Submodular Functions, Optimization, and Applications to Machine Learning
— Spring Quarter, Lecture 11 —

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

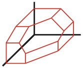
Prof. Jeff Bilmes

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

May 2nd, 2018

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

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



Cumulative Outstanding Reading

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

Announcements, Assignments, and Reminders

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

Class Road Map - EE563

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

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

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

Definition 11.2.6 (subvector)

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

Definition 11.2.7 (P-basis)

Given a compact set $P \subseteq \mathcal{R}_+^E$, for any $x \in \mathbb{R}_+^E$, a subvector y of x is called a P-basis of x if y maximal in P.
In other words, y is a P-basis of x if y is a maximal P-contained subvector of x.

Here, by y being "maximal", we mean that there exists no $z > y$ (more precisely, no $z \geq y + \epsilon \mathbf{1}_e$ for some $e \in E$ and $\epsilon > 0$) having the properties of y (the properties of y being: in P, and a subvector of x).
In still other words: y is a P-basis of x if:

1. $y \leq x$ (y is a subvector of x); and
2. $y \in P$ and $y + \epsilon \mathbf{1}_e \notin P$ for all $e \in E$ where $y(e) < x(e)$ and $\forall \epsilon > 0$ (y is maximal P-contained).

A vector form of rank

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

$$\text{rank}(A) = \max \{|I| : I \subseteq A, I \in \mathcal{I}\} = \max_{I \in \mathcal{I}} |A \cap I| \qquad (11.23)$$

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

$$\text{rank}(x) = \max \left(y(E) : y \leq x, y \in P\right) = \max_{y \in P} (x \wedge y)(E) \qquad (11.24)$$

  where $y \leq x$ is componentwise inequality ($y_i \leq x_i, \forall i$), and where $(x \wedge y) \in \mathbb{R}_+^E$ has $(x \wedge y)(i) = \min(x(i), y(i))$.
- Sometimes use $\text{rank}_P(x)$ to make P explicit.
- If $\mathcal{B}_x$ is the set of P-bases of x, than $\text{rank}(x) = \max_{y \in \mathcal{B}_x} y(E)$.
- If $x \in P$, then $\text{rank}(x) = x(E)$ (x is its own unique self P-basis).
- If $x_{\text{min}} = \min_{x \in P} x(E)$, and $x \leq x_{\text{min}}$ what then? $-\infty$?
- In general, might be hard to compute and/or have ill-defined properties. Next, we look at an object that restrains and cultivates this form of rank.

Polymatroidal polyhedron (or a "polymatroid")

Definition 11.2.1 (polymatroid)

A polymatroid is a compact set $P \subseteq \mathbb{R}_+^E$ satisfying

1. $0 \in P$
2. If $y \leq x \in P$ then $y \in P$ (called down monotone).
3. For every $x \in \mathbb{R}_+^E$, any maximal vector $y \in P$ with $y \leq x$ (i.e., any P-basis of x), has the same component sum $y(E)$

- Vectors within P (i.e., any $y \in P$) are called independent, and any vector outside of P is called dependent.
- Since all P-bases of x have the same component sum, if $\mathcal{B}_x$ is the set of P-bases of x, than $\text{rank}(x) = y(E)$ for any $y \in \mathcal{B}_x$.

Matroid and Polymatroid: side-by-side

A Matroid is:

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

A Polymatroid is:

1. a compact set $P \subseteq \mathbb{R}_+^E$
2. zero containing, $\mathbf{0} \in P$
3. down monotone, $0 \leq y \leq x \in P \Rightarrow y \in P$
4. any maximal vector y in P, bounded by another vector x, has the same vector rank (any maximal independent subvector $y \leq x$ has same sum $y(E)$).

A polymatroid function's polyhedron is a polymatroid.

Theorem 11.2.1

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

$$y^x(E) = \textit{rank}(x) \triangleq \max\left(y(E) : y \leq x, y \in P_f^+\right) = \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \quad (11.10)$$

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

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

$$\text{rank}\left(\frac{1}{\epsilon}\mathbf{1}_{E \setminus B}\right) = f(E \setminus B) = \max\left\{y(E \setminus B) : y \in P_f^+\right\} \quad (11.11)$$

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

A polymatroid is a polymatroid function's polytope

- So, when f is a polymatroid function, P_f^+ is a polymatroid.
- Is it the case that, conversely, for any polymatroid P, there is an associated polymatroidal function f such that $P = P_f^+$?

Theorem 11.2.1

For any polymatroid P (compact subset of $\mathbb{R}_+^E$, zero containing, down-monotone, and $\forall x \in \mathbb{R}_+^E$ any maximal independent subvector $y \leq x$ has same component sum $y(E) = \text{rank}(x)$), there is a polymatroid function $f : 2^E \to \mathbb{R}$ (normalized, monotone non-decreasing, submodular) such that $P = P_f^+$ where $P_f^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\}$.

Tight sets $\mathcal{D}(y)$ are closed, and max tight set $\text{sat}(y)$

Recall the definition of the set of tight sets at $y \in P_f^+$:

$$\mathcal{D}(y) \triangleq \{A : A \subseteq E,\ y(A) = f(A)\} \tag{11.19}$$

Theorem 11.2.1

For any $y \in P_f^+$, with f a polymatroid function, then $\mathcal{D}(y)$ is closed under union and intersection.

Proof.

We have already proven this as part of Theorem ?? □

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

$$\text{sat}(y) \stackrel{\text{def}}{=} \bigcup \{T : T \in \mathcal{D}(y)\} \tag{11.20}$$

Join $\vee$ and meet $\wedge$ for $x, y \in \mathbb{R}_+^E$

- For $x, y \in \mathbb{R}_+^E$, define vectors $x \wedge y \in \mathbb{R}_+^E$ and $x \vee y \in \mathbb{R}_+^E$ such that, for all $e \in E$

$$(x \vee y)(e) = \max(x(e), y(e)) \tag{11.19}$$

$$(x \wedge y)(e) = \min(x(e), y(e)) \tag{11.20}$$

Hence,

$$x \vee y \triangleq \Big(\max\big(x(e_1), y(e_1)\big), \max\big(x(e_2), y(e_2)\big), \dots, \max\big(x(e_n), y(e_n)\big)\Big)$$

and similarly

$$x \wedge y \triangleq \Big(\min\big(x(e_1), y(e_1)\big), \min\big(x(e_2), y(e_2)\big), \dots, \min\big(x(e_n), y(e_n)\big)\Big)$$

- From this, we can define things like an lattices, and other constructs.

