EE595A – Submodular functions, their optimization and applications – Spring 2011

Prof. Jeff Bilmes

University of Washington, Seattle
Department of Electrical Engineering
Spring Quarter, 2011
http://ssli.ee.washington.edu/~bilmes/ee595a_spring_2011/

Lecture 6 - April 15th, 2011

Announcements

- Reminder, HW1 is on the web page now, at http://ssli.ee.washington.edu/~bilmes/ee595a_spring_2011/hw1.pdf
- It is due, Tuesday, April 26th, 11:45pm
- All submissions must be done electronically, via our drop box. See the link https://catalyst.uw.edu/collectit/dropbox/bilmes/14888, or look at the homework on the web page.

Matroid and the greedy algorithm

- Let $\mathcal{I}$ be a set of subsets of E that is down-closed. Consider a non-negative modular weight function $w : E \to \mathbb{R}_+$, and we want to find the $A \in \mathcal{I}$ that maximizes $w(A)$.
- Greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.

Theorem 2.1

Let $\mathcal{I}$ be a non-empty collection of subsets of a set E, down-closed (i.e., an independence system). Then the pair $(E, \mathcal{I})$ is a matroid if and only if for each weight function $w \in \mathcal{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

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 keeping 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 keeping 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 keeping 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 keeping 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 keeping 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.
- We can instead do as small as possible thus giving us a minimum weight independent set/base.

Convex Polyhedra

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

Convex Polyhedra

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

Definition 3.1

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

$$P = \{x : Ax \leq b\} \tag{1}$$

Convex Polyhedra

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

Definition 3.1

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

$$P = \{x : Ax \leq b\} \tag{1}$$

- 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

Definition 3.2

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 3.2

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 $x = \sum_i \lambda_i x_i$.

- We define the convex hull operator as follows:

$$\text{conv}(x_1, x_2, \dots, x_k) \stackrel{\text{def}}{=} \left\{ \sum_{i=1}^{k} \lambda_i x_i : \forall i, \; \lambda_i \geq 0, \text{ and } \sum_i \lambda_i = 1 \right\} \tag{2}$$

Convex Polytope - key representation theorem

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

Theorem 3.3

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*

$$P = \{x : Ax \leq b\} \tag{3}$$

Convex Polytope - key representation theorem

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

Theorem 3.3

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*

$$P = \{x : Ax \leq b\} \tag{3}$$

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

Linear Programming

Theorem 3.4 (weak duality)

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

$$\max \{c^\intercal x | Ax \leq b\} \leq \min \{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \tag{4}$$

Linear Programming

Theorem 3.4 (weak duality)

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

$$\max\{c^\intercal x | Ax \leq b\} \leq \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \quad (4)$$

Theorem 3.5 (strong duality)

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

$$\max\{c^\intercal x | Ax \leq b\} = \min\{y^\intercal b : y \geq 0, y^\intercal A = c^\intercal\} \quad (5)$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{6}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{6}$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \tag{7}$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \quad (6)$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \quad (7)$$

$$\min\{c^\intercal x | x \geq 0, Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A \leq c^\intercal\} \quad (8)$$

Linear Programming duality forms

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

$$\max\{c^\intercal x | x \geq 0, Ax \leq b\} = \min\{y^\intercal b | y \geq 0, y^\intercal A \geq c^\intercal\} \tag{6}$$

$$\max\{c^\intercal x | x \geq 0, Ax = b\} = \min\{y^\intercal b | y^\intercal A \geq c^\intercal\} \tag{7}$$

$$\min\{c^\intercal x | x \geq 0, Ax \geq b\} = \max\{y^\intercal b | y \geq 0, y^\intercal A \leq c^\intercal\} \tag{8}$$

$$\min\{c^\intercal x | Ax \geq b\} = \max\{t^\intercal b | y \geq 0, y^\intercal A = c^\intercal\} \tag{9}$$

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.*

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

$$x(A) = \sum_{a \in A} x_a \tag{10}$$

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

$$x(A) = \sum_{a \in A} x_a \tag{10}$$

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

$$\mathbf{1}_A \stackrel{\text{def}}{=} \left\{ x \in \{0,1\}^E : x_i = 1 \text{ iff } i \in A \right\} \tag{11}$$

equivalently,

$$\mathbf{1}_A(j) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{if } j \in A \\ 0 & \text{if } j \notin A \end{cases} \tag{12}$$

Matroid

Matroid definition once again.

