Submodular Functions, Optimization, and Applications to Machine Learning

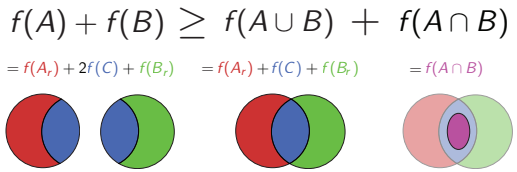
— Fall Quarter, Lecture 18 —

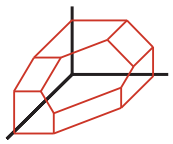
http://www.ee.washington.edu/people/faculty/bilmes/classes/ee563_spring_2018/

Prof. Jeff Bilmes

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

Dec 2nd, 2020





Class Road Map - EE563

- L1(9/30): Motivation, Applications, Definitions, Properties
- L2(10/5): Sums concave(modular), uses (diversity/costs, feature selection), information theory
- L3(10/7): Monge, More Definitions, Graph and Combinatorial Examples,
- L4(10/12): Graph & Combinatorial Examples, Matrix Rank, Properties, Other Defs, Independence
- L5(10/14): Properties, Defs of Submodularity, Independence
- L6(10/19): Matroids, Matroid Examples, Matroid Rank,
- L7(10/21): Matroid Rank, More on Partition Matroid, Laminar Matroids, System of Distinct Reps, Transversals
- L8(10/26): Transversal Matroid, Matroid and representation, Dual Matroid
- L9(10/28): Other Matroid Properties, Combinatorial Geometries, Matroid and Greedy, Polyhedra, Matroid Polytopes
- L10(11/2): Matroid Polytopes, Matroids $\rightarrow$ Polymatroids
- L11(11/4): Matroids $\rightarrow$ Polymatroids, Polymatroids
- L12(11/9): Polymatroids, Polymatroids and Greedy
- L–(11/11): Veterans Day, Holiday
- L13(11/16): Polymatroids and Greedy, Possible Polytopes, Extreme Points, Cardinality Constrained Maximization
- L14(11/18): Cardinality Constrained Maximization, Curvature
- L15(11/23): Curvature, Submodular Max w. Other Constraints, Start Cont. Extensions
- L16(11/25): Submodular Max w. Other Constraints, Cont. Extensions, Lovász extension
- L17(11/30): Choquet Integration, Non-linear Measure/Aggregation, Definitions/Properties, Examples.
- L18(12/2): Multilinear Extension, Submodular Max/polyhedral, Most Violated Ineq., Matroids Closure/Sat
- L19(12/7):
- L20(12/9):
- L21(12/14): final meeting (presentations) maximization.

Last day of instruction, Fri. Dec 11th. Finals Week: Dec 12-18, 2020

Rest of class

- Homework 4 posted, due Thursday Dec 17th, 2020, 11:55pm.
- Final project paper proposal, due Sunday Dec 6th, 11:59pm.
- Final project 4-page paper and presentation slides, due Sunday Dec 13th, 11:59pm.
- Final project presentation, Monday Dec 14th, starting at 10:30am.
- Final project: Read and present a recent (past 5 years) paper on submodular/supermodular optimization. Paper should have both a theoretical and practical component. What is due: (1) 4-page paper summary, and (2) 10 minute presentation about the paper, will be giving presentations on Monday 12/14/2020. You must choose your paper before the 14th (this will be HW5), and you must turn in your slides and 4-page paper (this will be HW6).
- Recall, grades will be based on a combination of a final project (40%) and the four homeworks (60%).

One slide review of convex closure/L.E./CI

- convex closure $\check{f}(x) = \min_{p \in \triangle^n(x)} E_{S \sim p}[f(S)]$, where where $\triangle^n(x) = \left\{ p \in \mathbb{R}^{2^n} : \sum_{S \subseteq V} p_S = 1,\ p_S \geq 0 \forall S \subseteq V,\ \& \ \sum_{S \subseteq V} p_S \mathbf{1}_S = x \right\}$
- "Edmonds" extension $\breve{f}(w) = \max(wx : x \in B_f)$
- Lovász extension $f_{\mathsf{LE}}(w) = \sum_{i=1}^m \lambda_i f(E_i)$, with λ_i such that $w = \sum_{i=1}^m \lambda_i \mathbf{1}_{E_i}$
- $\tilde{f}(w) = \max_{\sigma \in \Pi_{[m]}} w^\intercal c^\sigma$, $\Pi_{[m]}$ set of $m!$ permutations of $[m]$, $\sigma \in \Pi_{[m]}$ a permutation, c^σ vector with $c_i^\sigma = f(E_{\sigma_i}) - f(E_{\sigma_{i-1}})$, $E_{\sigma_i} = \{e_{\sigma_1}, e_{\sigma_2}, \dots, e_{\sigma_i}\}$.
- Choquet integral $C_f(w) = \sum_{i=1}^m (w_{e_i} - w_{e_{i+1}}) f(E_i)$
- $\tilde{f}(w) = \int_{-\infty}^{+\infty} \hat{f}(\alpha) d\alpha$, $\hat{f}(\alpha) = \begin{cases} f(\{w \geq \alpha\}) & \text{if } \alpha \geq 0 \\ f(\{w \geq \alpha\}) - f(E) & \text{if } \alpha < 0 \end{cases}$
- All the same when f is submodular.

Concave closure

- The concave closure is defined as:

$$\hat{f}(x) = \max_{p \in \triangle^n(x)} \sum_{S \subseteq V} p_S f(S) \tag{18.1}$$

where $\triangle^n(x) =$
$\left\{ p \in \mathbb{R}^{2^n} : \sum_{S \subseteq V} p_S = 1, p_S \geq 0 \forall S \subseteq V, \ \& \ \sum_{S \subseteq V} p_S \mathbf{1}_S = x \right\}$

Concave closure

- The concave closure is defined as:

$$\hat{f}(x) = \max_{p \in \triangle^n(x)} \sum_{S \subseteq V} p_S f(S) \tag{18.1}$$

where $\triangle^n(x) =$
$\left\{ p \in \mathbb{R}^{2^n} : \sum_{S \subseteq V} p_S = 1, p_S \geq 0 \forall S \subseteq V, \ \& \ \sum_{S \subseteq V} p_S \mathbf{1}_S = x \right\}$

- This is tight at the hypercube vertices, concave, and the concave envelope for the dual reasons as the convex closure.

Concave closure

- The concave closure is defined as:

$$\hat{f}(x) = \max_{p \in \triangle^n(x)} \sum_{S \subseteq V} p_S f(S) \tag{18.1}$$

where $\triangle^n(x) = \left\{ p \in \mathbb{R}^{2^n} : \sum_{S \subseteq V} p_S = 1, p_S \geq 0 \forall S \subseteq V, \ \& \ \sum_{S \subseteq V} p_S \mathbf{1}_S = x \right\}$

- This is tight at the hypercube vertices, concave, and the concave envelope for the dual reasons as the convex closure.
- Unlike the convex extension, the concave closure is defined by the Lovász extension iff f is a supermodular function.

Concave closure

- The concave closure is defined as:

$$\hat{f}(x) = \max_{p \in \triangle^n(x)} \sum_{S \subseteq V} p_S f(S) \tag{18.1}$$

where $\triangle^n(x) = \left\{ p \in \mathbb{R}^{2^n} : \sum_{S \subseteq V} p_S = 1, p_S \geq 0 \forall S \subseteq V, \ \& \ \sum_{S \subseteq V} p_S \mathbf{1}_S = x \right\}$

- This is tight at the hypercube vertices, concave, and the concave envelope for the dual reasons as the convex closure.
- Unlike the convex extension, the concave closure is defined by the Lovász extension iff f is a supermodular function.
- When f is submodular, even evaluating $\hat{f}$ is NP-hard (rough intuition: submodular maximization is NP-hard (reduction to set cover), if we could evaluate $\hat{f}$ in poly time, we can maximize concave function to solve submodular maximization in poly time).

Multilinear extension

- Rather than the concave closure, multi-linear extension is used as a surrogate. For $x \in [0,1]^V = [0,1]^{[n]}$

$$\tilde{f}(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \in V \setminus S} (1 - x_i) = E_{S \sim x}[f(S)] \tag{18.2}$$

$P(i \in S) = x_i$

Multilinear extension

- Rather than the concave closure, multi-linear extension is used as a surrogate. For $x \in [0,1]^V = [0,1]^{[n]}$

$$\tilde{f}(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \in V \setminus S} (1 - x_i) = E_{S \sim x}[f(S)] \tag{18.2}$$

- Can be viewed as expected value of $f(S)$ where S is a random set distributed via x, so $\Pr(v \in S) = x_v$ and is independent of $\Pr(u \in S) = x_u$, $v \neq u$.

Multilinear extension

- Rather than the concave closure, multi-linear extension is used as a surrogate. For $x \in [0,1]^V = [0,1]^{[n]}$

$$\tilde{f}(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \in V \setminus S} (1 - x_i) = E_{S \sim x}[f(S)] \tag{18.2}$$

- Can be viewed as expected value of $f(S)$ where S is a random set distributed via x, so $\Pr(v \in S) = x_v$ and is independent of $\Pr(u \in S) = x_u$, $v \neq u$.
- This is tight at the hypercube vertices (immediate, since $\tilde{f}(\mathbf{1}_A)$ yields only one term in the sum non-zero, namely the one where $S = A$).

$\tilde{f}(1_A) = f(A)$

Multilinear extension

- Rather than the concave closure, multi-linear extension is used as a surrogate. For $x \in [0,1]^V = [0,1]^{[n]}$

$$\tilde{f}(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \in V \setminus S} (1 - x_i) = E_{S \sim x}[f(S)] \tag{18.2}$$

- Can be viewed as expected value of $f(S)$ where S is a random set distributed via x, so $\Pr(v \in S) = x_v$ and is independent of $\Pr(u \in S) = x_u$, $v \neq u$.
- This is tight at the hypercube vertices (immediate, since $f(\mathbf{1}_A)$ yields only one term in the sum non-zero, namely the one where $S = A$).
- Why called multilinear (multi-linear) extension? It is linear in each of its arguments (i.e., $\tilde{f}(x_1, x_2, \dots, \alpha x_k + \beta x'_k, \dots, x_n) = \alpha \tilde{f}(x_1, x_2, \dots, x_k, \dots, x_n) + \beta \tilde{f}(x_1, x_2, \dots, x'_k, \dots, x_n)$

Multilinear extension

- Rather than the concave closure, multi-linear extension is used as a surrogate. For $x \in [0,1]^V = [0,1]^{[n]}$

$$\tilde{f}(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{i \in V \setminus S} (1 - x_i) = E_{S \sim x}[f(S)] \tag{18.2}$$

ε $O(1/\varepsilon^2)$

- Can be viewed as expected value of $f(S)$ where S is a random set distributed via x, so $\Pr(v \in S) = x_v$ and is independent of $\Pr(u \in S) = x_u$, $v \neq u$.
- This is tight at the hypercube vertices (immediate, since $f(\mathbf{1}_A)$ yields only one term in the sum non-zero, namely the one where $S = A$).
- Why called multilinear (multi-linear) extension? It is linear in each of its arguments (i.e., $\tilde{f}(x_1, x_2, \ldots, \alpha x_k + \beta x'_k, \ldots, x_n) = \alpha \tilde{f}(x_1, x_2, \ldots, x_k, \ldots, x_n) + \beta \tilde{f}(x_1, x_2, \ldots, x'_k, \ldots, x_n)$
- This is unfortunately not concave. However there are some useful properties.

Multilinear extension

Lemma 18.3.1

Let $\tilde{f}(x)$ be the multilinear extension of a set function $f : 2^V \to \mathbb{R}$. Then:

- *If f is monotone non-decreasing, then $\frac{\partial \tilde{f}}{\partial x_v} \geq 0$ for all $v \in V$ within $[0,1]^V$ (i.e., $\tilde{f}$ is also monotone non-decreasing).*
- *If f is submodular, then $\tilde{f}$ has an antitone supergradient, i.e., $\frac{\partial^2 \tilde{f}}{\partial x_i \partial x_j} \leq 0$ for all $i, j \in V$ within $[0,1]^V$.*

Proof.

- First part (monotonicity). Choose $x \in [0,1]^V$ and let $S \sim x$ be random where x is treated as a distribution (so elements v is chosen with probability x_v independently of any other element).

...

Multilinear extension

. . . proof continued.

- Since $\tilde{f}$ is multilinear, derivative is a simple difference when only one argument varies, i.e.,

$$\frac{\partial \tilde{f}}{\partial x_v} = \tilde{f}(x_1, x_2, \dots, x_{v-1}, 1, x_{v+1}, \dots, x_n) \tag{18.3}$$

$$- \tilde{f}(x_1, x_2, \dots, x_{v_1}, 0, x_{v+1}, \dots, x_n) \tag{18.4}$$

$$= E_{S \sim x}[f(S + v)] - E_{S \sim x}[f(S - v)] \tag{18.5}$$

$$\geq 0 \tag{18.6}$$

where the final part follows due to monotonicity of each argument, i.e., $f(S + i) \geq f(S - i)$ for any S and $i \in V$.

