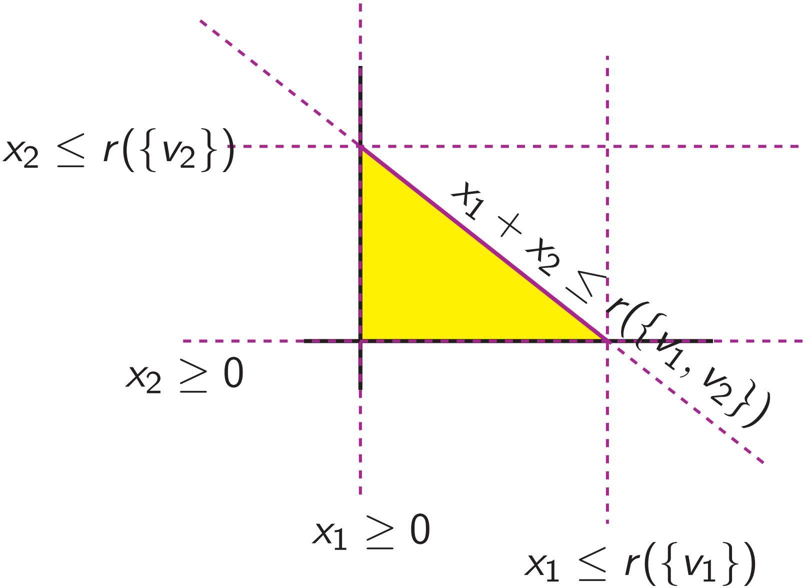
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{8.26}$$

- Because r is submodular, we have

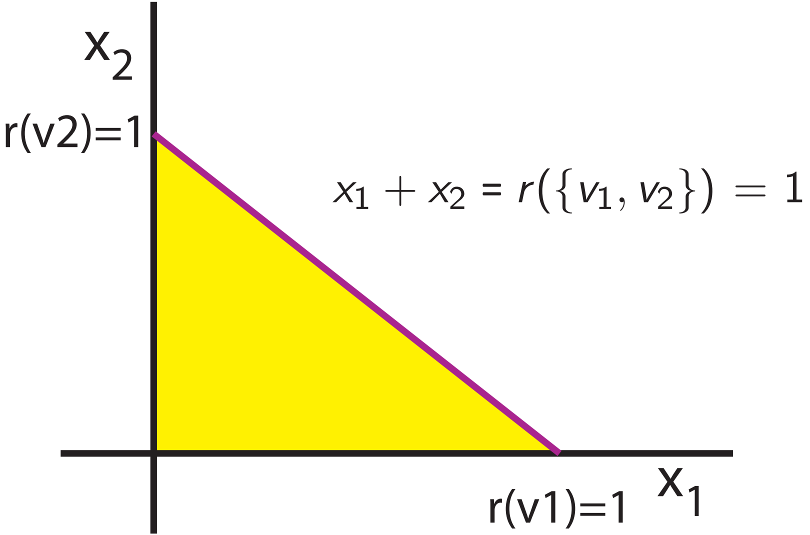
$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{8.27}$$