Definition 4.1 (Matroid)

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)

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

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{13}$$

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{13}$$

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

$$P_r = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{14}$$

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
$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{13}$$
- Now take the rank function r of M, and define the following polyhedron:
$$P_r = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{14}$$
- First note, take any $x \in P_{\text{ind. set}}$, then we have that $x \in P_r$ (or $P_{\text{ind. set}} \subseteq P_r$).

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.

Matroid

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

$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$

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

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{16}$$

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \quad (15)$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \quad (16)$$
$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \quad (17)$$

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{16}$$
$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{17}$$
$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{18}$$

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{16}$$
$$\leq \sum_i \lambda_i \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{17}$$
$$= \max_{j : I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{18}$$
$$= r(A) \tag{19}$$

Matroid

- If $x \in P_{\text{ind. set}}$, then
$$x = \sum_i \lambda_i \mathbf{1}_I \tag{15}$$
for some appropriate vector $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$.
- Clearly, for such x, $x \geq 0$.
- Now, for any $A \subseteq E$,
$$x(A) = x^\intercal \mathbf{1}_A = \sum_i \lambda_i \mathbf{1}_{I_i}{}^\intercal \mathbf{1}_A \tag{16}$$
$$\leq \sum_i \lambda_i \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{17}$$
$$= \max_{j: I_j \subseteq A} \mathbf{1}_{I_j}(E) \tag{18}$$
$$= r(A) \tag{19}$$
- Thus, $x \in P_r$.

Matroid Polyhedron in 2D

$$P_r = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{20}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{21}$$

$$x_1 \leq r(\{v_1\}) \tag{22}$$

$$x_2 \leq r(\{v_2\}) \tag{23}$$

$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{24}$$

Matroid Polyhedron in 2D

$$P_r = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{20}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \tag{21}$$

$$x_1 \leq r(\{v_1\}) \tag{22}$$

$$x_2 \leq r(\{v_2\}) \tag{23}$$

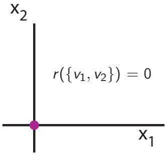
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{24}$$

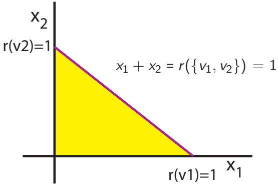
- Because r is submodular, we have

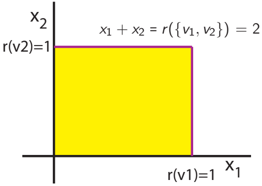
$$r(\{v_1\}) + r(\{v_2\}) \geq r(\{v_1, v_2\}) + r(\emptyset) \tag{25}$$