Side note: this means $\nabla \tilde{f}(x) = (\partial \tilde{f}/\partial x_{v_1}, \partial \tilde{f}/\partial x_{v_2}, \dots, \partial \tilde{f}/\partial x_{v_n})^\intercal$ has an easy expression.

$\square$

Multilinear extension

. . . proof continued.

- Second part of proof (antitone supergradient) also relies on simple consequence of multilinearity, namely multilinearity of the derivative as well. In this case

$$\frac{\partial^2 \tilde{f}}{\partial x_j \partial x_i} = \frac{\partial \tilde{f}}{\partial x_j}(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n) \tag{18.7}$$

$$- \frac{\partial \tilde{f}}{\partial x_j}(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) \tag{18.8}$$

$$= E_{S \sim x}[f(S + i + j) - f(S + i - j)] \tag{18.9}$$

$$- E_{S \sim x}[f(S - i + j) - f(S - i - j)] \tag{18.10}$$

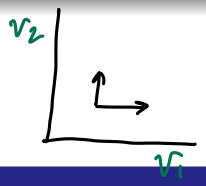
$$\leq 0 \tag{18.11}$$

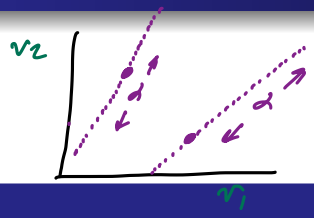
since by submodularity, we have

$$f(S + i - j) + f(S - i + j) \geq f(S + i + j) + f(S - i - j) \tag{18.12}$$

$\square$

Multilinear extension: some properties

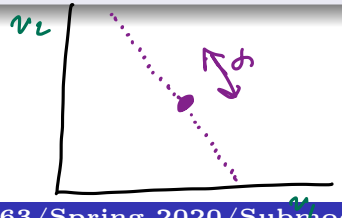




Corollary 18.3.2

let f be a function and $\tilde{f}$ its multilinear extension on $[0,1]^V$.

- *if f is monotone non-decreasing then $\tilde{f}$ is non-decreasing along any strictly non-negative direction (i.e., $\tilde{f}(x) \leq \tilde{f}(y)$ whenever $x \leq y$, or $\tilde{f}(x) \leq \tilde{f}(x + \epsilon \mathbf{1}_v)$ for any $v \in V$ and any $\epsilon \geq 0$.*
- *If f is submodular, then $\tilde{f}$ is concave along any non-negative direction (i.e., the function $g(\alpha) = \tilde{f}(x + \alpha z)$ is 1-D concave in α for any $z \in \mathbb{R}_+$).*
- *If f is submodular than $\tilde{f}$ is convex along any diagonal direction (i.e., the function $g(\alpha) = \tilde{f}(x + \alpha(\mathbf{1}_v - \mathbf{1}_u))$ is 1-D convex in α for any $u \neq v$.*



Review from lecture 12

The next two slide come from lecture 12.

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 $e \in E$

$$(x \vee y)(e) = \max(x(e), y(e)) \tag{18.1}$$

$$(x \wedge y)(e) = \min(x(e), y(e)) \tag{18.2}$$

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 $r(A) = \max(|I| : I \subseteq A : I \in \mathcal{I})$ is submodular.
- The vector rank function rank$(x) = \max\left(y(E) : y \leq x, y \in P\right)$ also satisfies a form of submodularity, namely one defined on the real lattice.

Theorem 18.3.1 (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 $u, v \in \mathbb{R}_+^E$

$$\textit{rank}(u) + \textit{rank}(v) \geq \textit{rank}(u \vee v) + \textit{rank}(u \wedge v) \tag{18.1}$$

- Note what happens when $u, v \in \{0,1\}^E \subseteq \mathbb{R}_+^E$.

Submodularity on the real lattice

- In general, we can define "submodularity" on the real lattice.

Submodularity on the real lattice

- In general, we can define "submodularity" on the real lattice.

Definition 18.3.3 (submodularity)

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be submodular (on the real lattice) if for all $x, y \in \mathbb{R}^n$ we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{18.13}$$

where $x \vee y$ (resp. $x \wedge y$) is the element-wise max (resp. min) of x and y.

Submodularity on the real lattice

- In general, we can define "submodularity" on the real lattice.

Definition 18.3.3 (submodularity)

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be submodular (on the real lattice) if for all $x, y \in \mathbb{R}^n$ we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{18.13}$$

where $x \vee y$ (resp. $x \wedge y$) is the element-wise max (resp. min) of x and y.

- When f is twice differentiable, this condition is identical to the off-diagonal elements of the Hessian of f being non-positive, i.e.,

$$\forall i \neq j, \forall x \in \mathbb{R}^n, \quad \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \leq 0 \tag{18.14}$$

See Topkis "Supermodularity and Complementarity", 1998. Note, non-positives along the diagonal are not required.

Submodularity on the real lattice

- In general, we can define "submodularity" on the real lattice.

Definition 18.3.3 (submodularity)

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be submodular (on the real lattice) if for all $x, y \in \mathbb{R}^n$ we have

$$f(x) + f(y) \geq f(x \vee y) + f(x \wedge y) \tag{18.13}$$

where $x \vee y$ (resp. $x \wedge y$) is the element-wise max (resp. min) of x and y.

- When f is twice differentiable, this condition is identical to the off-diagonal elements of the Hessian of f being non-positive, i.e.,

$$\forall i \neq j, \forall x \in \mathbb{R}^n, \quad \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \leq 0 \tag{18.14}$$

  See Topkis "Supermodularity and Complementarity", 1998. Note, non-positives along the diagonal are not required.
- This is neither the same condition as convexity nor concavity.

From Lecture 4

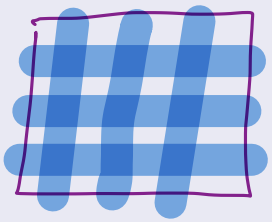
Recall quadratic structures from Lecture 4 (next frame).

Submodularity, Quadratic Structures, and Cuts

Lemma 18.3.1

Let $\mathbf{M} \in \mathbb{R}^{n \times n}$ *be a symmetric matrix and* $m \in \mathbb{R}^n$ *be a vector. Then* $f : 2^V \to \mathbb{R}$ *defined as*

$m(X) + \sum_{i,j \in X} M_{i,j}$



$$f(X) = m^\intercal \mathbf{1}_X + \frac{1}{2}\mathbf{1}_X^\intercal \mathbf{M} \mathbf{1}_X \qquad (18.2)$$

$x \in \mathbb{R}^n$

$f(x) = x^T M x$

is submodular <u>*iff*</u> *the off-diagonal elements of* M *are non-positive.*

Proof.

- Given a complete graph $G = (V, E)$, recall that $E(X)$ is the edge set with both vertices in $X \subseteq V(G)$, and that $|E(X)|$ is supermodular.
- Non-negative modular weights $w^+ : E \to \mathbb{R}_+$, $w(E(X))$ is also supermodular, so $-w(E(X))$ is submodular.
- f is a modular function $m^\intercal \mathbf{1}_A = m(A)$ added to a weighted submodular function, hence f is submodular.

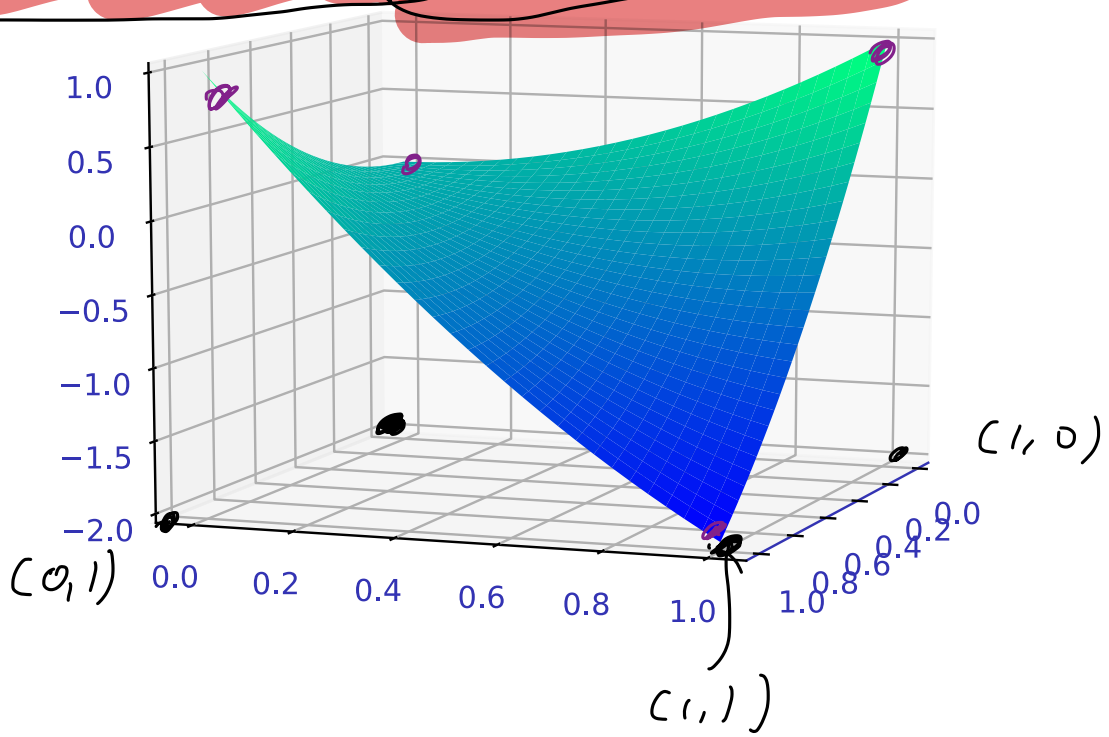
...

Example non-convex non-concave submodular quadratic

- Simply set $M = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.

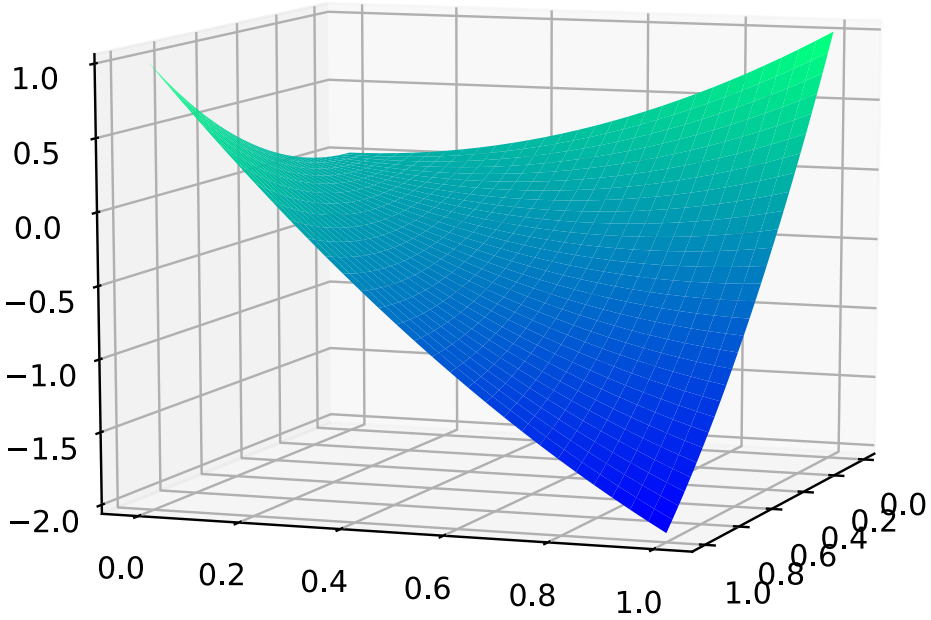
Example non-convex non-concave submodular quadratic

- Simply set $M = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are 3 and -1.



Example non-convex non-concave submodular quadratic

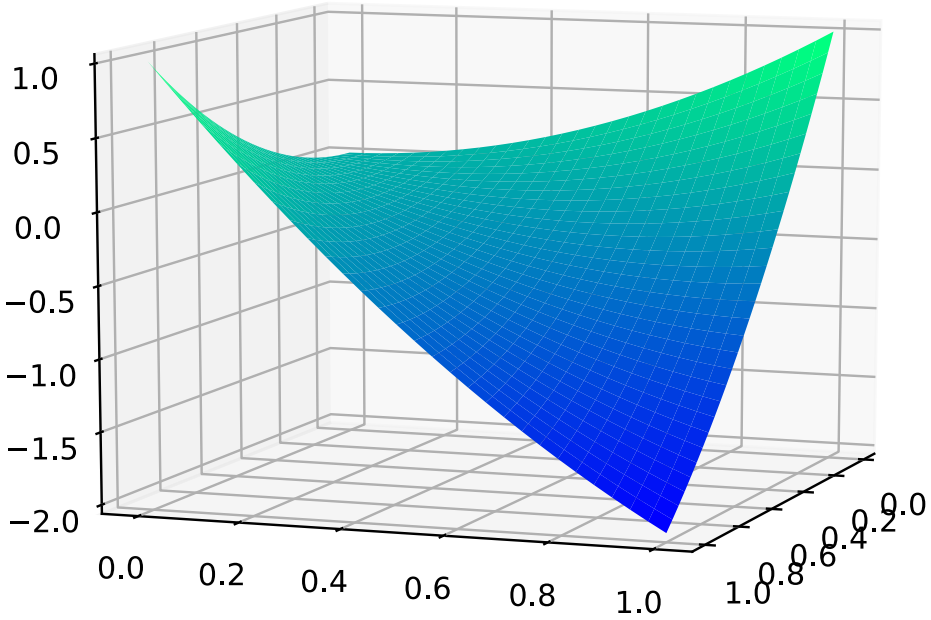
- Simply set $M = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are 3 and -1.