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

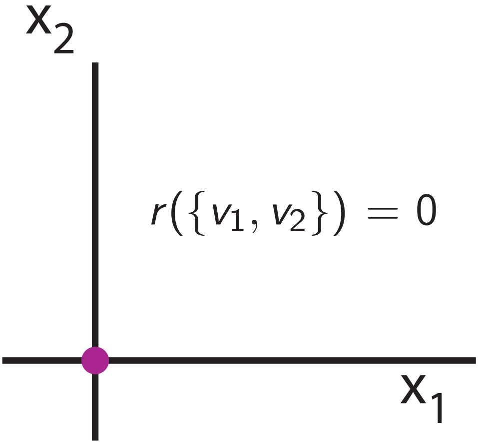
Matroid Polyhedron in 2D



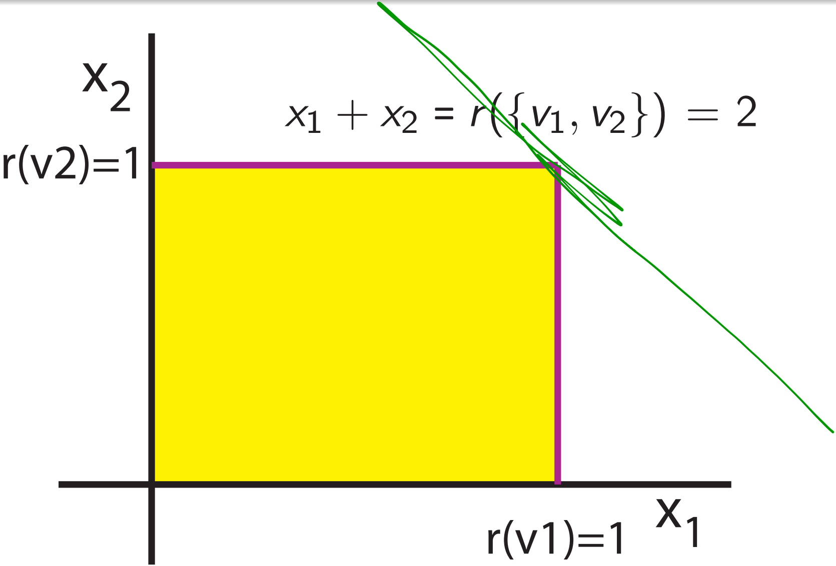
Matroid Polyhedron in 2D



Matroid Polyhedron in 2D

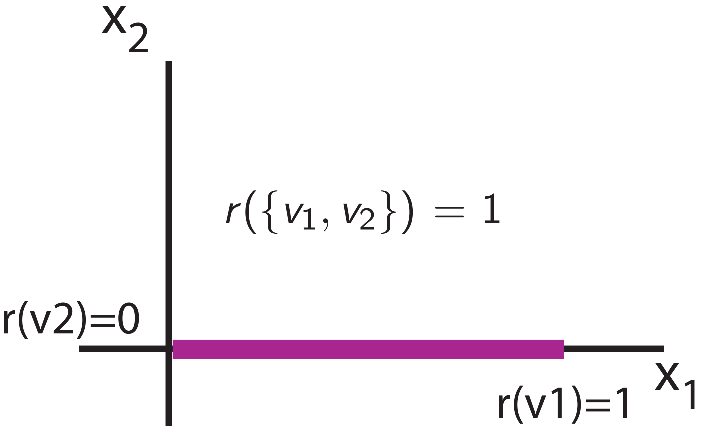


Matroid Polyhedron in 2D

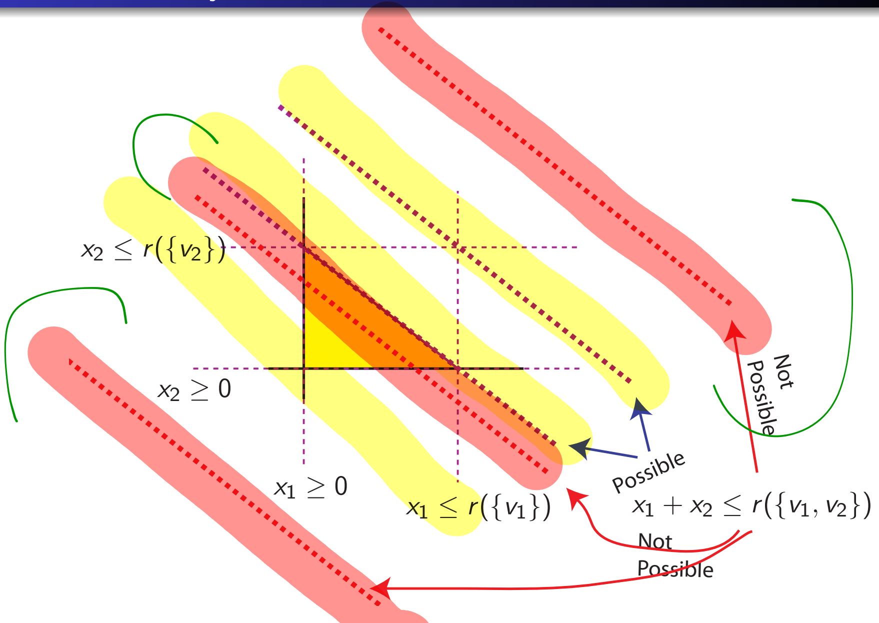


Matroid Polyhedron in 2D

And, if v2 is a loop ...