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

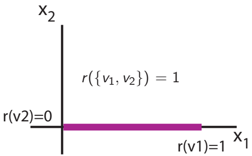
Matroid Polyhedron in 2D

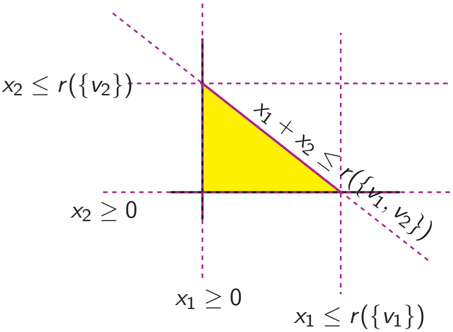




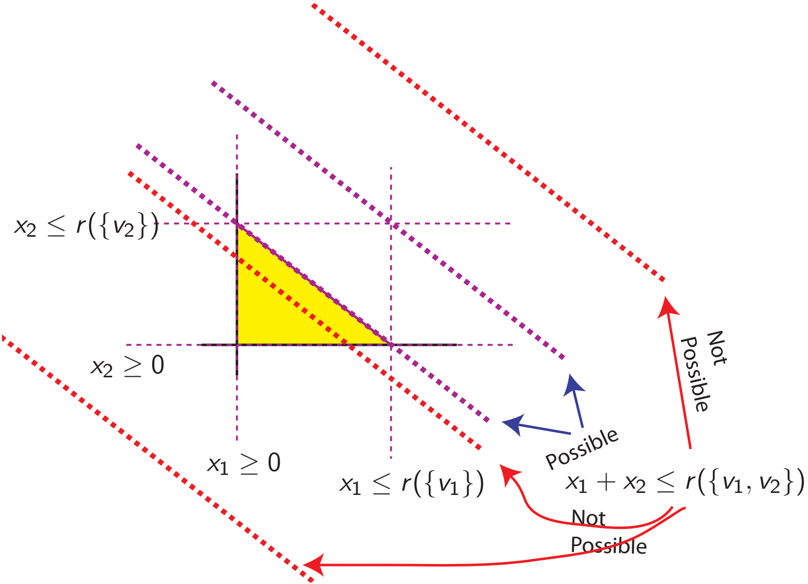


And, if v2 is a loop ...





Matroid Polyhedron in 2D



Matroid Polyhedron in 3D

$$P_r = \left\{ x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E \right\} \tag{26}$$

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

$$x_1 \geq 0 \text{ and } x_2 \geq 0 \text{ and } x_3 \geq 0 \tag{27}$$
$$x_1 \leq r(\{v_1\}) \tag{28}$$
$$x_2 \leq r(\{v_2\}) \tag{29}$$
$$x_3 \leq r(\{v_3\}) \tag{30}$$
$$x_1 + x_2 \leq r(\{v_1, v_2\}) \tag{31}$$
$$x_2 + x_3 \leq r(\{v_2, v_3\}) \tag{32}$$
$$x_1 + x_3 \leq r(\{v_1, v_3\}) \tag{33}$$
$$x_1 + x_2 + x_3 \leq r(\{v_1, v_2, v_3\}) \tag{34}$$

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

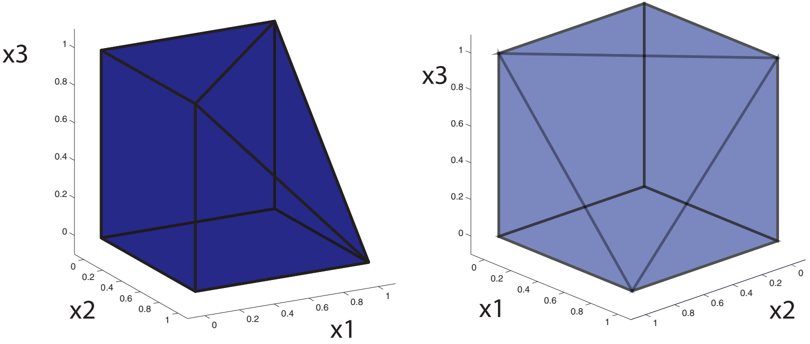
- Consider the simple cycle matroid on a graph consisting of a 3-cycle, $G = (V, E)$ with matroid $M = (E, \mathcal{I})$ where $I \in \mathcal{I}$ is a forest.
- So any set of either one or two edges is independent, and has rank equal to cardinality.

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

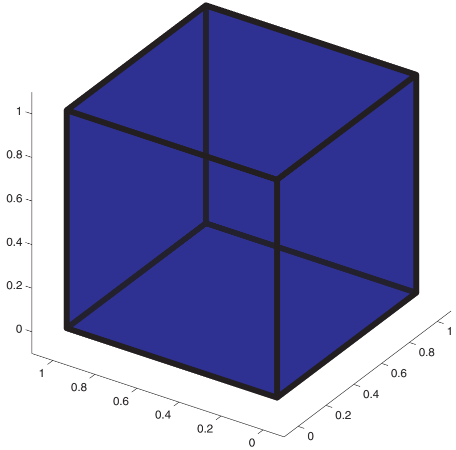


Matroid Polyhedron in 3D

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

Matroid Polyhedron in 3D

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

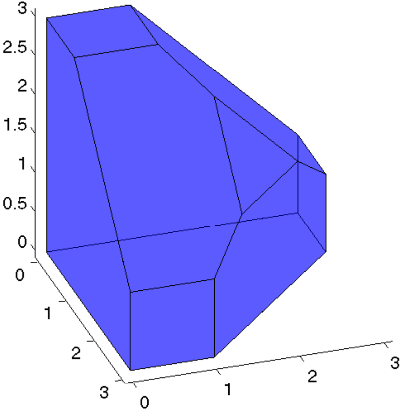


Another Polytope in 3D

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{35}$$

$$\subseteq P_r = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{36}$$

Matroid Independence Polyhedron

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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{35}$$