- But f is submodular on the reals since the off-diagonals of the Hessian are negative.

Example non-convex non-concave submodular quadratic

- Simply set $M = \begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are 3 and -1.



- But f is submodular on the reals since the off-diagonals of the Hessian are negative.
- Note if we define $f'(A) = f(\mathbf{1}_A)$, then $f' : 2^V \to \mathbb{R}$ is a submodular set function, hence f is a continuous extension of f'.

Alternate Definition

- We can equivalently define a submodular function using a definition analogous to the four-points property.

Alternate Definition

- We can equivalently define a submodular function using a definition analogous to the four-points property.

Definition 18.3.4 (submodularity)

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be submodular (on the real lattice) if for all $x \in \mathbb{R}^n$ and elements $i, j \in V$ and non-negative real values $\alpha_i, \alpha_j \geq 0$ we have that

$$f(x + \alpha_i \mathbf{1}_i) + f(x + \alpha_j \mathbf{1}_j) \geq f(x) + f(x + \alpha_i \mathbf{1}_i + \alpha_j \mathbf{1}_j) \tag{18.15}$$

$f'(A) = f(1_A)$

Alternate Definition

- We can equivalently define a submodular function using a definition analogous to the four-points property.

Definition 18.3.4 (submodularity)

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be submodular (on the real lattice) if for all $x \in \mathbb{R}^n$ and elements $i, j \in V$ and non-negative real values $\alpha_i, \alpha_j \geq 0$ we have that

$$f(x + \alpha_i \mathbf{1}_i) + f(x + \alpha_j \mathbf{1}_j) \geq f(x) + f(x + \alpha_i \mathbf{1}_i + \alpha_j \mathbf{1}_j) \tag{18.15}$$

Proposition 18.3.5

The two definitions of submodularity on the real lattice are the same.

Diminishing-Returns (DR) Submodularity

- What about diminishing returns?

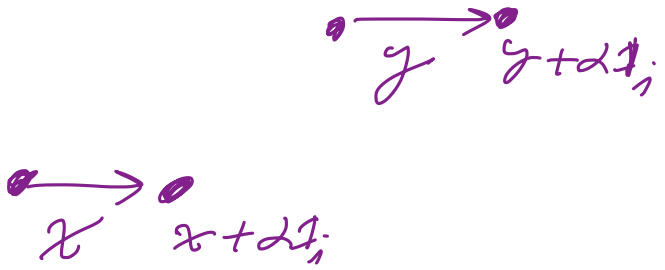
Diminishing-Returns (DR) Submodularity

- What about diminishing returns?

Definition 18.3.6

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be DR-submodular if for all $x, y \in \mathbb{R}^n$ with $x \leq y$ (element wise) and for any $i \in V$ and $\alpha > 0$, we have

$$f(x + \alpha \mathbf{1}_i) - f(x) \geq f(y + \alpha \mathbf{1}_i) - f(y) \tag{18.16}$$



Diminishing-Returns (DR) Submodularity

- What about diminishing returns?

Definition 18.3.6

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be DR-submodular if for all $x, y \in \mathbb{R}^n$ with $x \leq y$ (element wise) and for any $i \in V$ and $\alpha > 0$, we have

$$f(x + \alpha \mathbf{1}_i) - f(x) \geq f(y + \alpha \mathbf{1}_i) - f(y) \tag{18.16}$$

- For twice-differential functions,

Diminishing-Returns (DR) Submodularity

- What about diminishing returns?

Definition 18.3.6

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be DR-submodular if for all $x, y \in \mathbb{R}^n$ with $x \leq y$ (element wise) and for any $i \in V$ and $\alpha > 0$, we have

$$f(x + \alpha \mathbf{1}_i) - f(x) \geq f(y + \alpha \mathbf{1}_i) - f(y) \tag{18.16}$$

- For twice-differential functions,
- When f is twice differentiable, there is a similar property to the above. Namely, DR-submodularity is in this case identical to all the elements of the Hessian of f being non-positive, i.e.,

$$\forall i, j, \forall x \in \mathbb{R}^n, \quad \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \leq 0 \tag{18.17}$$

Note, non-positives along the diagonal (and everywhere) are required.

Diminishing-Returns (DR) Submodularity

- What about diminishing returns?

Definition 18.3.6

A given function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be DR-submodular if for all $x, y \in \mathbb{R}^n$ with $x \leq y$ (element wise) and for any $i \in V$ and $\alpha > 0$, we have

$$f(x + \alpha \mathbf{1}_i) - f(x) \geq f(y + \alpha \mathbf{1}_i) - f(y) \tag{18.16}$$

- For twice-differential functions,
- When f is twice differentiable, there is a similar property to the above. Namely, DR-submodularity is in this case identical to all the elements of the Hessian of f being non-positive, i.e.,

$$\forall i, j, \forall x \in \mathbb{R}^n, \quad \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \leq 0 \tag{18.17}$$

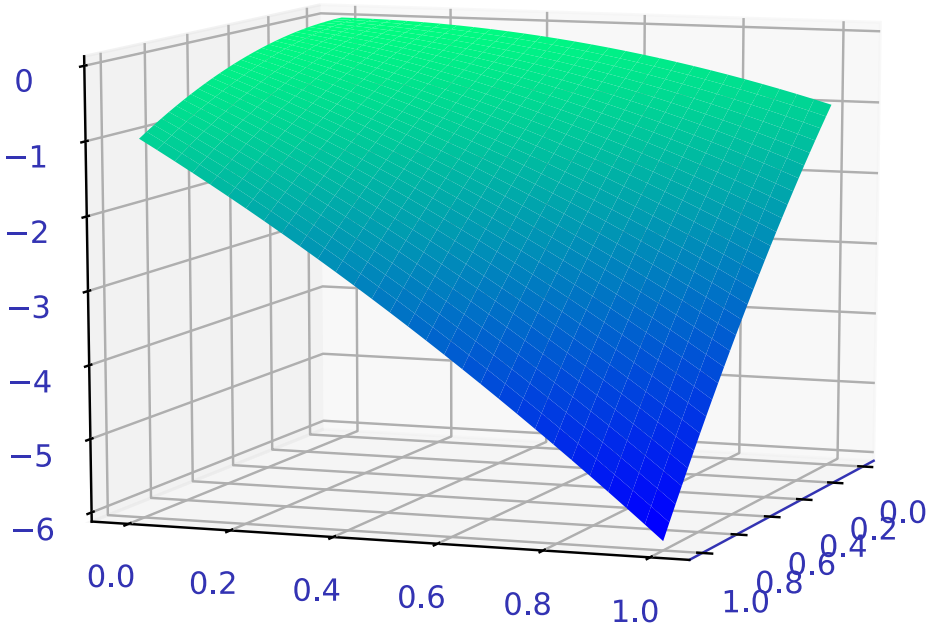
  Note, non-positives along the diagonal (and everywhere) are required.
- Again, neither the same condition as convexity nor concavity.

Example non-convex non-concave DR-Submodular quadratic

- Simply set $M = \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$.

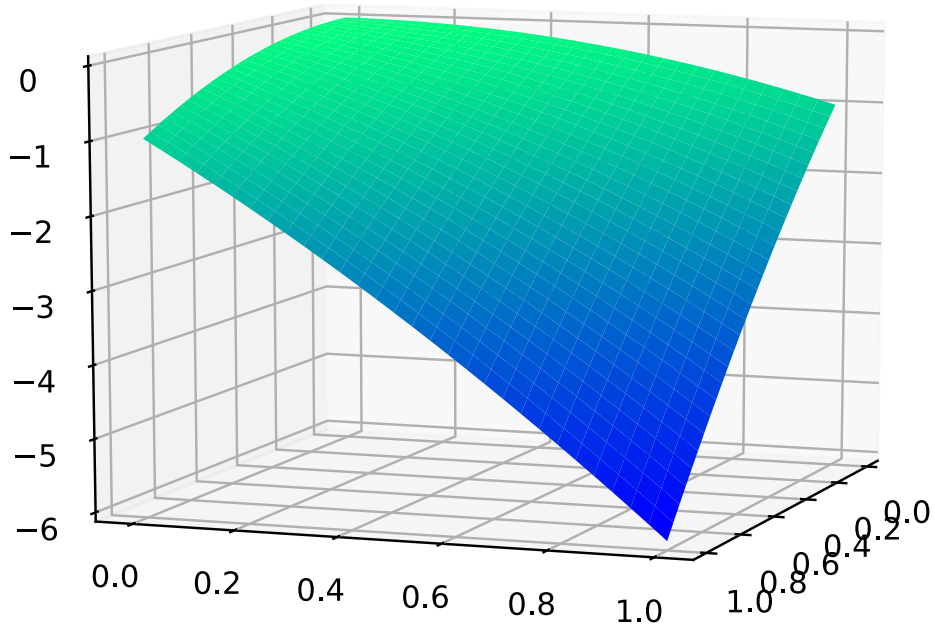
Example non-convex non-concave DR-Submodular quadratic

- Simply set $M = \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are -3 and 1.



Example non-convex non-concave DR-Submodular quadratic

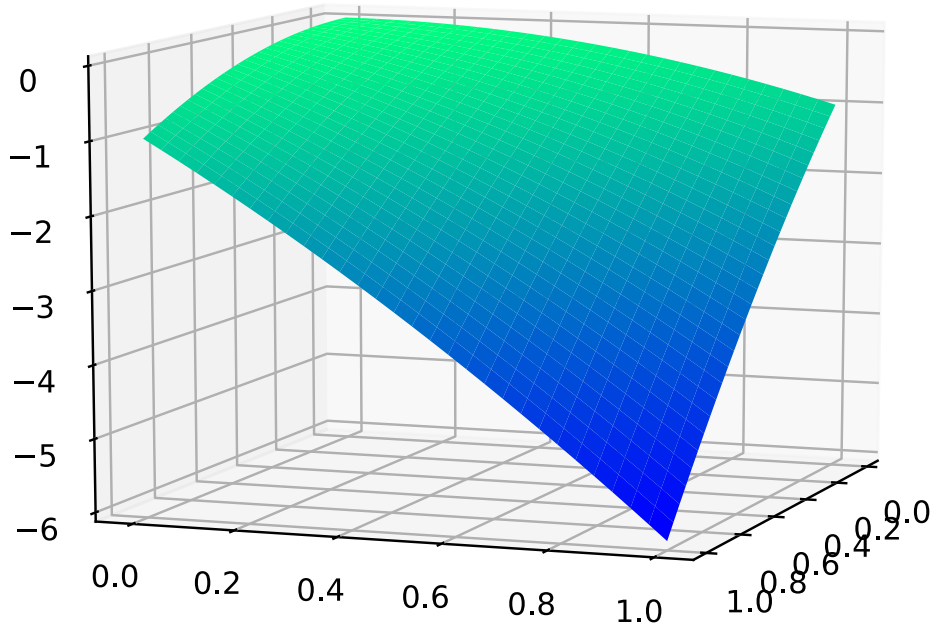
- Simply set $M = \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are -3 and 1.



- Again, f is submodular on the reals since the off-diagonals of the Hessian are negative.

Example non-convex non-concave DR-Submodular quadratic

- Simply set $M = \begin{pmatrix} -1 & -2 \\ -2 & -1 \end{pmatrix}$.
- Then $f(x) = x^\intercal \mathbf{M} x$ is neither convex nor concave but it is submodular since eigenvalues are -3 and 1.



- Again, f is submodular on the reals since the off-diagonals of the Hessian are negative.

but, vice versa is true on $\{0,1\}^V$, i.e., hypercube vertices.

- In general, DR-submodularity implies submodularity but not vice versa.

on the real lattice.

DR-submodularity Examples

- Proposition: in R2, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = \min(x_1, x_2)$, $f(x_1, x_2) = \max(x_1, x_2)$, $f(x_1, x_2) = x_1 * x_2$, and $f(x_1, x_2) = \|x_1 - x_2|^p$, for $p \geq 1$, are all DR-submodular.

Check

DR-submodularity Examples

- Proposition: in R2, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = \min(x_1, x_2)$, $f(x_1, x_2) = \max(x_1, x_2)$, $f(x_1, x_2) = x_1 * x_2$, and $f(x_1, x_2) = \|x_1 - x_2|^p$, for $p \geq 1$, are all DR-submodular.
- Proposition: multilinear extension of a submodular function is DR-submodular.