Vector rank, rank(x), is submodular

- Recall that the matroid rank function is submodular.

Vector rank, rank(x), is submodular

- Recall that the matroid rank function is submodular.
- The vector rank function rank(x) also satisfies a form of submodularity, namely one defined on the real lattice.

Vector rank, rank(x), is submodular

- Recall that the matroid rank function is submodular.
- The vector rank function rank(x) also satisfies a form of submodularity, namely one defined on the real lattice.

Theorem 11.3.1 (vector rank and submodularity)

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

$$\text{rank}(u) + \text{rank}(v) \geq \text{rank}(u \vee v) + \text{rank}(u \wedge v) \tag{11.1}$$

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.

...

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

Proof of Theorem 11.3.1.

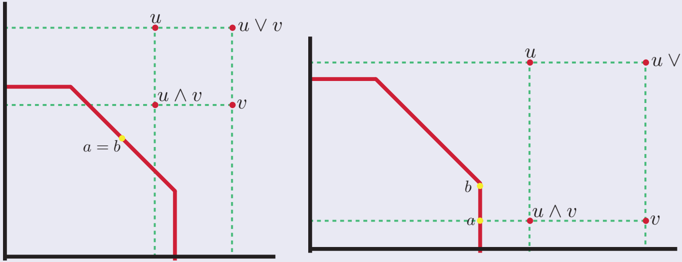
- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$

...

Vector rank rank(x) is submodular, proof

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.



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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...
- ... and since $b \leq u \vee v$, we get

$$a + b \tag{11.2}$$

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...
- ... and since $b \leq u \vee v$, we get

$$a + b = b \wedge u \wedge v + b \tag{11.2}$$

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...
- ... and since $b \leq u \vee v$, we get

$$a + b = b \wedge u \wedge v + b = b \wedge u + b \wedge v \tag{11.2}$$

...

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

Proof of Theorem 11.3.1.

- Let $a \in \mathbb{R}_+^E$ be a P-basis of $u \wedge v$, so $\text{rank}(u \wedge v) = a(E)$.
- Claim: By the polymatroid property, $\exists$ an independent $b \in P$ such that: $a \leq b \leq u \vee v$ and also such that $\text{rank}(b) = b(E) = \text{rank}(u \vee v)$, so b is a P-basis of $u \vee v$, and thus $b \leq u \vee v$.
- Given $e \in E$, if $a(e)$ is maximal due to P, then $a(e) = b(e)$ $\leq \min(u(e), v(e))$.
- If $a(e)$ is maximal due to $(u \wedge v)(e)$, then $a(e) = \min(u(e), v(e)) \leq b(e)$.
- Therefore, in either case, $a = b \wedge (u \wedge v)$ ...
- ... and since $b \leq u \vee v$, we get

$$a + b = b \wedge u \wedge v + b = b \wedge u + b \wedge v \tag{11.2}$$

How? With $b \leq u \vee v$, three cases: 1) b is minimum ($a + b = b + b$); 2) u is minimum with $b \leq v$ ($a + b = u + b$); 3) v is minimum with $b \leq u$ ($a + b = v + b$).

...

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

. . . proof of Theorem 11.3.1.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.

$\square$

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

. . . proof of Theorem 11.3.1.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.
- Hence,
$$\text{rank}(u \wedge v) + \text{rank}(u \vee v)$$

$\square$

Vector rank rank(x) is submodular, proof

. . . proof of Theorem 11.3.1.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.
- Hence,

$$\text{rank}(u \wedge v) + \text{rank}(u \vee v) = a(E) + b(E) \tag{11.3}$$

$\square$

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

...proof of Theorem 11.3.1.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.
- Hence,

$$\text{rank}(u \wedge v) + \text{rank}(u \vee v) = a(E) + b(E) \tag{11.3}$$

$$= (b \wedge u)(E) + (b \wedge v)(E) \tag{11.4}$$

$\square$

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

...proof of Theorem 11.3.1.

- b is independent, and $b \wedge u$ and $b \wedge v$ are independent subvectors of u and v respectively, so $(b \wedge u)(E) \leq \text{rank}(u)$ and $(b \wedge v)(E) \leq \text{rank}(v)$.
- Hence,

$$\begin{aligned} \text{rank}(u \wedge v) + \text{rank}(u \vee v) &= a(E) + b(E) && (11.3) \\ &= (b \wedge u)(E) + (b \wedge v)(E) && (11.4) \\ &\leq \text{rank}(u) + \text{rank}(v) && (11.5) \end{aligned}$$

$\square$

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 11.3.1 and the proof of Theorem 6.5.1 that the standard matroid rank function is submodular.

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 11.3.1 and the proof of Theorem 6.5.1 that the standard matroid rank function is submodular.
- Next, we prove Theorem 10.4.2, that any polymatroid polytope P has a polymatroid function f such that $P = P_f^+$.

A polymatroid function's polyhedron vs. a polymatroid.

- Note the remarkable similarity between the proof of Theorem 11.3.1 and the proof of Theorem 6.5.1 that the standard matroid rank function is submodular.
- Next, we prove Theorem 10.4.2, that any polymatroid polytope P has a polymatroid function f such that $P = P_f^+$.
- Given this result, we can conclude that a polymatroid is really an extremely natural polyhedral generalization of a matroid. This was all realized by Jack Edmonds in the mid 1960s (and published in 1969 in his landmark paper "Submodular Functions, Matroids, and Certain Polyhedra").

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.

…

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\max} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\max} > 0$ when P is non-empty, and $\alpha_{\max} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\max}\mathbf{1}_E)$.

. . .

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\text{max}}$.