$$\subseteq P_r = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{36}$$

- 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

$$P_{\text{ind. set}} = \text{conv}\left\{\cup_{I \in \mathcal{I}} \mathbf{1}_I\right\} \tag{35}$$

$$\subseteq P_r = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r(A), \forall A \subseteq E\right\} \tag{36}$$

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

Matroid Independence Polyhedron

Theorem 4.2

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

$$\max\{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{37}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{38}$$

Matroid Independence Polyhedron

Proof.

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

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3) \begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n) \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \tag{39}$$

...

Matroid Independence Polyhedron

Proof.

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

$$\begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix} = (w_1 - w_2)\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + (w_2 - w_3)\begin{pmatrix} 1 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + (w_{n-1} - w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 0 \end{pmatrix} + (w_n)\begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{pmatrix} \quad (39)$$

- If we take w in decreasing order, then each coefficient of the vectors is non-negative (except possibly the last one, w_n).

...

Matroid Independence Polyhedron

Proof.

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

. . .

Matroid Independence Polyhedron

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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{40}$$

Note that

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

Matroid Independence Polyhedron

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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{40}$$

- Define the set I as

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{41}$$

...

Matroid Independence Polyhedron

Proof.

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

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \dots, v_i\} \tag{40}$$

- Define the set I as

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{41}$$

- Therefore, I is the output of the greedy algorithm. *since the items v_i are coming in decreasing order, and we only choose the ones that increase the rank, which means they don't violate independence.*

...

Matroid Independence Polyhedron

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 $i = 0, \ldots, n$

$$U_i \stackrel{\text{def}}{=} \{v_1, v_2, \ldots, v_i\} \tag{40}$$

- Define the set I as

$$I \stackrel{\text{def}}{=} \{v_i | r(U_i) > r(U_{i-1})\} \tag{41}$$

- Therefore, I is the output of the greedy algorithm.
- And therefore, I is a maximum weight independent set.

...

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

...

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

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

$$w(I) = \sum_{v \in I} w(v) = \tag{45}$$

...

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{44}$$

(45)

...

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{44}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) \tag{45}$$

. . .

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \dots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{44}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{45}$$

...

Matroid Independence Polyhedron

Proof.

- Now, we define λ_i as follows

$$\lambda_i \stackrel{\text{def}}{=} w(v_i) - w(v_{i+1}) \text{ for } i = 1, \ldots, n-1 \tag{42}$$

$$\lambda_n \stackrel{\text{def}}{=} w(v_n) \tag{43}$$

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

$$w(I) = \sum_{v \in I} w(v) = \sum_{i=1}^{n} w(v_i)\big(r(U_i) - r(U_{i-1})\big) \tag{44}$$

$$= w(v_n)r(U_n) + \sum_{i=1}^{n-1}\big(w(v_i) - w(v_{i+1})\big)r(U_i) = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{45}$$

- Since we took $v_1, v_2, \ldots$ in decreasing order, for all i, we have $\lambda_i \geq 0$

□

Linear Program LP

Consider the linear programming primal problem

$$\begin{array}{rll} \text{maximize} & w^\intercal x & \\ \text{s.t.} & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{array} \tag{46}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{46}$$

And its dual:

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{s.t.} \quad & y_U \geq 0 && (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \chi^U \geq w \end{aligned} \tag{47}$$

Linear Program LP

Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{aligned} \tag{46}$$

And its dual:

$$\begin{aligned} \text{minimize} \quad & \textstyle\sum_{U \subseteq V} y_U r(U), \\ \text{s.t.} \quad & y_U \geq 0 & (\forall U \subseteq V) \\ & \textstyle\sum_{U \subseteq V} y_U \chi^U \geq w \end{aligned} \tag{47}$$