DR-submodularity Examples

- Proposition: in R2, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = x_1 + x_2$, $f(x_1, x_2) = \min(x_1, x_2)$, $f(x_1, x_2) = \max(x_1, x_2)$, $f(x_1, x_2) = x_1 * x_2$, and $f(x_1, x_2) = \|x_1 - x_2|^p$, for $p \geq 1$, are all DR-submodular.
- Proposition: multilinear extension of a submodular function is DR-submodular.
- DR-submodular, despite being neither convex nor concave, are possible to either exactly or approximately maximize, and minimize (and thus are an interesting class of non-convex/non-concave functions). *see Hassani, Soltanolkotabi, Karbasi, "Gradient Methods for Submodular Maximization", 2017.* and references therein.

Also see

- F. Bach. Submodular functions: from discrete to continous domains. arXiv preprint arXiv:1511.00394, 2015.

Submodular Max and polyhedral approaches

- Most of the approaches for submodular max have not used such an approach, probably due to the difficulty in computing the "concave extension" of a submodular function (the convex extension is easy, namely the Lovász extension).

Submodular Max and polyhedral approaches

- Most of the approaches for submodular max have not used such an approach, probably due to the difficulty in computing the "concave extension" of a submodular function (the convex extension is easy, namely the Lovász extension).
- We can achieve progress on this front using the "multilinear extension".

Multilinear extension (review)

Definition 18.4.1

For a set function $f : 2^V \to \mathbb{R}$, define its multilinear extension $F : [0,1]^V \to \mathbb{R}$ by

$$F(x) = \sum_{S \subseteq V} f(S) \prod_{i \in S} x_i \prod_{j \in V \setminus S} (1 - x_j) \tag{18.18}$$

- Note that $F(x) = Ef(\hat{x})$ where $\hat{x}$ is a random binary vector over $\{0,1\}^V$ with elements independent w. probability x_i for $\hat{x}_i$.
- While this is defined for any set function, we have:

Lemma 18.4.2

Let $F : [0,1]^V \to \mathbb{R}$ be multilinear extension of set function $f : 2^V \to \mathbb{R}$, then

- *If f is monotone non-decreasing, then $\frac{\partial F}{\partial x_i} \geq 0$ for all $i \in V$, $x \in [0,1]^V$.*
- *If f is submodular, then $\frac{\partial^2 F}{\partial x_i \partial x_j} \leq 0$ for all $i, j\ inV$, $x \in [0,1]^V$.*

Combinatorial Submodular Max

- We saw that monotone submodular max subject to a cardinality constraint has a $1 - 1/e$ approximation guarantee using the simple (combinatorial) greedy algorithm.

Combinatorial Submodular Max

- We saw that monotone submodular max subject to a cardinality constraint has a $1 - 1/e$ approximation guarantee using the simple (combinatorial) greedy algorithm.
- We also saw that monotone submodular max subject to p simultaneous matroid independence constraints has a $1/(1+p)$ approximation guarantee (so $1/2$ for one matroid). Can this be improved upon?

Combinatorial Submodular Max

- We saw that monotone submodular max subject to a cardinality constraint has a $1 - 1/e$ approximation guarantee using the simple (combinatorial) greedy algorithm.
- We also saw that monotone submodular max subject to p simultaneous matroid independence constraints has a $1/(1+p)$ approximation guarantee (so $1/2$ for one matroid). Can this be improved upon?
- The answer is yes.

Submodular Max and Polyhedral Approaches

- Basic idea for $\max(f(A) : A \in \mathcal{I})$: Given a set of constraints $\mathcal{I}$, we form a polytope $P_{\mathcal{I}}$ such that $\{\mathbf{1}_I : I \in \mathcal{I}\} \subseteq P_{\mathcal{I}}$ (e.g., $P_{\mathcal{I}} = \text{conv}\{\mathbf{1}_I : I \in \mathcal{I}\}$ where $\mathcal{I}$ are the independent sets of a matroid).
- We find $\max_{x \in P_{\mathcal{I}}} \tilde{f}(x)$ where $\tilde{f}(x)$ is the multi-linear extension of f, to find a fractional solution x^*. This can't be done exactly but can be done with a $1 - 1/e$ guarantee for the continuous optimum.
- We then round x^* to a point on the hypercube, thus giving us a solution to the discrete problem. The rounding keeps the $1 - 1/e$ guarantee for the discrete solution.

Continuous Greedy Algorithm

- For any down-closed polytope $P \ni 0$, we define a continuous process where $x_t \in P$ moves continuously from time $t = 0$ to $t = 1$ according to the following differential equation. We start at $x_0 = 0 \in P$ and at the end $x_1 \in P$ approximates $\max_{x \in P_{\mathcal{I}}} F(x)$.

Continuous Greedy Algorithm

- For any down-closed polytope $P \ni 0$, we define a continuous process where $x_t \in P$ moves continuously from time $t = 0$ to $t = 1$ according to the following differential equation. We start at $x_0 = 0 \in P$ and at the end $x_1 \in P$ approximates $\max_{x \in P_{\mathcal{I}}} F(x)$.

1 Start with $x_0 \leftarrow 0 \in P$;

2 Utilize $\frac{\partial x}{\partial t} = v_{\text{max}}(x)$ where $v_{\text{max}}(x) = \text{argmax}_{y \in P}(\langle y, \nabla \tilde{f}(x) \rangle)$ from $t = 0$ to $t = 1$.

Continuous Greedy Algorithm

- For any down-closed polytope $P \ni 0$, we define a continuous process where $x_t \in P$ moves continuously from time $t = 0$ to $t = 1$ according to the following differential equation. We start at $x_0 = 0 \in P$ and at the end $x_1 \in P$ approximates $\max_{x \in P_{\mathcal{I}}} \tilde{f}(x)$.

---

1 Start with $x_0 \leftarrow 0 \in P$;

2 Utilize $\frac{\partial x}{\partial t} = v_{\text{max}}(x)$ where $v_{\text{max}}(x) = \text{argmax}_{y \in P}(\langle y, \nabla \tilde{f}(x) \rangle)$ from $t = 0$ to $t = 1$.

- Claim is that, in the above, $x_1 \in P$ and $\tilde{f}(x_1) \geq (1 - 1/e) \max_{x \in P_{\mathcal{I}}} \tilde{f}(x)$.

$x_{t=1}$

OPT

Continuous Greedy Algorithm

- For any down-closed polytope $P \ni 0$, we define a continuous process where $x_t \in P$ moves continuously from time $t = 0$ to $t = 1$ according to the following differential equation. We start at $x_0 = 0 \in P$ and at the end $x_1 \in P$ approximates $\max_{x \in P_{\mathcal{I}}} F(x)$.

---

1 Start with $x_0 \leftarrow 0 \in P$;

2 Utilize $\frac{\partial x}{\partial t} = v_{\text{max}}(x)$ where $v_{\text{max}}(x) = \text{argmax}_{y \in P}(\langle y, \nabla \tilde{f}(x) \rangle)$ from $t = 0$ to $t = 1$.

- Claim is that, in the above, $x_1 \in P$ and $\tilde{f}(x_1) \geq (1 - 1/e) \max_{x \in P_{\mathcal{I}}} F(x)$.
- In practice, line 2 is done using in iterative procedure analogous to gradient descent getting $x_{i+1} \leftarrow x_i + \nu v_{\text{max}}(x_i)$ where ν is rate parameter and i is the iteration number. Note the argmax can be easily solve using the same Edmond's greedy algorithm we've discussed.

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:
- 1) constant factor approximation algorithm for $\max\{\tilde{f}(x) : x \in P\}$ for any down-monotone solvable polytope P and $\tilde{f}$ multilinear extension of any non-negative submodular function.

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:
- 1) constant factor approximation algorithm for $\max\{F(x) : x \in P\}$ for any down-monotone solvable polytope P and F multilinear extension of any non-negative submodular function.
- 2) A randomized rounding (pipage rounding) scheme to obtain an integer solution

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:
- 1) constant factor approximation algorithm for $\max\{F(x) : x \in P\}$ for any down-monotone solvable polytope P and F multilinear extension of any non-negative submodular function.
- 2) A randomized rounding (pipage rounding) scheme to obtain an integer solution
- 3) An optimal $(1 - 1/e)$ instance of their rounding scheme that can be used for a variety of interesting independence systems, including $O(1)$ knapsacks, k matroids and $O(1)$ knapsacks, a k-matchoid and ℓ sparse packing integer programs, and unsplittable flow in paths and trees.

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:
- 1) constant factor approximation algorithm for $\max\{F(x) : x \in P\}$ for any down-monotone solvable polytope P and F multilinear extension of any non-negative submodular function.
- 2) A randomized rounding (pipage rounding) scheme to obtain an integer solution
- 3) An optimal $(1 - 1/e)$ instance of their rounding scheme that can be used for a variety of interesting independence systems, including $O(1)$ knapsacks, k matroids and $O(1)$ knapsacks, a k-matchoid and ℓ sparse packing integer programs, and unsplittable flow in paths and trees.
- Also, J. Vondrak showed that this scheme achieves the $\frac{1}{c_f}(1 - e^{-c_f})$ curvature based bound for any matroid, which matches the bound we had earlier for uniform matroids with standard greedy. $\frac{1}{1+c_f}$

Other results

- In a paper by Chekuri, Vondrak, and Zenklusen, they show:
- 1) constant factor approximation algorithm for $\max\{F(x) : x \in P\}$ for any down-monotone solvable polytope P and F multilinear extension of any non-negative submodular function.
- 2) A randomized rounding (pipage rounding) scheme to obtain an integer solution
- 3) An optimal $(1 - 1/e)$ instance of their rounding scheme that can be used for a variety of interesting independence systems, including $O(1)$ knapsacks, k matroids and $O(1)$ knapsacks, a k-matchoid and ℓ sparse packing integer programs, and unsplittable flow in paths and trees.
- Also, J. Vondrak showed that this scheme achieves the $\frac{1}{c_f}(1 - e^{-c_f})$ curvature based bound for any matroid, which matches the bound we had earlier for uniform matroids with standard greedy.
- In general, one needs to do Monte-Carlo methods to estimate the multilinear extension to avoid needing to do exponential number of function queries (so further approximations would apply, but we get polynomial number of queries in $1/\epsilon^2$ for additional ϵ approximation).

Some Important Papers Along These Lines

- Vondrák, J.: Optimal approximation for the submodular welfare problem in the value oracle model. In: Proceedings, ACM STOC, pp. 67-74 (2008) *Early paper that showing approach works for submodular welfare problem*
- J. Vondrák. Submodularity and curvature: the optimal algorithm, in RIMS Kokyuroku Bessatsu, Workshop on combinatorial optimization, Kyoto 2008. *Shows curvature based bounds for continuous greedy with multilinear extension.*
- Calinescu, G., Chekuri, C., Pál, M., Vondrák, J.: Maximizing a monotone submodular function subject to a matroid constraint. SIAM J. Comput., 40 (2011), pp. 1740-1766. *followup paper that shows $1 - 1/e$ algorithm for monotone submodular max subject to a single matroid constraint.*
- Chekuri C, Vondrák J, Zenklusen R (2011b) Submodular function maximization via the multilinear relaxation and contention resolution schemes. Fortnow L, Vadhan SP, ed. Proc. 43rd ACM Sympos. Theory Comput., STOC '11 (ACM, New York), 783-792. *Very detailed and long paper that allows f to be non-monotone, uses using different rounding methods.*
- M. Feldman, J. Naor, and R. Schwartz, A unified continuous greedy algorithm for submodular maximization, in Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science, 2011, pp. 570-579. *Another paper with similar purpose (e.g., non-monotone) but different analysis.*

Review from lecture 11

The next slide comes from lecture 11.

A polymatroid function's polyhedron is a polymatroid.

Theorem 18.5.1

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

$$y^x(E) = \textit{rank}(x) \triangleq \max\left(y(E) : y \leq x, y \in P_f^+\right)$$
$$= \min\left(x(A) + f(E \setminus A) : A \subseteq E\right) \tag{18.10}$$

$= \min(x(E \setminus A) + f(A) : A \subseteq E) = \cdots$

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

Taking $E \setminus B = \text{supp}(x)$ (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:

$$\text{rank}\left(\frac{1}{\epsilon}\mathbf{1}_{E\setminus B}\right) = f(E \setminus B) = \max\left\{y(E \setminus B) : y \in P_f^+\right\} \tag{18.11}$$

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

Review from lecture 12

The next slide comes from lecture 12.

Matroid instance of Theorem ??

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

Corollary 18.5.2

We have that:

$$\max\{y(E) : y \in P_{\textit{ind. set}}(M), y \leq x\} = \min\{r_M(A) + x(E \setminus A) : A \subseteq E\} \tag{18.21}$$

where r_M is the matroid rank function of some matroid.