Matroid Polyhedron in 2D



Matroid Polyhedron in 3D

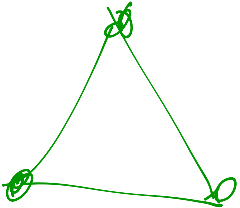
$$P_r^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\} \tag{8.28}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{8.29}$$
$$x_1 \leq r(\{v_1\}) \tag{8.30}$$
$$x_2 \leq r(\{v_2\}) \tag{8.31}$$
$$x_3 \leq r(\{v_3\}) \tag{8.32}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{8.33}$$
$$x_2 + x_3 \leq r(\{v_2, v_3\}) \tag{8.34}$$
$$x_1 + x_3 \leq r(\{v_1, v_3\}) \tag{8.35}$$
$$x_1 + x_2 + x_3 \leq r(\{v_1, v_2, v_3\}) \tag{8.36}$$

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

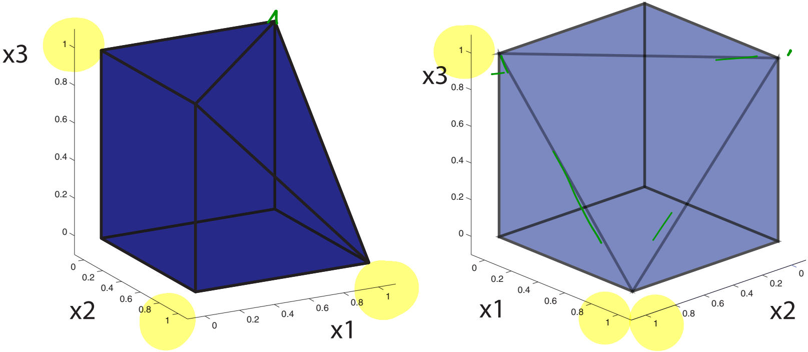
- Consider the simple cycle matroid on a graph consisting of a 3-cycle, $G = (V, E)$ with matroid $M = (E, \mathcal{I})$ where $I \in \mathcal{I}$ is a forest.
- So any set of either one or two edges is independent, and has rank equal to cardinality.

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
$(\{e_1, e_2, e_3\}, \{\emptyset, \{e_1\}, \{e_2\}, \{e_3\}, \{e_1, e_2\}, \{e_1, e_3\}, \{e_2, e_3\}\})$.

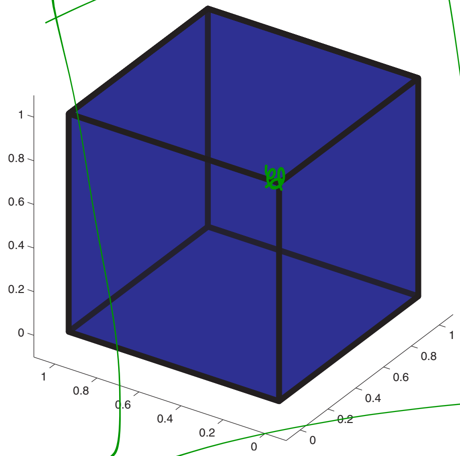


Matroid Polyhedron in 3D

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

Matroid Polyhedron in 3D

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

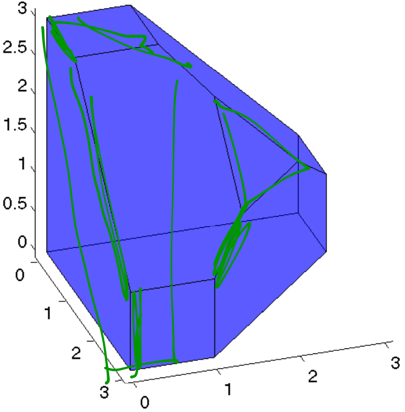


Another Polytope in 3D

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

$$P_{\text{ind. set}} = \operatorname{conv}\left\{\cup_{I\in\mathcal{I}}\{\mathbf{1}_I\}\right\}$$
$$\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{8.37}$$

Matroid Independence Polyhedron

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

$$\begin{aligned} P_{\text{ind. set}} &= \operatorname{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \end{aligned} \tag{8.37}$$

- 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

$$\begin{aligned} P_{\text{ind. set}} &= \operatorname{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \\ &\subseteq P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \end{aligned} \quad (8.37)$$

- 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 8.6.1

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{8.38}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{8.39}$$

Maximum weight independent set via weighted rank

Proof.