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}}\mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \le x(E) \le \alpha_{\text{max}}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}}\mathbf{1}_A) \tag{11.6}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\text{max}}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) \tag{11.6}$$

- Then f is submodular since

$f(A) + f(B)$

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha\to\infty} \text{rank}(\alpha\mathbf{1}_E) = \text{rank}(\alpha_{\text{max}}\mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\text{max}}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}}\mathbf{1}_A) \tag{11.6}$$

- Then f is submodular since

$$f(A) + f(B) = \text{rank}(\alpha_{\text{max}}\mathbf{1}_A) + \text{rank}(\alpha_{\text{max}}\mathbf{1}_B) \tag{11.7}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\max} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\max} > 0$ when P is non-empty, and $\alpha_{\max} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\max}\mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \le x(E) \le \alpha_{\max}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\max}\mathbf{1}_A) \tag{11.6}$$

- Then f is submodular since

$$f(A) + f(B) = \text{rank}(\alpha_{\max}\mathbf{1}_A) + \text{rank}(\alpha_{\max}\mathbf{1}_B) \tag{11.7}$$

$$\ge \text{rank}(\alpha_{\max}\mathbf{1}_A \vee \alpha_{\max}\mathbf{1}_B) + \text{rank}(\alpha_{\max}\mathbf{1}_A \wedge \alpha_{\max}\mathbf{1}_B) \tag{11.8}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\max} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\max} > 0$ when P is non-empty, and $\alpha_{\max} = \lim_{\alpha\to\infty} \text{rank}(\alpha\mathbf{1}_E) = \text{rank}(\alpha_{\max}\mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\max}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\max}\mathbf{1}_A) \tag{11.6}$$

- Then f is submodular since

$$f(A) + f(B) = \text{rank}(\alpha_{\max}\mathbf{1}_A) + \text{rank}(\alpha_{\max}\mathbf{1}_B) \tag{11.7}$$
$$\geq \text{rank}(\alpha_{\max}\mathbf{1}_A \vee \alpha_{\max}\mathbf{1}_B) + \text{rank}(\alpha_{\max}\mathbf{1}_A \wedge \alpha_{\max}\mathbf{1}_B) \tag{11.8}$$
$$= \text{rank}(\alpha_{\max}\mathbf{1}_{A\cup B}) + \text{rank}(\alpha_{\max}\mathbf{1}_{A\cap B}) \tag{11.9}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- We are given a polymatroid P.
- Define $\alpha_{\text{max}} \triangleq \max\{x(E) : x \in P\}$, and note that $\alpha_{\text{max}} > 0$ when P is non-empty, and $\alpha_{\text{max}} = \lim_{\alpha\to\infty} \text{rank}(\alpha \mathbf{1}_E) = \text{rank}(\alpha_{\text{max}}\mathbf{1}_E)$.
- Hence, for any $x \in P$, and $\forall e \in E$, we have $x(e) \leq x(E) \leq \alpha_{\text{max}}$.
- Define a function $f : 2^V \to \mathbb{R}$ as, for any $A \subseteq E$,

$$f(A) \triangleq \text{rank}(\alpha_{\text{max}}\mathbf{1}_A) \tag{11.6}$$

- Then f is submodular since

$$\begin{aligned}
f(A) + f(B) &= \text{rank}(\alpha_{\text{max}}\mathbf{1}_A) + \text{rank}(\alpha_{\text{max}}\mathbf{1}_B) && (11.7)\\
&\geq \text{rank}(\alpha_{\text{max}}\mathbf{1}_A \vee \alpha_{\text{max}}\mathbf{1}_B) + \text{rank}(\alpha_{\text{max}}\mathbf{1}_A \wedge \alpha_{\text{max}}\mathbf{1}_B) && (11.8)\\
&= \text{rank}(\alpha_{\text{max}}\mathbf{1}_{A\cup B}) + \text{rank}(\alpha_{\text{max}}\mathbf{1}_{A\cap B}) && (11.9)\\
&= f(A \cup B) + f(A \cap B) && (11.10)
\end{aligned}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).

...

Proof of Theorem ??

Proof of Theorem ??.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).
- Hence, f is a polymatroid function.

. . .

Proof of Theorem ??

Proof of Theorem ??.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).
- Hence, f is a polymatroid function.
- Definition: for any $A \subseteq E$, define $x_A \in \mathbb{R}_+^E$ as

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \tag{11.11}$$

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

...

Proof of Theorem ??

Proof of Theorem ??.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).
- Hence, f is a polymatroid function.
- Definition: for any $A \subseteq E$, define $x_A \in \mathbb{R}_+^E$ as

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \tag{11.11}$$

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

- Hence $x_A(A) = x(A)$ and $x_A(E \setminus A) = 0$.

...

Proof of Theorem ??

Proof of Theorem ??.

- Moreover, we have that f is non-negative, normalized with $f(\emptyset) = 0$, and monotone non-decreasing (since rank is monotone).
- Hence, f is a polymatroid function.
- Definition: for any $A \subseteq E$, define $x_A \in \mathbb{R}_+^E$ as

$$x_A(e) = \begin{cases} x(e) & \text{if } e \in A \\ 0 & \text{else} \end{cases} \tag{11.11}$$

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

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

$$P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\} \tag{11.12}$$

...

Proof of Theorem ??

Proof of Theorem ??.

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

...

Proof of Theorem ??

Proof of Theorem ??.

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

$$
\begin{aligned}
x(A) &\leq \max \{ z(A) : z \in P, z_A \leq \alpha_{\mathsf{max}} \mathbf{1}_A \} && (11.13) \\
&= \max \{ z(A) : z \in P, z \leq \alpha_{\mathsf{max}} \mathbf{1}_A \} && (11.14) \\
&\leq \max \{ z(E) : z \in P, z \leq \alpha_{\mathsf{max}} \mathbf{1}_A \} && (11.15) \\
&= \mathsf{rank}(\alpha_{\mathsf{max}} \mathbf{1}_A) && (11.16) \\
&= f(A) && (11.17)
\end{aligned}
$$

Therefore $x \in P_f^+$.

...

Proof of Theorem ??

Proof of Theorem ??.

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

$$x(A) \leq \max \{z(A) : z \in P, z_A \leq \alpha_{\text{max}} \mathbf{1}_A\} \tag{11.13}$$
$$= \max \{z(A) : z \in P, z \leq \alpha_{\text{max}} \mathbf{1}_A\} \tag{11.14}$$
$$\leq \max \{z(E) : z \in P, z \leq \alpha_{\text{max}} \mathbf{1}_A\} \tag{11.15}$$
$$= \text{rank}(\alpha_{\text{max}} \mathbf{1}_A) \tag{11.16}$$
$$= f(A) \tag{11.17}$$

  Therefore $x \in P_f^+$.
- Hence, $P \subseteq P_f^+$.