Most violated inequality problem in matroid polytope case

- Consider

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E\right\} \tag{18.19}$$

Most violated inequality problem in matroid polytope case

- Consider

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E\right\} \tag{18.19}$$

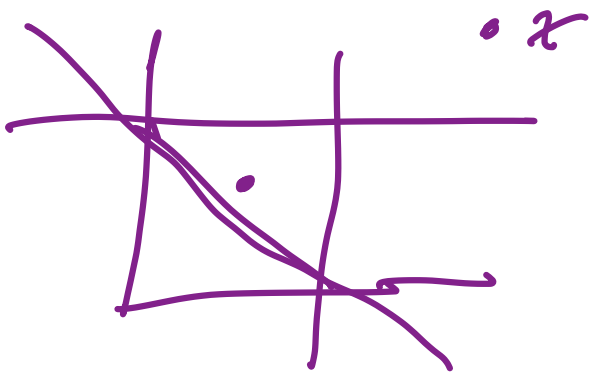
- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_r^+$.

Most violated inequality problem in matroid polytope case

- Consider

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E\right\} \tag{18.19}$$

- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_r^+$.
- Hence, there must be a set of $\mathcal{W} \subseteq 2^V$, each member of which corresponds to a violated inequality, i.e., equations of the form $x(A) > r_M(A)$ for $A \in \mathcal{W}$.



Most violated inequality problem in matroid polytope case

- Consider

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E\right\} \tag{18.19}$$

- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_r^+$.
- Hence, there must be a set of $\mathcal{W} \subseteq 2^V$, each member of which corresponds to a violated inequality, i.e., equations of the form $x(A) > r_M(A)$ for $A \in \mathcal{W}$.
- The most violated inequality when x is considered w.r.t. P_r^+ corresponds to the set A that maximizes $x(A) - r_M(A)$, i.e., the most violated inequality is valuated as:

$$\max\left\{x(A) - r_M(A) : A \in \mathcal{W}\right\} = \max\left\{x(A) - r_M(A) : A \subseteq E\right\} \tag{18.20}$$

Most violated inequality problem in matroid polytope case

- Consider

$$P_r^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq r_M(A), \forall A \subseteq E\right\} \tag{18.19}$$

- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_r^+$.
- Hence, there must be a set of $\mathcal{W} \subseteq 2^V$, each member of which corresponds to a violated inequality, i.e., equations of the form $x(A) > r_M(A)$ for $A \in \mathcal{W}$.
- The most violated inequality when x is considered w.r.t. P_r^+ corresponds to the set A that maximizes $x(A) - r_M(A)$, i.e., the most violated inequality is valuated as:

$$\max\left\{x(A) - r_M(A) : A \in \mathcal{W}\right\} = \max\left\{x(A) - r_M(A) : A \subseteq E\right\} \tag{18.20}$$

- Since x is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min\left\{r_M(A) + x(E \setminus A) : A \subseteq E\right\} \tag{18.21}$$

Most violated inequality/polymatroid membership/SFM

- Consider

$$P_f^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\right\} \tag{18.22}$$

Most violated inequality/polymatroid membership/SFM

- Consider

$$P_f^+ = \left\{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\right\} \tag{18.22}$$

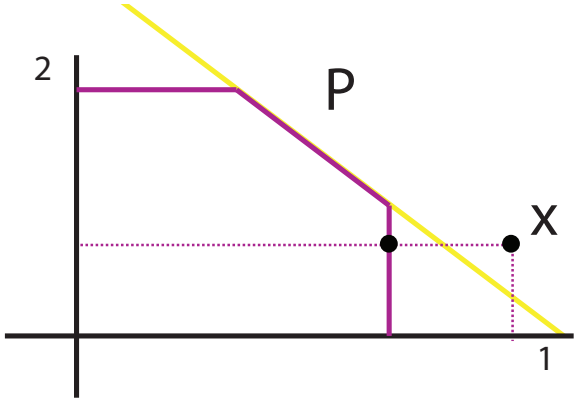
- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_f^+$.

Most violated inequality/polymatroid membership/SFM

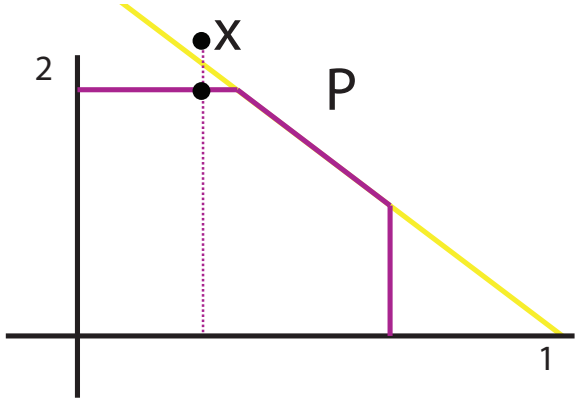
- Consider

$$P_f^+ = \{x \in \mathbb{R}^E : x \geq 0, x(A) \leq f(A), \forall A \subseteq E\} \tag{18.22}$$

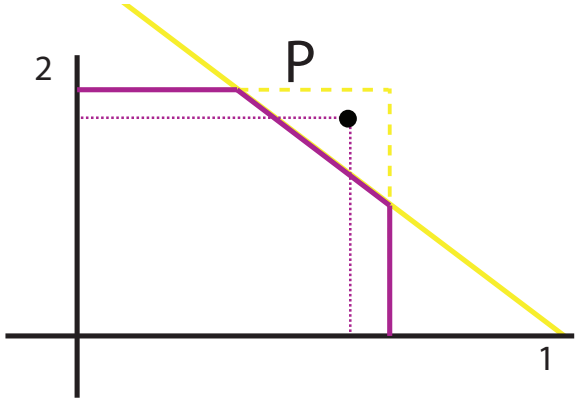
- Suppose we have any $x \in \mathbb{R}_+^E$ such that $x \notin P_f^+$.
- Hence, there must be a set of $\mathcal{W} \subseteq 2^V$, each member of which corresponds to a violated inequality, i.e., equations of the form $x(A) > r_M(A)$ for $A \in \mathcal{W}$.



$\mathcal{W} = \{\{1\}, \{1,2\}\}$



$\mathcal{W} = \{\{2\}, \{1,2\}\}$



$\mathcal{W} = \{\{1,2\}\}$

Most violated inequality/polymatroid membership/SFM

- The most violated inequality when x is considered w.r.t. P_f^+ corresponds to the set A that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max\{x(A) - f(A) : A \in \mathcal{W}\} = \max\{x(A) - f(A) : A \subseteq E\} \quad (18.23)$$

Most violated inequality/polymatroid membership/SFM

- The most violated inequality when x is considered w.r.t. P_f^+ corresponds to the set A that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max\{x(A) - f(A) : A \in \mathcal{W}\} = \max\{x(A) - f(A) : A \subseteq E\} \quad (18.23)$$

- Since x is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min\{f(A) + x(E \setminus A) : A \subseteq E\} \quad (18.24)$$

Most violated inequality/polymatroid membership/SFM

- The most violated inequality when x is considered w.r.t. P_f^+ corresponds to the set A that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max\{x(A) - f(A) : A \in \mathcal{W}\} = \max\{x(A) - f(A) : A \subseteq E\} \quad (18.23)$$

- Since x is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min\{f(A) + x(E \setminus A) : A \subseteq E\} \quad (18.24)$$

- More importantly, $\min\{f(A) + x(E \setminus A) : A \subseteq E\}$ is a form of submodular function minimization, namely $\min\{f(A) - x(A) : A \subseteq E\}$ for a submodular f and $x \in \mathbb{R}_+^E$, consisting of a difference of polymatroid and modular function (so $f - x$ is no longer necessarily monotone, nor positive).

Most violated inequality/polymatroid membership/SFM

- The most violated inequality when x is considered w.r.t. P_f^+ corresponds to the set A that maximizes $x(A) - f(A)$, i.e., the most violated inequality is valuated as:

$$\max\{x(A) - f(A) : A \in \mathcal{W}\} = \max\{x(A) - f(A) : A \subseteq E\} \quad (18.23)$$

- Since x is modular and $x(E \setminus A) = x(E) - x(A)$, we can express this via a min as in;:

$$\min\{f(A) + x(E \setminus A) : A \subseteq E\} \quad (18.24)$$

- More importantly, $\min\{f(A) + x(E \setminus A) : A \subseteq E\}$ is a form of submodular function minimization, namely $\min\{f(A) - x(A) : A \subseteq E\}$ for a submodular f and $x \in \mathbb{R}_+^E$, consisting of a difference of polymatroid and modular function (so $f - x$ is no longer necessarily monotone, nor positive).
- We will ultimately answer how general this form of SFM is.

Review from Lecture 6

The following three slides are review from lecture 6.

Matroids, other definitions using matroid rank $r : 2^V \to \mathbb{Z}_+$

Definition 18.6.3 (closed/flat/subspace)

A subset $A \subseteq E$ is closed in (equivalently, a flat or a subspace of) a matroid M if for all $x \in E \setminus A$, $r(A \cup \{x\}) = r(A) + 1$.

Definition: A hyperplane is a flat of rank $r(M) - 1$.

Definition 18.6.4 (closure)

Given $A \subseteq E$, the closure (or span) of A, is defined by $\text{span}(A) = \{b \in E : r(A \cup \{b\}) = r(A)\}$. $\supseteq A$

Therefore, a closed set A has $\text{span}(A) = A$, and the span of a set is closed.

Definition 18.6.5 (circuit)

A subset $A \subseteq E$ is circuit or a cycle if it is an inclusionwise-minimal dependent set (i.e., if $r(A) < |A|$ and for any $a \in A$, $r(A \setminus \{a\}) = |A| - 1$).

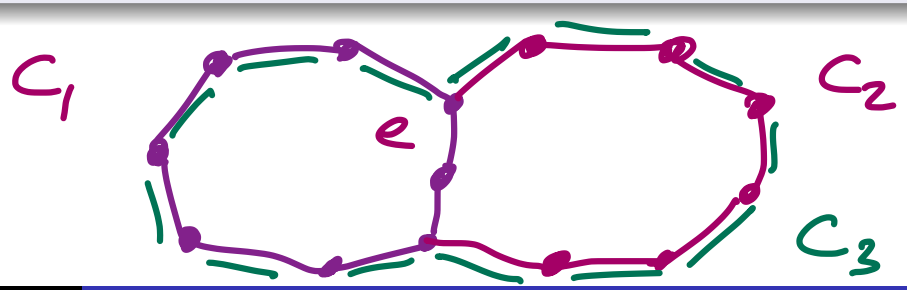
Matroids by circuits

A set is independent if and only if it contains no circuit. Therefore, it is not surprising that circuits can also characterize a matroid.

Theorem 18.6.3 (Matroid by circuits)

Let E be a set and $\mathcal{C}$ be a collection of subsets of E that satisfy the following three properties:

1. *(C1): $\emptyset \notin \mathcal{C}$*
2. *(C2): if $C_1, C_2 \in \mathcal{C}$ and $C_1 \subseteq C_2$, then $C_1 = C_2$.*
3. *(C3): if $C_1, C_2 \in \mathcal{C}$ with $C_1 \neq C_2$, and $e \in C_1 \cap C_2$, then there exists a $C_3 \in \mathcal{C}$ such that $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\}$.*



Matroids by circuits

Several circuit definitions for matroids.

Theorem 18.6.3 (Matroid by circuits)

Let E be a set and $\mathcal{C}$ be a collection of nonempty subsets of E, such that no two sets in $\mathcal{C}$ are contained in each other. Then the following are equivalent.

1. *$\mathcal{C}$ is the collection of circuits of a matroid;*
2. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$;*
3. *if $C, C' \in \mathcal{C}$, and $x \in C \cap C'$, and $y \in C \setminus C'$, then $(C \cup C') \setminus \{x\}$ contains a set in $\mathcal{C}$ containing y;*

Again, think about this for a moment in terms of linear spaces and matrices, and spanning trees.

Fundamental circuits in matroids

Lemma 18.6.1

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in M.

Proof.

$\square$

Fundamental circuits in matroids

Lemma 18.6.1

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in M.

Proof.

- Suppose, to the contrary, that there are two distinct circuits C_1, C_2 such that $(C_1 \cup C_2) \subseteq I \cup \{e\}$.

$\square$

Fundamental circuits in matroids

Lemma 18.6.1

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in M.

Proof.

- Suppose, to the contrary, that there are two distinct circuits C_1, C_2 such that $C_1 \cup C_2 \subseteq I \cup \{e\}$.
- Then $e \in C_1 \cap C_2$, and by (C2), there is a circuit C_3 of M s.t. $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\} \subseteq I$

$\square$

Fundamental circuits in matroids

Lemma 18.6.1

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in M.

Proof.

- Suppose, to the contrary, that there are two distinct circuits C_1, C_2 such that $C_1 \cup C_2 \subseteq I \cup \{e\}$.
- Then $e \in C_1 \cap C_2$, and by (C2), there is a circuit C_3 of M s.t. $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\} \subseteq I$
- This contradicts the independence of I.

$\square$

Fundamental circuits in matroids

Lemma 18.6.1

Let $I \in \mathcal{I}(M)$, and $e \in E$, then $I \cup \{e\}$ contains at most one circuit in M.

Proof.