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

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2)\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3)\begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \tag{8.40}$$

...

Maximum weight independent set via weighted rank

Proof.

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

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2)\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3)\begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \tag{8.40}$$

- 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 $(v_1, v_2, \dots, v_n)$ such that $w(v_1) \geq w(v_2) \geq \dots \geq w(v_n)$

. . .

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 $i = 0, \dots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \dots, v_i\} \tag{8.41}$$

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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{8.41}$$

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

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{8.42}$$

...

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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{8.41}$$

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

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{8.42}$$

- 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 $(v_1, v_2, \ldots, v_n)$ 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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{8.41}$$

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

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{8.42}$$

- 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 (even a base, actually).

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \ldots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

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

$$w(I) = \sum_{v \in I} w(v) = \tag{8.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{8.45}$$

(8.46)

…

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{8.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) \tag{8.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{8.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{8.46}$$

...

Maximum weight independent set via weighted rank

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{8.43}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{8.44}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{8.45}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{8.46}$$

- 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

$$\begin{aligned} \text{maximize} \quad & w^{\intercal} x \\ \text{subject to} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{8.47}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^{\intercal} x \\ \text{subject to} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{8.47}$$

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

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{subject to} \quad & y_U \geq 0 && (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \end{aligned} \tag{8.48}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{subject to} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{8.47}$$

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

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{subject to} \quad & y_U \geq 0 && (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \mathbf{1}_U \geq w \end{aligned} \tag{8.48}$$

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

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{8.49}$$

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{8.49}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r^+ \tag{8.50}$$

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

$$\begin{array}{rll} \text{maximize} & w^\intercal x & \\ \text{s.t.} & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{array} \tag{8.49}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r^+ \tag{8.50}$$

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.,

$$\max w^\intercal x \text{ s.t. } x \in P_{\text{ind. set}} \leq \max w^\intercal x \text{ s.t. } x \in P_r^+. \tag{8.51}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \quad (8.52)$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \quad (8.53)$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \quad (8.54)$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{8.52}$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \tag{8.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{8.54}$$

- Theorem 8.6.1 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{8.55}$$

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

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{8.52}$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \tag{8.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{8.54}$$

- Theorem 8.6.1 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{8.55}$$

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

- Therefore, we also have

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \geq \alpha_{\text{min}} \tag{8.56}$$

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{8.52}$$

$$\leq \max\{w^\intercal x : x \in P_r^+\} \tag{8.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{8.54}$$

- Therefore, all the inequalities above are equalities.

Polytope equivalence

- Hence, we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{8.52}$$

$$= \max\{w^\intercal x : x \in P_r^+\} \tag{8.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\text{min}} = \min\left\{\sum_{U\subseteq V} y_U r(U) : y \geq 0, \sum_{U\subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{8.54}$$

- 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:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{8.52}$$

$$= \max\left\{w^\intercal x : x \in P_r^+\right\} \tag{8.53}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \mathbf{1}_U \geq w\right\} \tag{8.54}$$

- 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 8.6.2

$$P_r^+ = P_{\textit{ind. set}} \tag{8.57}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \tag{8.58}$$

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

$$P_{\text{ind. set}} = \operatorname{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \tag{8.58}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{8.59}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}}\{\mathbf{1}_I\}\right\} \tag{8.58}$$

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

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{8.59}$$

Theorem 8.6.3

$$P_r^+ = P_{\textit{ind. set}} \tag{8.60}$$

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 8.52, the LP problem with exponential number of constraints $\max\{w^\intercal x : x \in P_{\text{ind. set}}\}$ 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 8.52, the LP problem with exponential number of constraints $\max\{w^\intercal x : x \in P_{\text{ind. set}}\}$ 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 8.6.4

The LP problem $\max\{w^\intercal x : x \in P_{\text{ind. set}}\}$ can be solved exactly using the greedy algorithm.