. . .

Proof of Theorem ??

Proof of Theorem ??.

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

$$\begin{aligned} x(A) &\leq \max\left\{z(A) : z \in P, z_A \leq \alpha_{\mathsf{max}} \mathbf{1}_A\right\} && (11.13)\\ &= \max\left\{z(A) : z \in P, z \leq \alpha_{\mathsf{max}} \mathbf{1}_A\right\} && (11.14)\\ &\leq \max\left\{z(E) : z \in P, z \leq \alpha_{\mathsf{max}} \mathbf{1}_A\right\} && (11.15)\\ &= \mathsf{rank}(\alpha_{\mathsf{max}} \mathbf{1}_A) && (11.16)\\ &= f(A) && (11.17) \end{aligned}$$

  Therefore $x \in P_f^+$.
- Hence, $P \subseteq P_f^+$.
- We will next show that $P_f^+ \subseteq P$ to complete the proof.

...

Proof of Theorem ??

Proof of Theorem ??.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).

...

Proof of Theorem ??

Proof of Theorem ??.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).
- Suppose $x \notin P$.

. . .

Proof of Theorem ??

Proof of Theorem ??.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).
- Suppose $x \notin P$. Then, choose y to be a P-basis of x that maximizes the number of y elements strictly less than the corresponding x element. I.e., that maximizes $|N(y)|$, where

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{11.18}$$

...

Proof of Theorem ??

Proof of Theorem ??.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).
- Suppose $x \notin P$. Then, choose y to be a P-basis of x that maximizes the number of y elements strictly less than the corresponding x element. I.e., that maximizes $|N(y)|$, where

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{11.18}$$

- Choose w between y and x, so that

$$y \leq w \triangleq (y + x)/2 \leq x \tag{11.19}$$

so y is also a P-basis of w.

...

Proof of Theorem ??

Proof of Theorem ??.

- Let $x \in P_f^+$ be chosen arbitrarily (goal is to show that $x \in P$).
- Suppose $x \notin P$. Then, choose y to be a P-basis of x that maximizes the number of y elements strictly less than the corresponding x element. I.e., that maximizes $|N(y)|$, where

$$N(y) = \{e \in E : y(e) < x(e)\} \tag{11.18}$$

- Choose w between y and x, so that

$$y \leq w \triangleq (y + x)/2 \leq x \tag{11.19}$$

so y is also a P-basis of w.
- Hence, $\text{rank}(x) = \text{rank}(w) = y(E)$, and the set of P-bases of w are also P-bases of x.

...

Proof of Theorem ??

Proof of Theorem ??.

- Now, we have

$$y(N(y)) < w(N(y)) \leq f(N(y)) = \text{rank}(\alpha_{\text{max}} \mathbf{1}_{N(y)}) \tag{11.20}$$

the last inequality follows since $w \leq x \in P_f^+$, and $y \leq w$.

...

Proof of Theorem ??

Proof of Theorem ??.

- Now, we have

$$y(N(y)) < w(N(y)) \leq f(N(y)) = \text{rank}(\alpha_{\text{max}}\mathbf{1}_{N(y)}) \tag{11.20}$$

  the last inequality follows since $w \leq x \in P_f^+$, and $y \leq w$.
- Thus, $y \wedge x_{N(y)}$ is not a P-basis of $w \wedge x_{N(y)}$ since, over $N(y)$, it is neither tight at w nor tight at the rank (i.e., not a maximal independent subvector on $N(y)$).

...

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),
- hence $\hat{y}(e) < y(e)$ for some $e \notin N(y)$.

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),
- hence $\hat{y}(e) < y(e)$ for some $e \notin N(y)$.
- Thus, $\hat{y}$ is a base of x, which violates the maximality of $|N(y)|$.

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),
- hence $\hat{y}(e) < y(e)$ for some $e \notin N(y)$.
- Thus, $\hat{y}$ is a base of x, which violates the maximality of $|N(y)|$.
- This contradiction means that we must have had $x \in P$.

$\square$

Proof of Theorem ??

Proof of Theorem ??.

- We can extend $y \wedge x_{N(y)}$ to be a P-basis of $w \wedge x_{N(y)}$ since $y \wedge x_{N(y)} < w \wedge x_{N(y)}$.
- This P-basis, in turn, can be extended to be a P-basis $\hat{y}$ of w & x.
- Now, we have $\hat{y}(N(y)) > y(N(y))$,
- and also that $\hat{y}(E) = y(E)$ (since both are P-bases),
- hence $\hat{y}(e) < y(e)$ for some $e \notin N(y)$.
- Thus, $\hat{y}$ is a base of x, which violates the maximality of $|N(y)|$.
- This contradiction means that we must have had $x \in P$.
- Therefore, $P_f^+ = P$.

$\square$

More on polymatroids

Theorem 11.3.2

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*

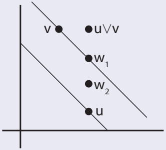
More on polymatroids

Theorem 11.3.2

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*
2. *If $u, v \in P$ (i.e., are independent) and $u(E) < v(E)$, then there exists a vector $w \in P$ such that*

$$u < w \leq u \vee v \tag{11.21}$$



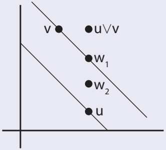
More on polymatroids

Theorem 11.3.2

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*
2. *If $u, v \in P$ (i.e., are independent) and $u(E) < v(E)$, then there exists a vector $w \in P$ such that*

$$u < w \leq u \vee v \qquad (11.21)$$



Corollary 11.3.3

The independent vectors of a polymatroid form a convex polyhedron in $\mathbb{R}_+^E$.

Review

- The next slide comes from lecture 6.

Matroids by bases

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

Theorem 11.3.3 (Matroid (by bases))

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

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

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

More on polymatroids

For any compact set P, b is a base of P if it is a maximal subvector within P. Recall the bases of matroids. In fact, we can define a polymatroid via vector bases (analogous to how a matroid can be defined via matroid bases).