- Suppose, to the contrary, that there are two distinct circuits C_1, C_2 such that $C_1 \cup C_2 \subseteq I \cup \{e\}$.
- Then $e \in C_1 \cap C_2$, and by (C2), there is a circuit C_3 of M s.t. $C_3 \subseteq (C_1 \cup C_2) \setminus \{e\} \subseteq I$
- This contradicts the independence of I.

$\square$

In general, let $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (commonly called the fundamental circuit in M w.r.t. I and e).

$e \in \text{span}(I) \setminus I$, $C(I, e)$ is a circuit.

Matroids: The Fundamental Circuit

- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in M w.r.t. I and e, if it exists).

Matroids: The Fundamental Circuit

- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in M w.r.t. I and e, if it exists).
- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).

Matroids: The Fundamental Circuit

- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in M w.r.t. I and e, if it exists).
- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).
- If $e \in I$, then $I + e = I$ doesn't create a circuit. In such cases, $C(I, e)$ is not really defined.

Matroids: The Fundamental Circuit

- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in M w.r.t. I and e, if it exists).
- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).
- If $e \in I$, then $I + e = I$ doesn't create a circuit. In such cases, $C(I, e)$ is not really defined.
- In such cases, we define $C(I, e) = \{e\}$, and we will soon see why.

Matroids: The Fundamental Circuit

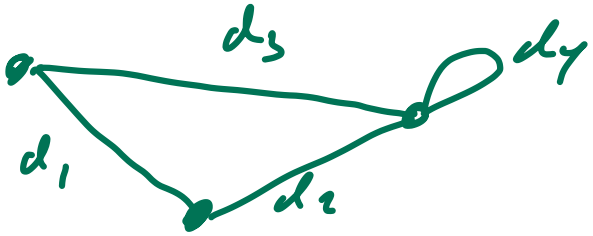
- Define $C(I, e)$ be the unique circuit associated with $I \cup \{e\}$ (the fundamental circuit in M w.r.t. I and e, if it exists).
- If $e \in \text{span}(I) \setminus I$, then $C(I, e)$ is well defined ($I + e$ creates one circuit).
- If $e \in I$, then $I + e = I$ doesn't create a circuit. In such cases, $C(I, e)$ is not really defined.
- In such cases, we define $C(I, e) = \{e\}$, and we will soon see why.
- If $e \notin \text{span}(I)$ (i.e., when $I + e$ is independent), then we set $C(I, e) = \emptyset$.

Union of matroid bases of a set

Lemma 18.6.2

Let $\mathcal{B}(D)$ be the set of bases of any set D. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free (i.e., no dependent singleton elements) set $D \subseteq E$, we have:

$$\bigcup_{B \in \mathcal{B}(D)} B = D. \tag{18.25}$$



$\text{rank}(\{d_1, d_2, d_3, d_4\}) = 2$

$\mathcal{B}(D) = \{\{d_1, d_2\}, \{d_2, d_3\}, \{d_3, d_1\}\}$

$\bigcup_{B \in \mathcal{B}(D)} = D \setminus \{d_4\}$

Union of matroid bases of a set

Lemma 18.6.2

Let $\mathcal{B}(D)$ be the set of bases of any set D. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free (i.e., no dependent singleton elements) set $D \subseteq E$, we have:

$$\bigcup_{B \in \mathcal{B}(D)} B = D. \tag{18.25}$$

Proof.

- Define $D' \triangleq \bigcup_{B \in \mathcal{B}(D)} \subseteq D$, suppose $\exists d \in D$ such that $d \notin D'$.

$\square$

Union of matroid bases of a set

Lemma 18.6.2

Let $\mathcal{B}(D)$ be the set of bases of any set D. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free (i.e., no dependent singleton elements) set $D \subseteq E$, we have:

$$\bigcup_{B \in \mathcal{B}(D)} B = D. \tag{18.25}$$

Proof.

- Define $D' \triangleq \bigcup_{B \in \mathcal{B}(D)} \subseteq D$, suppose $\exists d \in D$ such that $d \notin D'$.
- Hence, $\forall B \in \mathcal{B}(D)$ we have $d \notin B$, and $B + d$ must contain a <u>single</u> circuit for any B, namely $C(B, d)$.

$\square$

Union of matroid bases of a set

Lemma 18.6.2

Let $\mathcal{B}(D)$ be the set of bases of any set D. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free (i.e., no dependent singleton elements) set $D \subseteq E$, we have:

$$\bigcup_{B \in \mathcal{B}(D)} B = D. \tag{18.25}$$

Proof.

- Define $D' \triangleq \bigcup_{B \in \mathcal{B}(D)} \subseteq D$, suppose $\exists d \in D$ such that $d \notin D'$.
- Hence, $\forall B \in \mathcal{B}(D)$ we have $d \notin B$, and $B + d$ must contain a <u>single</u> circuit for any B, namely $C(B, d)$.
- Then choose $d' \in C(B, d)$ with $d' \neq d$.

$\square$

Union of matroid bases of a set

Lemma 18.6.2

Let $\mathcal{B}(D)$ be the set of bases of any set D. Then, given matroid $\mathcal{M} = (E, \mathcal{I})$, and any loop-free (i.e., no dependent singleton elements) set $D \subseteq E$, we have:

$$\bigcup_{B \in \mathcal{B}(D)} B = D. \tag{18.25}$$

Proof.

- Define $D' \triangleq \bigcup_{B \in \mathcal{B}(D)} \subseteq D$, suppose $\exists d \in D$ such that $d \notin D'$.
- Hence, $\forall B \in \mathcal{B}(D)$ we have $d \notin B$, and $B + d$ must contain a <u>single</u> circuit for any B, namely $C(B, d)$.
- Then choose $d' \in C(B, d)$ with $d' \neq d$.
- Then $B + d - d'$ is independent size-$|B|$ subset of D and hence spans D, and thus is a d-containing member of $\mathcal{B}(D)$, contradicting $d \notin D'$. $\square$

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).
- We wish to generalize closure to polymatroids.

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function f.

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function f.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function f.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.
- That is, we saw in Lecture 11 that for any $A, B \in \mathcal{D}(x)$, we have that $A \cup B \in \mathcal{D}(x)$ and $A \cap B \in \mathcal{D}(x)$, which can constitute a join and meet.

The sat function = Polymatroid Closure

- In a matroid, closure (span) of a set A are all items that A spans (eq. that depend on A).
- We wish to generalize closure to polymatroids.
- Consider $x \in P_f$ for polymatroid function f.
- Again, recall, tight sets are closed under union and intersection, and therefore form a distributive lattice.
- That is, we saw in Lecture 11 that for any $A, B \in \mathcal{D}(x)$, we have that $A \cup B \in \mathcal{D}(x)$ and $A \cap B \in \mathcal{D}(x)$, which can constitute a join and meet.
- Recall, for a given $x \in P_f$, we have defined this tight family as

$$\mathcal{D}(x) = \{A : A \subseteq E, x(A) = f(A)\} \tag{18.26}$$

The sat function = Polymatroid Closure

- Now given $x \in P_f^+$:

$$\mathcal{D}(x) = \{A : A \subseteq E, x(A) = f(A)\} \tag{18.27}$$
$$= \{A : f(A) - x(A) = 0\} \tag{18.28}$$

The sat function = Polymatroid Closure

- Now given $x \in P_f^+$:

$$\mathcal{D}(x) = \{A : A \subseteq E, x(A) = f(A)\} \tag{18.27}$$
$$= \{A : f(A) - x(A) = 0\} \tag{18.28}$$

- Since $x \in P_f^+$ and f is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $\mathcal{D}(x)$ are the zero-valued minimizers (if any) of $f'(A)$.

$x \in P_f^+$, $\forall A$, $x(A) \leq f(A)$

$\therefore \min_A f(A) - x(A) \geq 0$

The sat function = Polymatroid Closure

- Now given $x \in P_f^+$:

$$\mathcal{D}(x) = \{A : A \subseteq E, x(A) = f(A)\} \tag{18.27}$$
$$= \{A : f(A) - x(A) = 0\} \tag{18.28}$$

- Since $x \in P_f^+$ and f is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $\mathcal{D}(x)$ are the zero-valued minimizers (if any) of $f'(A)$.
- The zero-valued minimizers of f' are thus closed under union and intersection.

The sat function = Polymatroid Closure

- Now given $x \in P_f^+$:

$$\mathcal{D}(x) = \{A : A \subseteq E, x(A) = f(A)\} \tag{18.27}$$
$$= \{A : f(A) - x(A) = 0\} \tag{18.28}$$

- Since $x \in P_f^+$ and f is presumed to be polymatroid function, we see $f'(A) = f(A) - x(A)$ is a non-negative submodular function, and $\mathcal{D}(x)$ are the zero-valued minimizers (if any) of $f'(A)$.
- The zero-valued minimizers of f' are thus closed under union and intersection.
- In fact, this is true for all minimizers of a submodular function as stated in the next theorem.

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

M is a set of sets.

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

$\forall A \in \mathcal{M}$,

$f(A) \leq f(B), \forall B \subseteq E.$

$\therefore \forall A, B \in \mathcal{M}, f(A) = f(B).$

$\square$

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

Since A and B are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.

$\square$

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

Since A and B are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.

By submodularity, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{18.29}$$

but also $f(A) + f(B) \leq f(A \cup B) + f(A \cap B)$ $\square$

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

Since A and B are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.
By submodularity, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{18.29}$$

Hence, we must have $f(A) = f(B) = f(A \cup B) = f(A \cap B)$. $\square$

$\therefore \exists$ largest minimizer $\bigcup_{A \in \mathcal{M}} A$, smallest minimizer $\bigcap_{A \in \mathcal{M}} A$

Minimizers of a Submodular Function form a lattice

Theorem 18.7.1

For arbitrary submodular f, the minimizers are closed under union and intersection. That is, let $\mathcal{M} = \operatorname{argmin}_{X \subseteq E} f(X)$ be the set of minimizers of f. Let $A, B \in \mathcal{M}$. Then $A \cup B \in \mathcal{M}$ and $A \cap B \in \mathcal{M}$.

Proof.

Since A and B are minimizers, we have $f(A) = f(B) \leq f(A \cap B)$ and $f(A) = f(B) \leq f(A \cup B)$.
By submodularity, we have

$$f(A) + f(B) \geq f(A \cup B) + f(A \cap B) \tag{18.29}$$

Hence, we must have $f(A) = f(B) = f(A \cup B) = f(A \cap B)$. $\square$

Thus, the minimizers of a submodular function form a lattice, and there is a maximal and a minimal minimizer of every submodular function.

The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).

The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

$$\mathsf{cl}(x) \stackrel{\text{def}}{=} \text{sat}(x) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \tag{18.30}$$

The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

$$\mathsf{cl}(x) \stackrel{\text{def}}{=} \text{sat}(x) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \tag{18.30}$$

$$= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \tag{18.31}$$

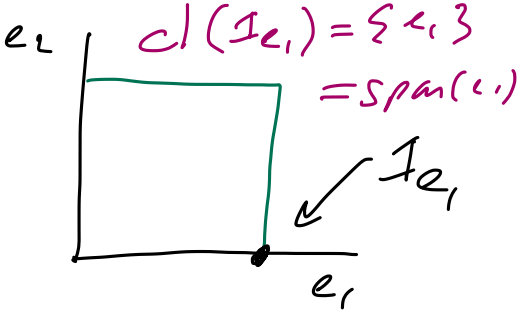
The sat function = Polymatroid Closure

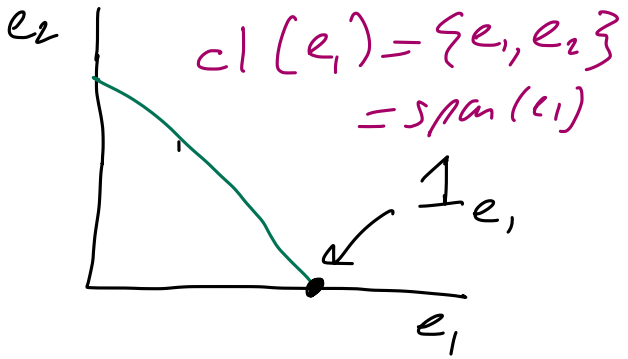
- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

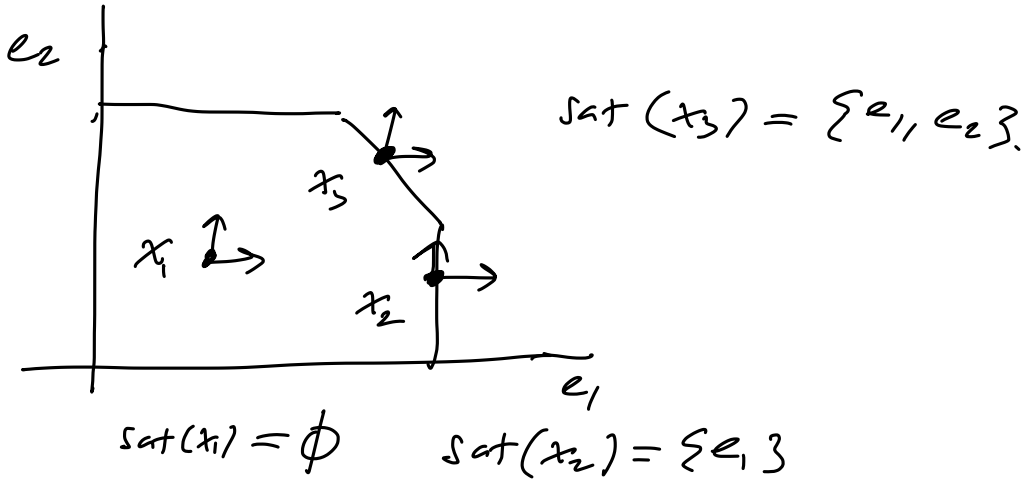
$$\mathsf{cl}(x) \stackrel{\text{def}}{=} \text{sat}(x) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \tag{18.30}$$

$$= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \tag{18.31}$$

$$= \{e : e \in E, \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f\} \tag{18.32}$$







The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

$$\begin{aligned} \mathsf{cl}(x) \stackrel{\text{def}}{=} \text{sat}(x) \stackrel{\text{def}}{=} & \bigcup \{A : A \in \mathcal{D}(x)\} && (18.30) \\ = & \bigcup \{A : A \subseteq E, x(A) = f(A)\} && (18.31) \\ = & \{e : e \in E, \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f\} && (18.32) \end{aligned}$$

- Hence, $\text{sat}(x)$ is the maximal (zero-valued) minimizer of the submodular function $f_x(A) \triangleq f(A) - x(A)$.