Note that this LP problem has an exponential number of constraints (since $P_{\text{ind. set}}$ 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 8.52, the LP problem with exponential number of constraints $\max\{w^\intercal x : x \in P_{\text{ind. set}}\}$ 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 8.6.4

The LP problem $\max\{w^\intercal x : x \in P_{\text{ind. set}}\}$ can be solved exactly using the greedy algorithm.

Note that this LP problem has an exponential number of constraints (since $P_{\text{ind. set}}$ 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 of bases of a matroid, rather than just independent sets.

Base Polytope Equivalence

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

$$x \geq 0 \tag{8.61}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{8.62}$$

$$x(V) = r(V) \tag{8.63}$$

Base Polytope Equivalence

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

$$x \geq 0 \tag{8.61}$$

$$x(A) \leq r(A) \ \forall A \subseteq V \tag{8.62}$$

$$x(V) = r(V) \tag{8.63}$$

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

Base Polytope Equivalence

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

$$x \geq 0 \tag{8.61}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{8.62}$$

$$x(V) = r(V) \tag{8.63}$$

- 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. 8.61- 8.63 above.

Base Polytope Equivalence

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

$$x \geq 0 \tag{8.61}$$

$$x(A) \leq r(A)\ \forall A \subseteq V \tag{8.62}$$

$$x(V) = r(V) \tag{8.63}$$

- 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. 8.61- 8.63 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 8.6.5

The spanning set polytope is determined by the following equations:

$$0 \le x_e \le 1 \qquad \text{for } e \in E \tag{8.64}$$

$$x(A) \ge r(E) - r(E \setminus A) \qquad \text{for } A \subseteq E \tag{8.65}$$

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

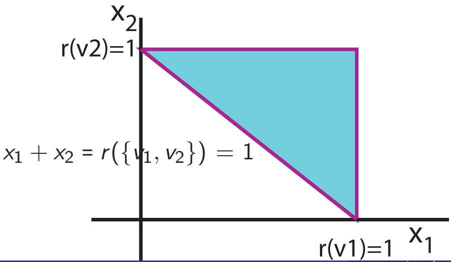
Theorem 8.6.5

The spanning set polytope is determined by the following equations:

$$0 \le x_e \le 1 \qquad \text{for } e \in E \tag{8.64}$$

$$x(A) \ge r(E) - r(E \setminus A) \qquad \text{for } A \subseteq E \tag{8.65}$$

- 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

$$x \in P_{\text{spanning}}(M) \Leftrightarrow 1 - x \in P_{\text{ind. set}}(M^*) \tag{8.66}$$

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:

$$x = \sum_i \lambda_i \mathbf{1}_{A_i} \tag{8.67}$$

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:

$$x = \sum_i \lambda_i \mathbf{1}_{A_i} \tag{8.67}$$

where A_i is spanning in M.

- Consider

$$\mathbf{1} - x = \mathbf{1}_E - x = \mathbf{1}_E - \sum_i \lambda_i \mathbf{1}_{A_i} = \sum_i \lambda_i \mathbf{1}_{E \setminus A_i}, \tag{8.68}$$

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

$$\mathbf{1} - x \geq 0 \tag{8.69}$$

$$\mathbf{1}_A - x(A) = |A| - x(A) \leq r_{M^*}(A) \text{ for } A \subseteq E \tag{8.70}$$

And we know the dual rank function is

$$r_{M^*}(A) = |A| + r_M(E \setminus A) - r_M(E) \tag{8.71}$$

...

Spanning set polytope

. . . proof continued.

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

$$\mathbf{1} - x \geq 0 \tag{8.69}$$

$$\mathbf{1}_A - x(A) = |A| - x(A) \leq r_{M^*}(A) \text{ for } A \subseteq E \tag{8.70}$$

And we know the dual rank function is

$$r_{M^*}(A) = |A| + r_M(E \setminus A) - r_M(E) \tag{8.71}$$

- giving

$$x(A) \geq r_M(E) - r_M(E \setminus A) \text{ for all } A \subseteq E \tag{8.72}$$

$\square$