Theorem 11.3.4

A polymatroid can equivalently be defined as a pair (E, P) where E is a finite ground set and $P \subseteq R_+^E$ is a compact non-empty set of independent vectors such that

1. *every subvector of an independent vector is independent (if $x \in P$ and $y \leq x$ then $y \in P$, i.e., down closed)*
2. *if b, c are bases of P and d is such that $b \wedge c < d < b$, then there exists an f, with $d \wedge c < f \leq c$ such that $d \vee f$ is a base of P*
3. *All of the bases of P have the same rank.*

Note, all three of the above are required for a polymatroid (a matroid analogy would require the equivalent of only the first two).

A word on terminology & notation

- Recall how a matroid is sometimes given as (E, r) where r is the rank function.

A word on terminology & notation

- Recall how a matroid is sometimes given as (E, r) where r is the rank function.
- We mention also that the term "polymatroid" is sometimes not used for the polytope itself, but instead but for the pair (E, f),

A word on terminology & notation

- Recall how a matroid is sometimes given as (E, r) where r is the rank function.
- We mention also that the term "polymatroid" is sometimes not used for the polytope itself, but instead but for the pair (E, f),
- But now we see that (E, f) is equivalent to a polymatroid polytope, so this is sensible.

Where are we going with this?

- Consider the right hand side of Theorem ??:
  $\min\left(x(A) + f(E \setminus A) : A \subseteq E\right)$

Where are we going with this?

- Consider the right hand side of Theorem ??:
  $\min \left(x(A) + f(E \setminus A) : A \subseteq E\right)$
- We are going to study this problem, and approaches that address it, as part of our ultimate goal which is to present strategies for submodular function minimization (that we will ultimately get to, in near future lectures).

Where are we going with this?

- Consider the right hand side of Theorem ??:
  $\min\left(x(A) + f(E \setminus A) : A \subseteq E\right)$
- We are going to study this problem, and approaches that address it, as part of our ultimate goal which is to present strategies for submodular function minimization (that we will ultimately get to, in near future lectures).
- As a bit of a hint on what's to come, recall that we can write it as: $x(E) + \min\left(f(A) - x(A) : A \subseteq E\right)$ where f is a polymatroid function.

Another Interesting Fact: Matroids from polymatroid functions

Theorem 11.3.5

Given integral polymatroid function f, let $(E, \mathcal{F})$ be a set system with ground set E and set of subsets $\mathcal{F}$ such that

$$\forall F \in \mathcal{F}, \ \ \forall \emptyset \subset S \subseteq F, |S| \leq f(S) \tag{11.22}$$

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

Proof.

Exercise □

And its rank function is Exercise.

Matroid instance of Theorem ??

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

Corollary 11.3.6

We have that:

$$\max\{y(E) : y \in P_{\textit{ind. set}}(M), y \le x\} = \min\{r_M(A) + x(E \setminus A) : A \subseteq E\} \tag{11.23}$$

where r_M is the matroid rank function of some matroid.

Review

- The next two slides come respectively from Lecture 11 and Lecture 10.

Polymatroidal polyhedron (or a "polymatroid")

Definition 11.4.1 (polymatroid)

A polymatroid is a compact set $P \subseteq \mathbb{R}_+^E$ satisfying

1. $0 \in P$
2. If $y \leq x \in P$ then $y \in P$ (called down monotone).
3. For every $x \in \mathbb{R}_+^E$, any maximal vector $y \in P$ with $y \leq x$ (i.e., any P-basis of x), has the same component sum $y(E)$

- Vectors within P (i.e., any $y \in P$) are called independent, and any vector outside of P is called dependent.
- Since all P-bases of x have the same component sum, if $\mathcal{B}_x$ is the set of P-bases of x, than $\text{rank}(x) = y(E)$ for any $y \in \mathcal{B}_x$.

Maximum weight independent set via greedy weighted rank

Theorem 11.4.5

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

$$\max \{w(I) | I \in \mathcal{I}\} = \sum_{i=1}^{n} \lambda_i r(U_i) \tag{11.8}$$

where $\lambda_i \geq 0$ satisfy

$$w = \sum_{i=1}^{n} \lambda_i \mathbf{1}_{U_i} \tag{11.9}$$

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.
- For a matroid, we saw that independence system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.
- For a matroid, we saw that independence system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.
- Stated succinctly, considering $\max\{w(I) : I \in \mathcal{I}\}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.
- For a matroid, we saw that independence system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.
- Stated succinctly, considering $\max \{w(I) : I \in \mathcal{I}\}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.
- Can we also characterize a polymatroid in this way?

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.
- For a matroid, we saw that independence system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.
- Stated succinctly, considering $\max \{w(I) : I \in \mathcal{I}\}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.
- Can we also characterize a polymatroid in this way?
- That is, if we consider $\max \left\{ wx : x \in P_f^+ \right\}$, where P_f^+ represents the "independent vectors", is it the case that P_f^+ is a polymatroid iff greedy works for this maximization?

Polymatroidal polyhedron and greedy

- Let $(E, \mathcal{I})$ be a set system and $w \in \mathbb{R}_+^E$ be a weight vector.
- Recall greedy algorithm: Set $A = \emptyset$, and repeatedly choose $y \in E \setminus A$ such that $A \cup \{y\} \in \mathcal{I}$ with $w(y)$ as large as possible, stopping when no such y exists.
- For a matroid, we saw that independence system $(E, \mathcal{I})$ is a matroid iff for each weight function $w \in \mathbb{R}_+^E$, the greedy algorithm leads to a set $I \in \mathcal{I}$ of maximum weight $w(I)$.
- Stated succinctly, considering $\max \{w(I) : I \in \mathcal{I}\}$, then $(E, \mathcal{I})$ is a matroid iff greedy works for this maximization.
- Can we also characterize a polymatroid in this way?
- That is, if we consider $\max \left\{ wx : x \in P_f^+ \right\}$, where P_f^+ represents the "independent vectors", is it the case that P_f^+ is a polymatroid iff greedy works for this maximization?
- Can we, ultimately, even relax things so that $w \in \mathbb{R}^E$?

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- Sort elements of E w.r.t. w so that, w.l.o.g. $E = (e_1, e_2, \dots, e_m)$ with $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$.