The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

$$\mathsf{cl}(x) \stackrel{\text{def}}{=} \mathrm{sat}(x) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \tag{18.30}$$

$$= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \tag{18.31}$$

$$= \{e : e \in E, \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f\} \tag{18.32}$$

- Hence, $\mathrm{sat}(x)$ is the maximal (zero-valued) minimizer of the submodular function $f_x(A) \triangleq f(A) - x(A)$.
- Eq. (18.32) says that sat consists of elements of E for point x that are P_f saturated (any additional positive movement, in that dimension, leaves P_f). We'll revisit this in a few slides.

The sat function = Polymatroid Closure

- Matroid closure is generalized by the unique maximal element in $\mathcal{D}(x)$, also called the polymatroid closure or sat (saturation function).
- For some $x \in P_f$, we have defined:

$$\mathsf{cl}(x) \stackrel{\text{def}}{=} \mathrm{sat}(x) \stackrel{\text{def}}{=} \bigcup \{A : A \in \mathcal{D}(x)\} \tag{18.30}$$

$$= \bigcup \{A : A \subseteq E, x(A) = f(A)\} \tag{18.31}$$

$$= \{e : e \in E, \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f\} \tag{18.32}$$

- Hence, $\mathrm{sat}(x)$ is the maximal (zero-valued) minimizer of the submodular function $f_x(A) \triangleq f(A) - x(A)$.
- Eq. (18.32) says that sat consists of elements of E for point x that are P_f saturated (any additional positive movement, in that dimension, leaves P_f). We'll revisit this in a few slides.
- First, we see how sat generalizes matroid closure.

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\mathrm{sat}(\mathbf{1}_I)$$

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\text{sat}(\mathbf{1}_I) = \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} \tag{18.34}$$

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\mathrm{sat}(\mathbf{1}_I) = \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} \tag{18.34}$$

$$= \bigcup \{A : A \subseteq E, \mathbf{1}_I(A) = r(A)\} \tag{18.35}$$

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\begin{aligned}
\mathrm{sat}(\mathbf{1}_I) &= \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} && (18.34)\\
&= \bigcup \{A : A \subseteq E, \mathbf{1}_I(A) = r(A)\} && (18.35)\\
&= \bigcup \{A : A \subseteq E, |I \cap A| = r(A)\} && (18.36)
\end{aligned}$$

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\begin{aligned} \text{sat}(\mathbf{1}_I) &= \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} && (18.34) \\ &= \bigcup \{A : A \subseteq E, \mathbf{1}_I(A) = r(A)\} && (18.35) \\ &= \bigcup \{A : A \subseteq E, |I \cap A| = r(A)\} && (18.36) \end{aligned}$$

- Notice that $\mathbf{1}_I(A) = |I \cap A| \leq |I|$.

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\begin{aligned} \mathrm{sat}(\mathbf{1}_I) &= \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} && (18.34) \\ &= \bigcup \{A : A \subseteq E, \mathbf{1}_I(A) = r(A)\} && (18.35) \\ &= \bigcup \{A : A \subseteq E, |I \cap A| = r(A)\} && (18.36) \end{aligned}$$

- Notice that $\mathbf{1}_I(A) = |I \cap A| \leq |I|$.
- Intuitively, consider an $A \supset I \in \mathcal{I}$ that doesn't increase rank, meaning $r(A) = r(I)$. If $r(A) = |I \cap A| = r(I \cap A)$, as in Eqn. (18.36), then A is in I's span, so should get $\mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I)$.

The sat function = Polymatroid Closure

- Consider matroid $(E, \mathcal{I}) = (E, r)$, some $I \in \mathcal{I}$. Then $\mathbf{1}_I \in P_r$ and

$$\mathcal{D}(\mathbf{1}_I) = \{A : \mathbf{1}_I(A) = r(A)\} \tag{18.33}$$

and

$$\begin{aligned} \mathrm{sat}(\mathbf{1}_I) &= \bigcup \{A : A \subseteq E, A \in \mathcal{D}(\mathbf{1}_I)\} && (18.34) \\ &= \bigcup \{A : A \subseteq E, \mathbf{1}_I(A) = r(A)\} && (18.35) \\ &= \bigcup \{A : A \subseteq E, |I \cap A| = r(A)\} && (18.36) \end{aligned}$$

- Notice that $\mathbf{1}_I(A) = |I \cap A| \leq |I|$.
- Intuitively, consider an $A \supset I \in \mathcal{I}$ that doesn't increase rank, meaning $r(A) = r(I)$. If $r(A) = |I \cap A| = r(I \cap A)$, as in Eqn. (18.36), then A is in I's span, so should get $\mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I)$.
- We formalize this next.

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

Proof.

- For $\mathbf{1}_I(I) = |I| = r(I)$, so $I \in \mathcal{D}(\mathbf{1}_I)$ and $I \subseteq \mathrm{sat}(\mathbf{1}_I)$. Also, $I \subseteq \mathrm{span}(I)$.

...

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

Proof.

- For $\mathbf{1}_I(I) = |I| = r(I)$, so $I \in \mathcal{D}(\mathbf{1}_I)$ and $I \subseteq \mathrm{sat}(\mathbf{1}_I)$. Also, $I \subseteq \mathrm{span}(I)$.
- Consider some $b \in \mathrm{span}(I) \setminus I$.

...

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

Proof.

- For $\mathbf{1}_I(I) = |I| = r(I)$, so $I \in \mathcal{D}(\mathbf{1}_I)$ and $I \subseteq \mathrm{sat}(\mathbf{1}_I)$. Also, $I \subseteq \mathrm{span}(I)$.
- Consider some $b \in \mathrm{span}(I) \setminus I$.
- Then $I \cup \{b\} \in \mathcal{D}(\mathbf{1}_I)$ since $\underbrace{\mathbf{1}_I(I \cup \{b\})}_{1_I(b)=0} = |I| = r(I \cup \{b\}) = r(I)$.

...

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

Proof.

- For $\mathbf{1}_I(I) = |I| = r(I)$, so $I \in \mathcal{D}(\mathbf{1}_I)$ and $I \subseteq \mathrm{sat}(\mathbf{1}_I)$. Also, $I \subseteq \mathrm{span}(I)$.
- Consider some $b \in \mathrm{span}(I) \setminus I$.
- Then $I \cup \{b\} \in \mathcal{D}(\mathbf{1}_I)$ since $\mathbf{1}_I(I \cup \{b\}) = |I| = r(I \cup \{b\}) = r(I)$.
- Thus, $b \in \mathrm{sat}(\mathbf{1}_I)$.

...

The sat function = Polymatroid Closure

Lemma 18.7.2 (Matroid $\mathrm{sat} : \mathbb{R}_+^E \to 2^E$ is the same as closure.)

$$\textit{For } I \in \mathcal{I}, \textit{ we have } \mathrm{sat}(\mathbf{1}_I) = \mathrm{span}(I) \tag{18.37}$$

Proof.

- For $\mathbf{1}_I(I) = |I| = r(I)$, so $I \in \mathcal{D}(\mathbf{1}_I)$ and $I \subseteq \mathrm{sat}(\mathbf{1}_I)$. Also, $I \subseteq \mathrm{span}(I)$.
- Consider some $b \in \mathrm{span}(I) \setminus I$.
- Then $I \cup \{b\} \in \mathcal{D}(\mathbf{1}_I)$ since $\mathbf{1}_I(I \cup \{b\}) = |I| = r(I \cup \{b\}) = r(I)$.
- Thus, $b \in \mathrm{sat}(\mathbf{1}_I)$.
- Therefore, $\mathrm{sat}(\mathbf{1}_I) \supseteq \mathrm{span}(I)$.

...

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \mathrm{sat}(\mathbf{1}_I) \setminus I$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \mathrm{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \mathrm{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.

$\square$

The sat function = Polymatroid Closure

. . . proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.
- Also, $r(A \cap I) = |A \cap I|$ since $A \cap I \in \mathcal{I}$.

$\square$

The sat function = Polymatroid Closure

. . . proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.
- Also, $r(A \cap I) = |A \cap I|$ since $A \cap I \in \mathcal{I}$.
- Hence, $r(A \cap I) = r(A) = r((A \cap I) \cup (A \setminus I))$ meaning $(A \setminus I) \subseteq \text{span}(A \cap I) \subseteq \text{span}(I)$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.
- Also, $r(A \cap I) = |A \cap I|$ since $A \cap I \in \mathcal{I}$.
- Hence, $r(A \cap I) = r(A) = r((A \cap I) \cup (A \setminus I))$ meaning $(A \setminus I) \subseteq \text{span}(A \cap I) \subseteq \text{span}(I)$.
- Since $b \in A \setminus I$, we get $b \in \text{span}(I)$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.
- Also, $r(A \cap I) = |A \cap I|$ since $A \cap I \in \mathcal{I}$.
- Hence, $r(A \cap I) = r(A) = r((A \cap I) \cup (A \setminus I))$ meaning $(A \setminus I) \subseteq \text{span}(A \cap I) \subseteq \text{span}(I)$.
- Since $b \in A \setminus I$, we get $b \in \text{span}(I)$.
- Thus, $\text{sat}(\mathbf{1}_I) \subseteq \text{span}(I)$.

$\square$

The sat function = Polymatroid Closure

...proof continued.

- Now, consider $b \in \text{sat}(\mathbf{1}_I) \setminus I$.
- Choose any $A \in \mathcal{D}(\mathbf{1}_I)$ with $b \in A$, thus $b \in A \setminus I$.
- Then $\mathbf{1}_I(A) = |A \cap I| = r(A) = r(A \cap I)$.
- Now $r(A) = |A \cap I| \leq |I| = r(I)$.
- Also, $r(A \cap I) = |A \cap I|$ since $A \cap I \in \mathcal{I}$.
- Hence, $r(A \cap I) = r(A) = r((A \cap I) \cup (A \setminus I))$ meaning $(A \setminus I) \subseteq \text{span}(A \cap I) \subseteq \text{span}(I)$.
- Since $b \in A \setminus I$, we get $b \in \text{span}(I)$.
- Thus, $\text{sat}(\mathbf{1}_I) \subseteq \text{span}(I)$.
- Hence $\text{sat}(\mathbf{1}_I) = \text{span}(I)$

$\square$

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$.

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$?

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$? No, it is not a vertex, or even a member, of P_r.

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$? No, it is not a vertex, or even a member, of P_r.
- $\mathrm{span}(\cdot)$ operates on more than just independent sets, so $\mathrm{span}(C)$ is perfectly sensible.

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$? No, it is not a vertex, or even a member, of P_r.
- $\text{span}(\cdot)$ operates on more than just independent sets, so $\text{span}(C)$ is perfectly sensible.
- Note $\text{span}(C) = \text{span}(B)$ where $\mathcal{I} \ni B \in \mathcal{B}(C)$ is a base of C.

The sat function = Polymatroid Closure

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$? No, it is not a vertex, or even a member, of P_r.
- $\mathrm{span}(\cdot)$ operates on more than just independent sets, so $\mathrm{span}(C)$ is perfectly sensible.
- Note $\mathrm{span}(C) = \mathrm{span}(B)$ where $\mathcal{I} \ni B \in \mathcal{B}(C)$ is a base of C.
- Then we have $\mathbf{1}_B \leq \mathbf{1}_C \leq \mathbf{1}_{\mathrm{span}(C)}$, and that $\mathbf{1}_B \in P_r$. We can then make the definition:

$$\mathrm{sat}(\mathbf{1}_C) \triangleq \mathrm{sat}(\mathbf{1}_B) \text{ for } B \in \mathcal{B}(C) \tag{18.38}$$

In which case, we also get $\mathrm{sat}(\mathbf{1}_C) = \mathrm{span}(C)$ (in general, could define $\mathrm{sat}(y) = \mathrm{sat}(\text{P-basis}(y))$).