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

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 & (v \in V) \\ & x(U) \leq r(U) & (\forall U \subseteq V) \end{aligned} \tag{48}$$

Linear Program LP

- Consider the linear programming primal problem

$$\begin{aligned} \text{maximize} \quad & w^{\intercal} x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{48}$$

- This is identical to the problem

$$\max w^{\intercal} x \text{ such that } x \in P_r \tag{49}$$

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

$$\begin{aligned} \text{maximize} \quad & w^\intercal x \\ \text{s.t.} \quad & x_v \geq 0 && (v \in V) \\ & x(U) \leq r(U) && (\forall U \subseteq V) \end{aligned} \tag{48}$$

- This is identical to the problem

$$\max w^\intercal x \text{ such that } x \in P_r \tag{49}$$

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.,

$$\max w^\intercal x \text{ s.t. } x \in P_{\text{ind. set}} \leq \max w^\intercal x \text{ s.t. } x \in P_r. \tag{50}$$

Polytope equivalence

- Now we have the following relations:

$$\max\left\{w(I) : I \in \mathcal{I}\right\} \leq \max\left\{w^\intercal x : x \in P_{\text{ind. set}}\right\} \tag{51}$$

$$\leq \max\left\{w^\intercal x : x \in P_r\right\} \tag{52}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \chi^U \geq w\right\} \tag{53}$$

Polytope equivalence

- Now we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \quad (51)$$

$$\leq \max\{w^\intercal x : x \in P_r\} \quad (52)$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \chi^U \geq w\right\} \quad (53)$$

- Theorem 4.2 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \quad (54)$$

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

Polytope equivalence

- Now we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \quad (51)$$

$$\leq \max\{w^\intercal x : x \in P_r\} \quad (52)$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \chi^U \geq w\right\} \quad (53)$$

- Theorem 4.2 states that

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \quad (54)$$

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

- Therefore, we also have

$$\max\{w(I) : I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \geq \alpha_{\min} \quad (55)$$

Polytope equivalence

- Now we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} \leq \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{51}$$

$$\leq \max\{w^\intercal x : x \in P_r\} \tag{52}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \chi^U \geq w\right\} \tag{53}$$

- Therefore, all the inequalities above are equalities.

Polytope equivalence

- Now we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^\intercal x : x \in P_{\text{ind. set}}\} \tag{51}$$

$$= \max\{w^\intercal x : x \in P_r\} \tag{52}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U \subseteq V} y_U r(U) : y \geq 0, \sum_{U \subseteq V} y_U \chi^U \geq w\right\} \tag{53}$$

- 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

- Now we have the following relations:

$$\max\{w(I) : I \in \mathcal{I}\} = \max\{w^{\intercal}x : x \in P_{\text{ind. set}}\} \tag{51}$$

$$= \max\{w^{\intercal}x : x \in P_r\} \tag{52}$$

$$\stackrel{\text{def}}{=} \alpha_{\min} = \min\left\{\sum_{U\subseteq V} y_U r(U) : y \geq 0, \sum_{U\subseteq V} y_U \chi^U \geq w\right\} \tag{53}$$

- 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 proved

Theorem 4.3

$$P_r = P_{ind.\ set} \tag{56}$$

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, looking at equations starting at Eq 51, we see that the 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, looking at equations starting at Eq 51, we see that the 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 4.4

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

- 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. By the same argument, this will be the same as the following polytope

$$x \geq 0 \tag{57}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{58}$$

$$x(V) = r(V) \tag{59}$$

Base Polytope Equivalence

- Consider convex hull of indicator vectors of bases of a matroid. By the same argument, this will be the same as the following polytope

$$x \geq 0 \tag{57}$$

$$x(A) \leq r(A) \; \forall A \subseteq V \tag{58}$$

$$x(V) = r(V) \tag{59}$$

- By essentially the same argument as above, we have shown that the convex hull of the incidence vectors of the bases of a matroid is a polytope that can be described by Eq. 57- 59 above.

Scratch Paper

Scratch Paper

Scratch Paper

Sources for Today's Lecture

Korte,Vygen-2005, Vondrak-2010, Schrijver-2003, Oxley-1992, Welsh-1973, Goemans-2010, Vazirani-2001