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- Sort elements of E w.r.t. w so that, w.l.o.g. $E = (e_1, e_2, \dots, e_m)$ with $w(e_1) \geq w(e_2) \geq \cdots \geq w(e_m)$.
- Let $k+1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.
  That is, we have

$$w(e_1) \geq w(e_2) \geq \cdots \geq w(e_k) > 0 \geq w(e_{k+1}) \geq \cdots \geq w(e_m) \tag{11.24}$$

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- Sort elements of E w.r.t. w so that, w.l.o.g. $E = (e_1, e_2, \dots, e_m)$ with $w(e_1) \geq w(e_2) \geq \dots \geq w(e_m)$.
- Let $k+1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.
- Next define partial accumulated sets E_i, for $i = 0 \dots m$, we have w.r.t. the above sorted order:

$$E_i \stackrel{\text{def}}{=} \{e_1, e_2, \dots e_i\} \tag{11.25}$$

(note $E_0 = \emptyset$, $f(E_0) = 0$, and $\underline{E \text{ and } E_i \text{ is always sorted w.r.t } w}$).

Polymatroidal polyhedron and greedy

- What is the greedy solution in this setting, when $w \in \mathbb{R}^E$?
- Sort elements of E w.r.t. w so that, w.l.o.g. $E = (e_1, e_2, \dots, e_m)$ with $w(e_1) \geq w(e_2) \geq \dots \geq w(e_m)$.
- Let $k+1$ be the first point (if any) at which we are non-positive, i.e., $w(e_k) > 0$ and $0 \geq w(e_{k+1})$.
- Next define partial accumulated sets E_i, for $i = 0 \dots m$, we have w.r.t. the above sorted order:

$$E_i \stackrel{\text{def}}{=} \{e_1, e_2, \dots e_i\} \tag{11.25}$$

(note $E_0 = \emptyset$, $f(E_0) = 0$, and E and E_i is always sorted w.r.t w).
- The greedy solution is the vector $x \in \mathbb{R}_+^E$ with elements defined as:

$$x(e_1) \stackrel{\text{def}}{=} f(E_1) = f(e_1) = f(e_1|E_0) = f(e_1|\emptyset) \tag{11.26}$$

$$x(e_i) \stackrel{\text{def}}{=} f(E_i) - f(E_{i-1}) = f(e_i|E_{i-1}) \text{ for } i = 2 \dots k \tag{11.27}$$

$$x(e_i) \stackrel{\text{def}}{=} 0 \text{ for } i = k+1 \dots m = |E| \tag{11.28}$$

Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$

Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$
- So $x(e_1) = f(e_1)$ and this corresponds to $w(e_1) \geq w(e_i)$ for all $i \neq 1$.

Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$
- So $x(e_1) = f(e_1)$ and this corresponds to $w(e_1) \geq w(e_i)$ for all $i \neq 1$.
- Hence, for the largest value of w (namely $w(e_1)$), we use for $x(e_1)$ the largest possible gain value of e_1 (namely $f(e_1|\emptyset) \geq f(e_1|A)$ for any $A \subseteq E \setminus \{e_1\}$).

Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$
- So $x(e_1) = f(e_1)$ and this corresponds to $w(e_1) \geq w(e_i)$ for all $i \neq 1$.
- Hence, for the largest value of w (namely $w(e_1)$), we use for $x(e_1)$ the largest possible gain value of e_1 (namely $f(e_1|\emptyset) \geq f(e_1|A)$ for any $A \subseteq E \setminus \{e_1\}$).
- For the next largest value of w (namely $w(e_2)$), we use for $x(e_2)$ the next largest gain value of e_2 (namely $f(e_2|e_1)$), while still ensuring (as we will soon see in Theorem 11.4.1) that the resulting $x \in P_f$.

Some Intuition: greedy and gain

- Note $x(e_i) = f(e_i|E_{i-1}) \leq f(e_i|E')$ for any $E' \subseteq E_{i-1}$
- So $x(e_1) = f(e_1)$ and this corresponds to $w(e_1) \geq w(e_i)$ for all $i \neq 1$.
- Hence, for the largest value of w (namely $w(e_1)$), we use for $x(e_1)$ the largest possible gain value of e_1 (namely $f(e_1|\emptyset) \geq f(e_1|A)$ for any $A \subseteq E \setminus \{e_1\}$).
- For the next largest value of w (namely $w(e_2)$), we use for $x(e_2)$ the next largest gain value of e_2 (namely $f(e_2|e_1)$), while still ensuring (as we will soon see in Theorem 11.4.1) that the resulting $x \in P_f$.
- This process continues, using the next largest possible gain of e_i for $x(e_i)$ while ensuring (as we will show) we do not leave the polytope, given the values we've already chosen for $x(e_{i'})$ for $i' < i$.

Polymatroidal polyhedron and greedy

Theorem 11.4.1

The vector $x \in \mathbb{R}_+^E$ as previously defined using the greedy algorithm maximizes wx over P_f^+, with $w \in \mathbb{R}_+^E$, if f is submodular.

Proof.

Polymatroidal polyhedron and greedy

Theorem 11.4.1

The vector $x \in \mathbb{R}_+^E$ as previously defined using the greedy algorithm maximizes wx over P_f^+, with $w \in \mathbb{R}_+^E$, if f is submodular.

Proof.

- Consider the LP strong duality equation:

$$\max(wx : x \in P_f^+) = \min\Big(\sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}_+^{2^E}, \sum_{A \subseteq E} y_A \mathbf{1}_A \geq w \Big) \tag{11.29}$$

Polymatroidal polyhedron and greedy

Theorem 11.4.1

The vector $x \in \mathbb{R}_+^E$ as previously defined using the greedy algorithm maximizes wx over P_f^+, with $w \in \mathbb{R}_+^E$, if f is submodular.

Proof.

- Consider the LP strong duality equation:

$$\max(wx : x \in P_f^+) = \min\Big(\sum_{A \subseteq E} y_A f(A) : y \in \mathbb{R}_+^{2^E}, \sum_{A \subseteq E} y_A \mathbf{1}_A \geq w\Big) \tag{11.29}$$

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

$$y_{E_i} \leftarrow w(e_i) - w(e_{i+1}) \text{ for } i = 1 \dots (m-1), \tag{11.30}$$

$$y_E \leftarrow w(e_m), \text{ and} \tag{11.31}$$

$$y_A \leftarrow 0 \text{ otherwise} \tag{11.32}$$

Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \le f(A), \forall A$).

...

Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \leq f(A), \forall A$).
- Order $A = (a_1, a_2, \ldots, a_k)$ based on order $(e_1, e_2, \ldots, e_m)$.

		a_1		a_2	a_3			a_4		a_5	$\ldots$	
e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	$\ldots$	e_m

...

Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \leq f(A), \forall A$).
- Order $A = (a_1, a_2, \dots, a_k)$ based on order $(e_1, e_2, \dots, e_m)$.

		a_1		a_2	a_3			a_4		a_5	$\dots$	
e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	$\dots$	e_m

- Define $e^{-1} : E \to \{1, \dots, m\}$ so that $e^{-1}(e_i) = i$.
  This means that with $A = \{a_1, a_2, \dots, a_k\}$, and $\forall j \leq k$

$$\{a_1, a_2, \dots, a_j\} \subseteq \left\{e_1, e_2, \dots, e_{e^{-1}(a_j)}\right\} \tag{11.33}$$

and

$$\{a_1, a_2, \dots, a_{j-1}\} \subseteq \left\{e_1, e_2, \dots, e_{e^{-1}(a_j)-1}\right\} \tag{11.34}$$

Also recall matlab notation: $a_{1:j} \equiv \{a_1, a_2, \dots, a_j\}$.
E.g., with $j = 4$ we get $e^{-1}(a_4) = 9$, and

$$\{a_1, a_2, a_3, a_4\} \subseteq \{e_1, e_2, \dots, e_9\} \tag{11.35}$$

Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \le f(A), \forall A$).
- Order $A = (a_1, a_2, \ldots, a_k)$ based on order $(e_1, e_2, \ldots, e_m)$.

		a_1		a_2	a_3			a_4		a_5	$\ldots$	
e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	$\ldots$	e_m

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

$$f(A) = \sum_{i=1}^{k} f(a_i | a_{1:i-1}) \tag{11.33}$$

$$\ge \sum_{i=1}^{k} f(a_i | e_{1:e^{-1}(a_i)-1}) \tag{11.34}$$

$$= \sum_{a \in A} f(a | e_{1:e^{-1}(a)-1}) = x(A) \tag{11.35}$$

$\ldots$

Polymatroidal polyhedron and greedy

Proof.

- We first will see that greedy $x \in P_f^+$ (that is $x(A) \leq f(A), \forall A$).
- Order $A = (a_1, a_2, \ldots, a_k)$ based on order $(e_1, e_2, \ldots, e_m)$.

		a_1		a_2	a_3			a_4		a_5	$\ldots$	
e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9	e_{10}	e_{11}	$\ldots$	e_m

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

$$f(A) = \sum_{i=1}^{k} f(a_i | a_{1:i-1}) \tag{11.33}$$

$$\geq \sum_{i=1}^{k} f(a_i | e_{1:e^{-1}(a_i)-1}) \tag{11.34}$$

$$= \sum_{a \in A} f(a | e_{1:e^{-1}(a)-1}) = x(A) \tag{11.35}$$

...

Polymatroidal polyhedron and greedy

Proof.

- y being dual feasible in Eq. 11.29 means: $y \geq 0$ and $\sum_{A \subseteq E} y_A \mathbf{1}_A \geq w$.

. . .

Polymatroidal polyhedron and greedy

Proof.

- y being dual feasible in Eq. 11.29 means: $y \geq 0$ and $\sum_{A \subseteq E} y_A \mathbf{1}_A \geq w$.
- Next, we check that y is dual feasible. Clearly, $y \geq 0$,

...

Polymatroidal polyhedron and greedy

Proof.

- y being dual feasible in Eq. 11.29 means: $y \geq 0$ and $\sum_{A \subseteq E} y_A \mathbf{1}_A \geq w$.
- Next, we check that y is dual feasible. Clearly, $y \geq 0$,
- and also, considering y component wise, for any i, we have that

$$\sum_{A: e_i \in A} y_A = \sum_{j \geq i} y_{E_j} = \sum_{j=i}^{m-1} (w(e_j) - w(e_{j+1})) + w(e_m) = w(e_i).$$

...

Polymatroidal polyhedron and greedy

Proof.

- y being dual feasible in Eq. 11.29 means: $y \geq 0$ and $\sum_{A \subseteq E} y_A \mathbf{1}_A \geq w$.
- Next, we check that y is dual feasible. Clearly, $y \geq 0$,
- and also, considering y component wise, for any i, we have that

$$\sum_{A : e_i \in A} y_A = \sum_{j \geq i} y_{E_j} = \sum_{j=i}^{m-1} (w(e_j) - w(e_{j+1})) + w(e_m) = w(e_i).$$

- Now optimality for x and y follows from strong duality, i.e.:

$$wx = \sum_{e \in E} w(e) x(e) = \sum_{i=1}^{m} w(e_i) f(e_i | E_{i-1}) = \sum_{i=1}^{m} w(e_i) \Big(f(E_i) - f(E_{i-1}) \Big)$$

$$= \sum_{i=1}^{m-1} f(E_i) \Big(w(e_i) - w(e_{i+1}) \Big) + f(E) w(e_m) = \sum_{A \subseteq E} y_A f(A)$$

...

Polymatroidal polyhedron and greedy

Proof.

- The equality in prev. Eq. follows via Abel summation:

$$
\begin{aligned}
wx &= \sum_{i=1}^{m} w_i x_i && (11.36)\\
&= \sum_{i=1}^{m} w_i \Big(f(E_i) - f(E_{i-1}) \Big) && (11.37)\\
&= \sum_{i=1}^{m} w_i f(E_i) - \sum_{i=1}^{m-1} w_{i+1} f(E_i) && (11.38)\\
&= w_m f(E_m) + \sum_{i=1}^{m-1} (w_i - w_{i+1}) f(E_i) && (11.39)
\end{aligned}
$$

$\square$

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

- When w contains negative elements, we have $x(e_i) = 0$ for $i = k+1, \ldots, m$, where k is the last positive element of w when it is sorted in decreasing order.

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

- When w contains negative elements, we have $x(e_i) = 0$ for $i = k+1, \dots, m$, where k is the last positive element of w when it is sorted in decreasing order.
- Exercise: show a modification of the previous proof that works for arbitrary $w \in \mathbb{R}^E$

Polymatroidal polyhedron and greedy

Theorem 11.4.1

Conversely, suppose P_f^+ is a polytope of form $P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if f is submodular.

Proof.