The sat function = Polymatroid Closure

Skip

- Now, consider a matroid (E, r) and some $C \subseteq E$ with $C \notin \mathcal{I}$, and consider $\mathbf{1}_C$. Is $\mathbf{1}_C \in P_r$? No, it is not a vertex, or even a member, of P_r.
- $\text{span}(\cdot)$ operates on more than just independent sets, so $\text{span}(C)$ is perfectly sensible.
- Note $\text{span}(C) = \text{span}(B)$ where $\mathcal{I} \ni B \in \mathcal{B}(C)$ is a base of C.
- Then we have $\mathbf{1}_B \leq \mathbf{1}_C \leq \mathbf{1}_{\text{span}(C)}$, and that $\mathbf{1}_B \in P_r$. We can then make the definition:

$$\text{sat}(\mathbf{1}_C) \triangleq \text{sat}(\mathbf{1}_B) \text{ for } B \in \mathcal{B}(C) \tag{18.38}$$

  In which case, we also get $\text{sat}(\mathbf{1}_C) = \text{span}(C)$ (in general, could define $\text{sat}(y) = \text{sat}(\text{P-basis}(y))$).
- However, consider the following form

$$\text{sat}(\mathbf{1}_C) = \bigcup \{A : A \subseteq E, |A \cap C| = r(A)\} \tag{18.39}$$

Exercise: is $\text{span}(C) = \text{sat}(\mathbf{1}_C)$? Prove or disprove it.

The sat function, span, and submodular function minimization

- Thus, for a matroid, $\mathrm{sat}(\mathbf{1}_I)$ is exactly the closure (or span) of I in the matroid. I.e., for matroid (E, r), we have $\mathrm{span}(I) = \mathrm{sat}(\mathbf{1}_B)$.

The sat function, span, and submodular function minimization

- Thus, for a matroid, $\mathrm{sat}(\mathbf{1}_I)$ is exactly the closure (or span) of I in the matroid. I.e., for matroid (E, r), we have $\mathrm{span}(I) = \mathrm{sat}(\mathbf{1}_B)$.
- Recall, for $x \in P_f$ and polymatroidal f, $\mathrm{sat}(x)$ is the maximal (by inclusion) minimizer of $f(A) - x(A)$, and thus in a matroid, span(I) is the maximal minimizer of the submodular function formed by $r(A) - \mathbf{1}_I(A)$.

The sat function, span, and submodular function minimization

- Thus, for a matroid, $\mathrm{sat}(\mathbf{1}_I)$ is exactly the closure (or span) of I in the matroid. I.e., for matroid (E, r), we have $\mathrm{span}(I) = \mathrm{sat}(\mathbf{1}_B)$.
- Recall, for $x \in P_f$ and polymatroidal f, $\mathrm{sat}(x)$ is the maximal (by inclusion) minimizer of $f(A) - x(A)$, and thus in a matroid, span(I) is the maximal minimizer of the submodular function formed by $r(A) - \mathbf{1}_I(A)$.
- Submodular function minimization can solve "span" queries in a matroid or "sat" queries (e.g., $\min_{A \subseteq V} f(A) - x(A)$) in a polymatroid.

sat, as tight polymatroidal elements

- We are given an $x \in P_f^+$ for submodular function f.

sat, as tight polymatroidal elements

- We are given an $x \in P_f^+$ for submodular function f.
- Recall that for such an x, $\text{sat}(x)$ is defined as

$$\text{sat}(x) = \bigcup \{A : x(A) = f(A)\} \tag{18.40}$$

sat, as tight polymatroidal elements

- We are given an $x \in P_f^+$ for submodular function f.
- Recall that for such an x, $\text{sat}(x)$ is defined as

$$\text{sat}(x) = \bigcup \{A : x(A) = f(A)\} \tag{18.40}$$

- We also have stated that $\text{sat}(x)$ can be defined as:

$$\text{sat}(x) = \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.41}$$

sat, as tight polymatroidal elements

- We are given an $x \in P_f^+$ for submodular function f.
- Recall that for such an x, $\text{sat}(x)$ is defined as

$$\text{sat}(x) = \bigcup \{A : x(A) = f(A)\} \tag{18.40}$$

- We also have stated that $\text{sat}(x)$ can be defined as:

$$\text{sat}(x) = \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.41}$$

- We next show more formally that these are the same.

sat, as tight polymatroidal elements

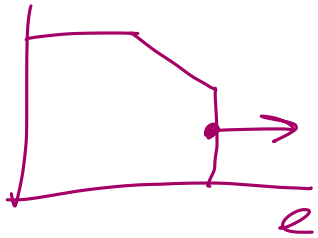
- Lets start with one definition and derive the other.

$$\mathrm{sat}(x)$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\text{sat}(x) \stackrel{\text{def}}{=} \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.42}$$



sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\text{sat}(x) \stackrel{\text{def}}{=} \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.42}$$

$$= \{ e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A) \} \tag{18.43}$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\text{sat}(x) \stackrel{\text{def}}{=} \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.42}$$

$$= \{ e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A) \} \tag{18.43}$$

$$= \{ e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A) \} \tag{18.44}$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\mathrm{sat}(x) \stackrel{\text{def}}{=} \left\{ e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+ \right\} \tag{18.42}$$

$$= \{ e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A) \} \tag{18.43}$$

$$= \{ e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A) \} \tag{18.44}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$.

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\mathrm{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \qquad (18.45)$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\mathrm{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \qquad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\mathrm{sat}(x)$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\mathrm{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \quad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) = f(A)\} \quad (18.46)$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\mathrm{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \quad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\begin{aligned}\mathrm{sat}(x) &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.46)\\ &= \{e : \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.47)\end{aligned}$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\text{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\text{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \quad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\begin{aligned}\text{sat}(x) &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.46)\\ &= \{e : \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.47)\end{aligned}$$

- So now, if A is any set such that $x(A) = f(A)$, then we clearly have

(18.48)

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\text{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\text{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \quad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\begin{aligned}\text{sat}(x) &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.46)\\ &= \{e : \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.47)\end{aligned}$$

- So now, if A is any set such that $x(A) = f(A)$, then we clearly have

$$\forall e \in A, e \in \text{sat}(x), \quad (18.48)$$

sat, as tight polymatroidal elements

- Lets start with one definition and derive the other.

$$\begin{aligned}\mathrm{sat}(x) &\stackrel{\text{def}}{=} \left\{e : \forall \alpha > 0, x + \alpha \mathbf{1}_e \notin P_f^+\right\} && (18.42)\\ &= \{e : \forall \alpha > 0, \exists A \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.43)\\ &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } (x + \alpha \mathbf{1}_e)(A) > f(A)\} && (18.44)\end{aligned}$$

- this last bit follows since $\mathbf{1}_e(A) = 1 \iff e \in A$. Continuing, we get

$$\mathrm{sat}(x) = \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) + \alpha > f(A)\} \quad (18.45)$$

- given that $x \in P_f^+$, meaning $x(A) \leq f(A)$ for all A, we must have

$$\begin{aligned}\mathrm{sat}(x) &= \{e : \forall \alpha > 0, \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.46)\\ &= \{e : \exists A \ni e \text{ s.t. } x(A) = f(A)\} && (18.47)\end{aligned}$$

- So now, if A is any set such that $x(A) = f(A)$, then we clearly have

$$\forall e \in A, e \in \mathrm{sat}(x), \text{ and therefore that } \mathrm{sat}(x) \supseteq A \quad (18.48)$$

sat, as tight polymatroidal elements

- ...and therefore, with sat as defined in Eq. (18.32),

$$\mathrm{sat}(x) \supseteq \bigcup \{A : x(A) = f(A)\} \tag{18.49}$$

sat, as tight polymatroidal elements

- ...and therefore, with sat as defined in Eq. (18.32),

$$\mathrm{sat}(x) \supseteq \bigcup \{A : x(A) = f(A)\} \tag{18.49}$$

- On the other hand, for any $e \in \mathrm{sat}(x)$ defined as in Eq. (18.47), since e is itself a member of a tight set, there is a set $A \ni e$ such that $x(A) = f(A)$, giving

$$\mathrm{sat}(x) \subseteq \bigcup \{A : x(A) = f(A)\} \tag{18.50}$$

sat, as tight polymatroidal elements

- ...and therefore, with sat as defined in Eq. (18.32),

$$\mathrm{sat}(x) \supseteq \bigcup \{A : x(A) = f(A)\} \tag{18.49}$$

- On the other hand, for any $e \in \mathrm{sat}(x)$ defined as in Eq. (18.47), since e is itself a member of a tight set, there is a set $A \ni e$ such that $x(A) = f(A)$, giving

$$\mathrm{sat}(x) \subseteq \bigcup \{A : x(A) = f(A)\} \tag{18.50}$$

- Therefore, the two definitions of sat are identical.

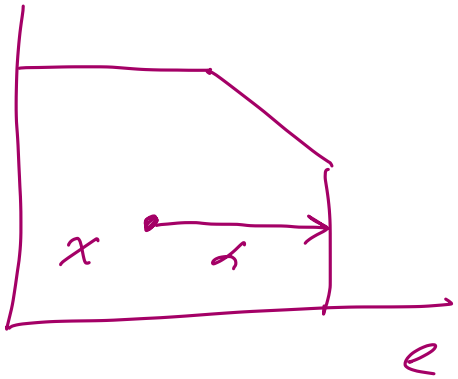
Saturation Capacity

- Another useful concept is saturation capacity which we develop next.

Saturation Capacity

- Another useful concept is saturation capacity which we develop next.
- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.51}$$



Saturation Capacity

- Another useful concept is saturation capacity which we develop next.
- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.51}$$

- This is identical to:

$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{18.52}$$

since any $B \subseteq E$ such that $e \notin B$ does not change in a $\mathbf{1}_e$ adjustment, meaning $(x + \alpha \mathbf{1}_e)(B) = x(B)$.

Saturation Capacity

- Another useful concept is saturation capacity which we develop next.
- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.51}$$

- This is identical to:

$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{18.52}$$

  since any $B \subseteq E$ such that $e \notin B$ does not change in a $\mathbf{1}_e$ adjustment, meaning $(x + \alpha \mathbf{1}_e)(B) = x(B)$.
- Again, this is identical to:

$$\max \{\alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\}\} \tag{18.53}$$

Saturation Capacity

- Another useful concept is saturation capacity which we develop next.
- For $x \in P_f$, and $e \in E$, consider finding

$$\max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.51}$$

- This is identical to:

$$\max \{\alpha : (x + \alpha \mathbf{1}_e)(A) \leq f(A), \forall A \supseteq \{e\}\} \tag{18.52}$$

since any $B \subseteq E$ such that $e \notin B$ does not change in a $\mathbf{1}_e$ adjustment, meaning $(x + \alpha \mathbf{1}_e)(B) = x(B)$.

- Again, this is identical to:

$$\max \{\alpha : x(A) + \alpha \leq f(A), \forall A \supseteq \{e\}\} \tag{18.53}$$

or

$$\max \{\alpha : \alpha \leq f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.54}$$

Saturation Capacity

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

Saturation Capacity

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.

Saturation Capacity

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.
- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \ni e\} \tag{18.56}$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.57}$$

Saturation Capacity

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.
- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \ni e\} \tag{18.56}$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.57}$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.

Saturation Capacity

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.
- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \ni e\} \tag{18.56}$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.57}$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, we have that: $e \in \text{sat}(x) \Leftrightarrow \hat{c}(x; e) = 0$.

Saturation Capacity

- The max is achieved when

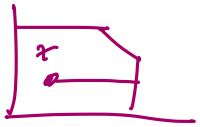
$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.
- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \ni e\} \tag{18.56}$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.57}$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, we have that: $e \in \text{sat}(x) \Leftrightarrow \hat{c}(x; e) = 0$.
- Note that any α with $0 \leq \alpha \leq \hat{c}(x; e)$ we have $x + \alpha \mathbf{1}_e \in P_f$.



Saturation Capacity

tool

- The max is achieved when

$$\alpha = \hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \supseteq \{e\}\} \tag{18.55}$$

- $\hat{c}(x; e)$ is known as the saturation capacity associated with $x \in P_f$ and e.
- Thus we have for $x \in P_f$,

$$\hat{c}(x; e) \stackrel{\text{def}}{=} \min \{f(A) - x(A), \forall A \ni e\} \tag{18.56}$$

$$= \max \{\alpha : \alpha \in \mathbb{R}, x + \alpha \mathbf{1}_e \in P_f\} \tag{18.57}$$

- We immediately see that for $e \in E \setminus \text{sat}(x)$, we have that $\hat{c}(x; e) > 0$.
- Also, we have that: $e \in \text{sat}(x) \Leftrightarrow \hat{c}(x; e) = 0$.
- Note that any α with $0 \leq \alpha \leq \hat{c}(x; e)$ we have $x + \alpha \mathbf{1}_e \in P_f$.
- We also see that computing $\hat{c}(x; e)$ is a form of submodular function minimization.