- Choose A and B arbitrarily, and then order elements of E as $(e_1, e_2, \ldots, e_m)$, with $E_i = (e_1, e_2, \ldots, e_i)$, so the following is true:

Polymatroidal polyhedron and greedy

Theorem 11.4.1

Conversely, suppose P_f^+ *is a polytope of form* $P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$*, then the greedy solution to* $\max(wx : x \in P)$ *is optimum only if* f *is submodular.*

Proof.

- Choose A and B arbitrarily, and then order elements of E as $(e_1, e_2, \ldots, e_m)$, with $E_i = (e_1, e_2, \ldots, e_i)$, so the following is true:
- For $1 \leq p \leq q \leq m$, $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p) = (A \cap B) \cup (B \setminus A)$

Polymatroidal polyhedron and greedy

Theorem 11.4.1

Conversely, suppose P_f^+ is a polytope of form $P_f^+ = \left\{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\right\}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if f is submodular.

Proof.

- Choose A and B arbitrarily, and then order elements of E as $(e_1, e_2, \ldots, e_m)$, with $E_i = (e_1, e_2, \ldots, e_i)$, so the following is true:
- For $1 \leq p \leq q \leq m$, $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p) = (A \cap B) \cup (B \setminus A)$
- Note, then we have $A \cap B = \{e_1, \ldots, e_k\} = E_k$, and $A \cup B = E_q$.

Polymatroidal polyhedron and greedy

Theorem 11.4.1

Conversely, suppose P_f^+ is a polytope of form $P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if f is submodular.

Proof.

- Choose A and B arbitrarily, and then order elements of E as $(e_1, e_2, \ldots, e_m)$, with $E_i = (e_1, e_2, \ldots, e_i)$, so the following is true:
- For $1 \leq p \leq q \leq m$, $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p) = (A \cap B) \cup (B \setminus A)$
- Note, then we have $A \cap B = \{e_1, \ldots, e_k\} = E_k$, and $A \cup B = E_q$.
- Define $w \in \{0, 1\}^m$ as:

$$w \stackrel{\text{def}}{=} \sum_{i=1}^{q} \mathbf{1}_{e_i} = \mathbf{1}_{A \cup B} \tag{11.40}$$

Polymatroidal polyhedron and greedy

Theorem 11.4.1

Conversely, suppose P_f^+ is a polytope of form $P_f^+ = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$, then the greedy solution to $\max(wx : x \in P)$ is optimum only if f is submodular.

Proof.

- Choose A and B arbitrarily, and then order elements of E as $(e_1, e_2, \ldots, e_m)$, with $E_i = (e_1, e_2, \ldots, e_i)$, so the following is true:
- For $1 \leq p \leq q \leq m$, $A = \{e_1, e_2, \ldots, e_k, e_{k+1}, \ldots, e_p\} = E_p$ and $B = \{e_1, e_2, \ldots, e_k, e_{p+1}, \ldots, e_q\} = E_k \cup (E_q \setminus E_p) = (A \cap B) \cup (B \setminus A)$
- Note, then we have $A \cap B = \{e_1, \ldots, e_k\} = E_k$, and $A \cup B = E_q$.
- Define $w \in \{0,1\}^m$ as:

$$w \stackrel{\text{def}}{=} \sum_{i=1}^{q} \mathbf{1}_{e_i} = \mathbf{1}_{A \cup B} \tag{11.40}$$

- Suppose optimum solution x is given by the greedy procedure.

Polymatroidal polyhedron and greedy

Proof.

- Then

$$\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k}(f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B) \tag{11.41}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Then

$$\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k} (f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B) \tag{11.41}$$

- and

$$\sum_{i=1}^{p} x_i = f(E_1) + \sum_{i=2}^{p} (f(E_i) - f(E_{i-1})) = f(E_p) = f(A) \tag{11.42}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Then

$$\sum_{i=1}^{k} x_i = f(E_1) + \sum_{i=2}^{k}(f(E_i) - f(E_{i-1})) = f(E_k) = f(A \cap B) \tag{11.41}$$

- and

$$\sum_{i=1}^{p} x_i = f(E_1) + \sum_{i=2}^{p}(f(E_i) - f(E_{i-1})) = f(E_p) = f(A) \tag{11.42}$$

- and

$$\sum_{i=1}^{q} x_i = f(E_1) + \sum_{i=2}^{q}(f(E_i) - f(E_{i-1})) = f(E_q) = f(A \cup B) \tag{11.43}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Thus, we have

$$x(B) = \sum_{i \in 1,\dots,k,p+1,\dots,q} x_i = \sum_{i : e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \tag{11.44}$$

...

Polymatroidal polyhedron and greedy

Proof.

- Thus, we have

$$x(B) = \sum_{i \in 1,\ldots,k,p+1,\ldots,q} x_i = \sum_{i:e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \tag{11.44}$$

- But given that the greedy algorithm gives the optimal solution to $\max(wx : x \in P_f^+)$, we have that $x \in P_f^+$ and thus $x(B) \leq f(B)$.

...

Polymatroidal polyhedron and greedy

Proof.

- Thus, we have

$$x(B) = \sum_{i \in 1,\ldots,k,p+1,\ldots,q} x_i = \sum_{i:e_i \in B} x_i = f(A \cup B) + f(A \cap B) - f(A) \tag{11.44}$$

- But given that the greedy algorithm gives the optimal solution to $\max(wx : x \in P_f^+)$, we have that $x \in P_f^+$ and thus $x(B) \leq f(B)$.
- Thus,

$$x(B) = f(A \cup B) + f(A \cap B) - f(A) = \sum_{i:e_i \in B} x_i \leq f(B) \tag{11.45}$$

ensuring the submodularity of f, since A and B are arbitrary.

□

Review from Lecture 8

- The next slide comes from lecture 8.

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 11.4.8

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

Polymatroidal polyhedron and greedy

- Thus, restating the above results into a single complete theorem, we have a result very similar to what we saw for matroids (i.e., Theorem 9.4.1)

Theorem 11.4.1

If $f : 2^E \to \mathbb{R}_+$ is given, and P is a polytope in $\mathbb{R}_+^E$ of the form $P = \{x \in \mathbb{R}_+^E : x(A) \leq f(A), \forall A \subseteq E\}$, then the greedy solution to the problem $\max(w^\intercal x : x \in P)$ is $\forall w$ optimum iff f is monotone non-decreasing submodular (i.e., iff P is a polymatroid).